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

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

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

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). Necessary and sufficient for f to be measurable are

(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 $x \in \mathbb{R}^*$.

Note I 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 $A_1, A_2, A_3, \ldots \in \mathcal{A}$, 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 $T_1, T_2, T_3, \ldots$ of the topology $\mathcal{T}$, and the Borel sets are the elements of the smallest σ-algebra generated by $\mathcal{T}$, $\sigma(\mathcal{T})$.

Note II 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

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

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

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

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

3. INTEGRATION

DEFINITION

Simple Functions Let $(\Omega, \mathcal{E}, \mu)$ be a measure space. A *simple function*, ψ, is a set function defined on elements ω of sample space Ω by

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

for real constants $a_1, ..., a_k$ and measurable sets $A_1, ..., A_k$, for some $k = 1, 2, 3, ...$, where $I_A(\omega)$ is the *indicator function*, where

$$I_A(\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A \end{cases}.$$

Note that any such simple function, can be re-expressed as a simple function defined for a **partition** of Ω, $E_1, ..., E_l$,

$$\psi(\omega) = \sum_{i=1}^{l} e_i I_{E_i}(\omega)$$

by suitable choice of the constants $e_1, ..., e_k$.
Let ψ be a **non-negative** simple function, $\psi : \Omega \longrightarrow \mathbb{R}^\star$,

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

for real constants $a_1, ..., a_k \geq 0$ and measurable sets $A_1, ..., A_k \in \mathcal{E}$, for some $k = 1, 2, 3, ...$

(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

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

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 $E \subset \Omega$, if $I_E(\omega)$ is the indicator function for set E, then we can also define

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

(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 $0 \le \psi_n(\omega) \le f(\omega)$, 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.

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

4. RESULTS FOR MEASURABLE FUNCTIONS

Theorem 1. ***MEASURABILITY UNDER COMPOSITION***
Let g_1 and g_2 be measurable functions on $E \subset \Omega$ with ranges in $\mathbb{R}^*$. Let f be a Borel function from $\mathbb{R}^* \times \mathbb{R}^*$ into $\mathbb{R}^*$. Then the composite function h, defined on E by

$$h(\omega) = f(g_1(\omega_1), g_2(\omega_2))$$

is measurable.

Proof. The function $g = (g_1, g_2)$ has domain E and range $\mathbb{R}^* \times \mathbb{R}^*$, and is measurable as g_1 and g_2 are measurable, and denote $h = f \circ g$ (the operator $\circ$ indicates composition, i.e.

$$h(\omega_1, \omega_2) = (f \circ g)(\omega_1, \omega_2) \qquad \text{if} \qquad h(\omega_1, \omega_2) = f(g(\omega_1, \omega_2)) = f(g_1(\omega_1), g_2(\omega_2)).$$

If $B \in \mathcal{B}$, then $f^{-1}(B)$ is a Borel set as f is a Borel function. Thus the inverse image under h,

$$h^{-1}(B) = g^{-1}\left(f^{-1}(B)\right)$$

is measurable as g_1 and g_2, and hence g, are measurable. ∎

Corollary 2. *If g is a measurable function from E into $\mathbb{R}^*$, and f is a continuous function from $\mathbb{R}^*$ into $\mathbb{R}^*$, then $h = f \circ g$ is measurable.*

Theorem 3. ***MEASURABILITY UNDER ELEMENTARY OPERATIONS***
Let g_1 and g_2 be measurable functions defined on $E \subset \Omega$ into $\mathbb{R}^*$, and let c be any real number. Then all of the following composite and other related functions are measurable

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

Proof. In each case, we examine the domain of the composite function to ensure measurability in the Borel σ-algebra. Consider $g_1 + g_2$; this is not defined on the set $\{\omega : g_1(\omega) = -g_2(\omega) = \pm\infty\}$ (as $\infty \pm \infty$ is not defined), but this set is measurable, and so is the domain of $g_1 + g_2$. Let $f(x_1, x_2) = x_1 + x_2$ be a continuous function defined on $\mathbb{R}^* \times \mathbb{R}^*$. Then, by Theorem 1 and its corollary, $g_1 + g_2$ is measurable. Taking $g_2 = c$ proves that $g_1 + c$ is measurable.

The function $g_1 g_2$ is defined everywhere on E; it's measurability follows from Theorem 1, setting $f(x_1, x_2) = x_1 x_2$. Setting $g_2 = c$ proves that $c g_1$ is measurable.

The function g_1/g_2 is defined everywhere except on the union of sets

$$\{\omega : g_1(\omega) = g_2(\omega) = 0\} \cup \{\omega : \pm g_1(\omega) = \pm g_2(\omega) = \infty\}$$

Similarly, if $c = 0$, $|g_1|^c$ is defined except on $\{\omega : g_1(\omega) = \pm\infty\}$; if $c < 0$, it is defined except on $\{\omega : g_1(\omega) = 0\}$. If $c > 0$, it is defined everywhere. All of these sets are measurable Thus, we consider in turn functions

$$f(x_1, x_2) = x_1/x_2 \qquad f(x) = x^c$$

and use Theorem 1.

The functions $g_1 \vee g_2, g_1 \wedge g_2$ are defined everywhere; so we consider functions

$$f(x_1, x_2) = \max\{x_1, x_2\} \qquad f(x_1, x_2) = \min\{x_1, x_2\}$$

and again use Theorem 1. Finally, setting $g_2 = 0$ yields the measurability of g_1^+ and g_1^-. ∎

Theorem 4. *If g_1 and g_2 are measurable functions on a common domain, then each of the sets*

$$\{\omega : g_1(\omega) < g_2(\omega)\} \qquad \{\omega : g_1(\omega) = g_2(\omega)\} \qquad \{\omega : g_1(\omega) > g_2(\omega)\}$$

is measurable.

Proof. Since g_1 and g_2 are measurable, then $f = g_1 - g_2$ is measurable, and thus the two sets

$$\{\omega : f(\omega) > 0\} \qquad \{\omega : f(\omega) = 0\}$$

are measurable. Since

$$\{\omega : g_1(\omega) < g_2(\omega)\} \equiv \{\omega : f(\omega) > 0\}$$

and

$$\{\omega : g_1(\omega) = g_2(\omega)\} \equiv \{\omega : f(\omega) = 0\} \cup \{\omega : g_1(\omega) = g_2(\omega) = \pm\infty\}$$

then $\{\omega : g_1(\omega) < g_2(\omega)\}$ and $\{\omega : g_1(\omega) = g_2(\omega)\}$ are measurable, and so is

$$\{\omega : g_1(\omega) \leq g_2(\omega)\} \equiv \{\omega : g_1(\omega) < g_2(\omega)\} \cup \{\omega : g_1(\omega) = g_2(\omega)\}.$$

∎

Theorem 5. ***MEASURABILITY UNDER LIMIT OPERATIONS***
If $\{g_n\}$ is a sequence of measurable functions, the functions $\sup_n g_n$ and $\inf_n g_n$ are measurable.

Proof. Let $g = \sup_n g_n$. Then for real x, consider

$$g_n^{-1}([-\infty, x]) \equiv \{\omega : g_n(\omega) \leq x\}$$

and

$$g^{-1}([-\infty, x]) \equiv \{\omega : g(\omega) \leq x\}.$$

If $g = \sup_n g_n$, then $g_n \leq g$ for all n, and

$$g(\omega) \leq x \Longrightarrow g_n(\omega) \leq x \qquad \text{so that} \qquad \omega \in g^{-1}([-\infty, x]) \Longrightarrow \omega \in g_n^{-1}([-\infty, x])$$

so that

$$g^{-1}([-\infty, x]) \subseteq g_n^{-1}([-\infty, x])$$

for all n. Thus, in fact

$$g^{-1}([-\infty, x]) = \bigcap_n g_n^{-1}([-\infty, x])$$

and hence g is measurable, as the intersection of measurable sets is measurable. The result for $\inf_n$ follows by noting that

$$\inf_n g_n = -\sup_n(-g_n).$$

∎

Theorem 6. ***MEASURABILITY UNDER LIMINF/LIMSUP***
If $\{g_n\}$ is a sequence of measurable functions, the functions $\limsup_n g_n$ and $\liminf_n g_n$ are measurable.

Proof. This follows from Theorem 5, as

$$\limsup_n g_n = \inf_k \left\{\sup_{n \geq k} g_n\right\} \qquad \text{and} \qquad \liminf_n g_n = \sup_k \left\{\inf_{n \geq k} g_n\right\}$$

∎

5. SIMPLE FUNCTIONS AND THEIR CONVERGENCE PROPERTIES.

Theorem 7. *A non-negative function on Ω is measurable if and only if it is the limit of an increasing sequence of non-negative simple functions.*

Proof. Suppose that g is a nonnegative measurable function. For each positive integer n, define the simple function ψ_n on Ω by

$$\psi_n(\omega) = \frac{m}{2^n} \qquad \text{if } \frac{m}{2^n} \le g(\omega) < \frac{m+1}{2^n}$$

for $m = 0, 1, 2, ..., 2^n - 1$, and

$$\psi_n(\omega) = n \qquad \text{if } n \le g(\omega).$$

Then $\{\psi_n\}$ is an increasing sequence of non-negative simple functions. Since

$$|\psi_n(\omega) - g(\omega)| < \frac{1}{2^n} \qquad \text{if } n > g(\omega)$$

and $\psi_n(\omega) = n$ if $g(\omega) = \infty$, then, for all ω,

$$\psi_n(\omega) \to g(\omega)$$

and we have found the sequence required for the result.

Now suppose that g is a limit of an increasing sequence of non-negative simple functions; it is measurable by Theorem 6. ∎

Theorem 8. *A function g defined on Ω is measurable if and only if it is the limit of a sequence of simple functions.*

Proof. Suppose that g is measurable. Then g^+ and g^- are measurable and non-negative, and thus can be represented as limits of simple functions $\{\psi_n^+\}$ and $\{\psi_n^-\}$, by the Theorem 7. Consider the sequence of simple functions defined by $\{\psi_n^+ - \psi_n^-\}$; this sequence converges to $g^+ - g^- = g$, and we have the sequence of simple functions required for the result.

Now suppose that g is a limit of a sequence of simple functions; it is measurable by Theorem 6. ∎

6. KEY THEOREMS

The following key theorems describe the behaviour of the Lebesgue-Stieltjes integral. In particular, the theorems specify when it is legitimate to exchange the order of limit and integral operators. In the theorems, we have a general measure space $(\Omega, \mathcal{F}, \nu)$, and measurable set $E \in \mathcal{F}$.

Theorem 9. ***Lebesgue Monotone Convergence Theorem***
If $\{f_n\}$ is an increasing sequence of nonnegative measurable functions, and if

$$\lim_{n\to\infty} f_n = f \qquad \textit{almost everywhere}$$

then

$$\lim_{n\to\infty} \int_E f_n d\nu = \int_E f d\nu$$

Proof. Let the (real) sequence $\{i_n\}$ be defined by

$$i_n = \int_E f_n d\nu.$$

Then, by a previous result

$$i_n = \int_E f_n d\nu \leq \int_E f_{n+1} d\nu = i_{n+1} \qquad \text{as } f_n \leq f_{n+1}$$

so $\{i_n\}$ is increasing. Let L denote the (possibly infinite) limit of $\{i_n\}$. Now, since $f_n \leq f$ almost everywhere for all n, we have (by the same previous result) that

$$\int_E f_n d\nu \leq \int_E f d\nu \Longrightarrow L \leq \int_E f d\nu. \tag{1}$$

Now consider constant c with $0 < c < 1$, and let ψ be any simple function satisfying $0 \leq \psi \leq f$. Let

$$E_n \equiv \{\omega : \omega \in E \text{ and } c\psi(\omega) \leq f_n(\omega)\}$$

and as $E_n \subseteq E$, E_n is measurable, and because $f_n \leq f_{n+1}$, $E_n \subseteq E_{n+1}$ for all n, so $\{E_n\}$ is increasing. Let the limit of the $\{E_n\}$ sequence be denoted

$$F = \bigcup_{i=1}^{\infty} E_n.$$

The set $E \cap F'$ has measure zero, because $\lim\limits_{n\to\infty} f_n = f$ a.e. and $0 \leq c\psi < \psi \leq f$. Hence, as $E_n \subseteq E$

$$\int_E f_n d\nu \geq \int_{E_n} f_n d\nu \geq \int_{E_n} c\psi d\nu = c\int_{E_n} \psi d\nu.$$

Taking the limit as $n \to \infty$,

$$L = \lim_{n\to\infty} \int_E f_n d\nu \geq c \lim_{n\to\infty} \int_{E_n} \psi d\nu = c\int_F \psi d\nu = c\int_E \psi d\nu$$

the final step following as $E \cap F'$ has measure zero. Thus, as this holds for all c such that $0 < c < 1$, we must have that

$$L \geq \int_E \psi d\nu$$

whenever $0 \leq \psi \leq f$. Hence L is an upper bound the integral of such a simple function on E. But, by the supremum definition from lectures, the integral of f with respect to ν on E is the **least** upper bound on the integral of such simple functions on E. Hence

$$L \geq \int_E f d\nu. \tag{2}$$

Thus, combining (1) and (2), we have that

$$L = \lim_{n\to\infty} \int_E f_n d\nu = \int_E f d\nu.$$

∎

Theorem 10. ***Fatou's Lemma*** *(or* ***Lebesgue-Fatou Theorem****)*
If $\{f_n\}$ is a sequence of non-negative measurable functions, and if

$$\liminf_{n\to\infty} f_n = f \qquad \textit{almost everywhere}$$

then

$$\int_E f d\nu \leq \liminf_{n\to\infty} \left\{ \int_E f_n d\nu \right\}$$

Proof. The function $\liminf_{n\to\infty} f_n$ is measurable. For $k = 1, 2, 3, ...$ let

$$h_k = \inf \{f_n : n \geq k\}.$$

Then, by definition of infimum, $h_k \leq f_k$ for all k, and thus

$$\int_E h_k d\nu \leq \int_E f_k d\nu \quad \text{for all } k \qquad \Longrightarrow \qquad \liminf_{k\to\infty} \left\{ \int_E h_k d\nu \right\} \leq \liminf_{k\to\infty} \left\{ \int_E f_k d\nu \right\}. \tag{3}$$

Now $\{h_k\}$ is an increasing sequence of non-negative functions, we have in the limit

$$\lim_{k\to\infty} h_k = \liminf_{n\to\infty} f_n = f$$

almost everywhere. Now, by the Monotone Convergence Theorem,

$$\lim_{k\to\infty} \left\{ \int_E h_k d\nu \right\} = \int_E \left\{ \lim_{k\to\infty} h_k \right\} d\nu = \int_E f d\nu$$

Hence, by (3),

$$\int_E f d\nu \leq \liminf_{k\to\infty} \left\{ \int_E f_k d\nu \right\}.$$

∎

Some corollaries follow immediately from this important theorem

1 If $E_1, E_2, ..., E_n$ are disjoint, with $\bigcup_{i=1}^{n} E_i \equiv E$, and f is non-negative, then

$$\int_E f d\nu = \sum_{i=1}^{n} \left\{ \int_{E_i} f d\nu \right\}$$

Proof: Let $\{\psi_k\}$ be an increasing sequence of simple functions that converge to f, where

$$\psi_k = \sum_{j=1}^{m_k} a_{kj} I_{A_{kj}}$$

say. Then,

$$\begin{aligned}
\int_E \psi_k d\nu &= \sum_{j=1}^{m_k} a_{kj} \nu(E \cap A_{kj}) = \sum_{j=1}^{m_k} \sum_{i=1}^{n} a_{kj} \nu(E_i \cap A_{kj}) \qquad \text{as the } E_i \text{ are disjoint} \\
&= \sum_{i=1}^{n} \left\{ \sum_{j=1}^{m_k} a_{kj} \nu(E_i \cap A_{kj}) \right\} = \sum_{i=1}^{n} \left\{ \int_{E_i} \psi_k d\nu \right\}
\end{aligned}$$

by hence the monotone convergence theorem,

$$\begin{aligned}\int_E f d\nu &= \lim_{k\to\infty}\left\{\int_E \psi_k d\nu\right\} = \lim_{k\to\infty}\left\{\sum_{i=1}^n\left\{\int_{E_i}\psi_k d\nu\right\}\right\} = \sum_{i=1}^n\left\{\lim_{k\to\infty}\left\{\int_{E_i}\psi_k d\nu\right\}\right\}\\ &= \sum_{i=1}^n\left\{\int_{E_i}\left\{\lim_{k\to\infty}\psi_k\right\}d\nu\right\} = \sum_{i=1}^n\left\{\int_{E_i} f d\nu\right\}.\end{aligned}$$

2 Now consider a **countable** (rather than merely finite) collection $\{E_i\}$ with $\bigcup\limits_{i=1}^{\infty} E_i \equiv E$. Then if f is non-negative

$$\int_E f d\nu = \sum_{i=1}^{\infty}\left\{\int_{E_i} f d\nu\right\}$$

Proof: For each positive integer n, let $A_n \equiv \bigcup\limits_{i=1}^{n} E_i$, and define $f_n = I_{A_n} f$. Then $\{f_n\}$ is an increasing sequence of non-negative functions, that converges to f (on E). Hence

$$\int_E f d\nu = \lim_{n\to\infty}\left\{\int_E f_n d\nu\right\} = \lim_{n\to\infty}\left\{\int_{A_n} f d\nu\right\} = \lim_{n\to\infty}\left\{\sum_{i=1}^n\left\{\int_{E_i} f d\nu\right\}\right\} = \sum_{i=1}^{\infty}\left\{\int_{E_i} f d\nu\right\}$$

3 Let f be a non-negative function on Ω. Then the function defined on $\mathcal{F}$ by

$$\varphi(E) = \int_E f d\nu$$

is a measure. The only part of the definition of a measure that needs verifying is the countable additivity, by the last result, we have directly that

$$\varphi\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty}\varphi(E_i)$$

when the $\{E_i\}$ are disjoint.

For the results above (and the results proved in lectures), we have considered only the integrals of non-negative measurable functions. We now extend them for general measurable functions, using the decomposition into positive and negative part functions $f = f^+ - f^-$ where both f^+ and f^- are measurable and non-negative, and we have

$$\int_E f d\nu = \int_E f^+ d\nu - \int_E f^- d\nu.$$

Recall that we say that f is integrable if both f^+ and f^- are integrable, and now denote the set of all functions integrable on E with respect to ν by $\mathcal{L}_E(\nu)$. From previous arguments we have that

$$f \in \mathcal{L}_E(\nu) \Leftrightarrow f^+ \text{ and } f^- \in \mathcal{L}_E(\nu)$$

Some results can be proved for the functions in this class.

LEMMA
If $\nu(E) = 0$, then

$$f \in \mathcal{L}_E(\nu) \qquad \text{and} \qquad \int_E f d\nu = 0$$

Proof. We have by definition

$$\int_E f d\nu = \int_E f^+ d\nu - \int_E f^- d\nu = 0 - 0 = 0$$

∎

LEMMA
If $f \in \mathcal{L}_{E_2}(\nu)$ and $E_1 \subset E_2$, then $f \in \mathcal{L}_{E_1}(\nu)$.

Proof. By a result from lectures

$$\int_{E_1} f^+ d\nu \leq \int_{E_2} f^+ d\nu \qquad \text{and} \qquad \int_{E_1} f^- d\nu \leq \int_{E_2} f^- d\nu$$

∎

LEMMA
If $\{E_n\}$ is a sequence of disjoint sets with $\bigcup_{n=1}^{\infty} E_n \equiv E$, and $f \in \mathcal{L}_E(\nu)$, then

$$\int_E f d\nu = \sum_{n=1}^{\infty} \left\{ \int_{E_n} f d\nu \right\}$$

Proof. The previous Lemma ensures that $f \in \mathcal{L}_{E_n}(\nu)$ as $E_n \subset E$ for all n. By using the result proved earlier, that if f is non-negative then

$$\int_E f d\nu = \sum_{n=1}^{\infty} \left\{ \int_{E_n} f d\nu \right\}$$

we use the positive and negative part decompositions

$$\begin{aligned}
\int_E f d\nu &= \int_E f^+ d\nu - \int_E f^- d\nu = \sum_{n=1}^{\infty} \left\{ \int_{E_n} f^+ d\nu \right\} - \sum_{n=1}^{\infty} \left\{ \int_{E_n} f^+ d\nu \right\} \\
&= \sum_{n=1}^{\infty} \left[\int_{E_n} f^+ d\nu - \int_{E_n} f^- d\nu \right] = \sum_{n=1}^{\infty} \left\{ \int_{E_n} \left(f^+ - f^- \right) d\nu \right\} = \sum_{n=1}^{\infty} \left\{ \int_{E_n} f d\nu \right\}
\end{aligned}$$

∎

Corollary 11. *If $f \in \mathcal{L}_\Omega(\nu)$, then the function φ defined on $\mathcal{F}$ by*

$$\varphi(E) = \int_E f d\nu$$

is additive.

Proof. As for previous result. ∎

LEMMA
If $f = g$ a.e. on E, and if $g \in \mathcal{L}_E(\nu)$, then $f \in \mathcal{L}_E(\nu)$ and

$$\int_E f d\nu = \int_E g d\nu$$

Proof. Define $A \equiv \{\omega : \omega \in E, f(\omega) = g(\omega)\}$. Then $E \cap A'$ has measure zero, and

$$\int_E f^+ d\nu = \int_A f^+ d\nu = \int_A g^+ d\nu = \int_E g^+ d\nu$$

and

$$\int_E f^- d\nu = \int_A f^- d\nu = \int_A g^- d\nu = \int_E g^- d\nu$$

Adding these equations, we have immediately that $f \in \mathcal{L}_E(\nu)$ and

$$\int_E f d\nu = \int_E g d\nu$$

∎

LEMMA
If $f \in \mathcal{L}_E(\nu)$ and c is any real number, then $cf \in \mathcal{L}_E(\nu)$ and

$$\int_E (cf)\, d\nu = c \int_E f d\nu$$

Proof. Consider only the non-trivial case $c \neq 0$. Suppose first $c > 0$, and let g be a non-negative function. For any simple function ψ, say

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

we have

$$\psi \leq g \Leftrightarrow c\psi \leq cg.$$

and

$$\int_E (c\psi)\, d\nu = \sum_{i=1}^{k} (ca_i)\, \nu(E \cap A_i) = c \sum_{i=1}^{k} a_i \nu(E \cap A_i) = c \int_E \psi d\nu$$

Therefore

$$\int_E (cf)\, d\nu = c \int_E f d\nu$$

by the supremum definition, and the result follows for $c > 0$ using this result, and the decomposition $cf = cf^+ - cf^-$. For $c < 0$, write

$$cf = (-c)\, f^- - (-c)\, f^+$$

so that the result follows, as $-c > 0$. ∎

LEMMA
If $f, g \in \mathcal{L}_E(\nu)$, then $f+g \in \mathcal{L}_E(\nu)$ and

$$\int_E (f+g)\, d\nu = \int_E f d\nu + \int_E g d\nu$$

Proof. We prove the result two several stages. First suppose that f and g are non-negative, and let $\left\{\psi_n^{(f)}\right\}$ and $\left\{\psi_n^{(g)}\right\}$ be increasing sequences of simple functions with limits f and g respectively. Then $\left\{\psi_n^{(f)} + \psi_n^{(g)}\right\}$ has limit $f+g$, and as

$$\int_E \left(\psi_n^{(f)} + \psi_n^{(g)}\right) d\nu = \int_E \psi_n^{(f)} d\nu + \int_E \psi_n^{(f)} d\nu$$

(see this result by using the measure definition of the integral of a simple function), we have, taking the limit as $n \to \infty$,

$$\int_E (f+g)\, d\nu = \int_E f d\nu + \int_E g d\nu.$$

Now consider the general case; define the following subsets of E

$$\begin{aligned}
E_1 &\equiv \{\omega : f(\omega) \geq 0, g(\omega) \geq 0\} \\
E_2 &\equiv \{\omega : f(\omega) < 0, g(\omega) \geq 0\} \\
E_3 &\equiv \{\omega : f(\omega) \geq 0, g(\omega) < 0, (f+g)(\omega) \geq 0\} \\
E_4 &\equiv \{\omega : f(\omega) < 0, g(\omega) \geq 0, (f+g)(\omega) \geq 0\} \\
E_5 &\equiv \{\omega : f(\omega) \geq 0, g(\omega) < 0, (f+g)(\omega) < 0\} \\
E_6 &\equiv \{\omega : f(\omega) < 0, g(\omega) \geq 0, (f+g)(\omega) < 0\}
\end{aligned}$$

Then $E_n, n = 1, 2, ..., 6$ are disjoint, and $\bigcup\limits_{n=1}^{6} E_n \equiv E$. By the Lemma **??**, proving that

$$\int_{E_n} (f+g)\, d\nu = \int_{E_n} f d\nu + \int_{E_n} g d\nu$$

for each n is sufficient to prove the result. The proofs for each separate case are very similar; so consider for example set E_3. Then on E, the functions $f, -g$ and $f+g$ are non-negative, and threfore by part one of this proof,

$$\int_{E_3} f d\nu = \int_{E_3} (-g)\, d\nu + \int_{E_3} (f+g)\, d\nu = -\int_{E_3} g d\nu + \int_{E_3} (f+g)\, d\nu$$

and the result follows. ∎

LEMMA
The function $f \in \mathcal{L}_E(\nu)$ if and only if $|f| \in \mathcal{L}_E(\nu)$. In this instance,

$$\left|\int_E f d\nu\right| \leq \int_E |f|\, d\nu.$$

Proof. We have identified previously that f is integrable if the positive and negative part functions are integrable, and this is the case if and only if the function

$$|f| = f^+ + f^-$$

is integrable. If this is the case, then

$$\left|\int_E f d\nu\right| = \left|\int_E f^+ - f^- d\nu\right| \leq \left|\int_E f^+ d\nu\right| + \left|\int_E f^- d\nu\right| = \int_E |f| \, d\nu$$

∎

Corollary 12. *If $g \in \mathcal{L}_E(\nu)$, and $|f| \leq g$, then $f \in \mathcal{L}_E(\nu)$*

LEMMA
If $f, g \in \mathcal{L}_E(\nu)$, and $f \leq g$ a.e. on E, then

$$\int_E f d\nu \leq \int_E g d\nu$$

that is, the Lebesgue-Stieltjes Integral operator preserves ordering of functions.

Proof. We have $g - f \geq 0$, so the result follows from Integral Result (e) from lectures, and Lemma 6.. ∎

Corollary 13. *If $\upsilon(E) < \infty$, and $m \leq f \leq M$ on E, for real values m and M, then by considering simple functions $\psi_m = mI_E$ and $\psi_M = MI_E$, for which $\psi_m \leq f \leq \psi_M$, we have*

$$m\upsilon(E) \leq \int_E f d\nu \leq M\upsilon(E)$$

LEMMA
Suppose $f, g \in \mathcal{L}_E(\nu)$, and that for $A \subset E$,

$$\int_A f d\nu \leq \int_A g d\nu.$$

Then $f \leq g$ a.e. on E.

Proof. Let $F_1 \equiv \{\omega : \omega \in E, f(\omega) \geq g(\omega)\}$, so that $f - g \geq 0$ on F_1. Thus, by the assumption of the Lemma,

$$\int_F (f - g) \, d\nu = 0$$

and hence by $f - g = 0$ or $f = g$ a.e. on F_1, by Integral Result (f) from lectures. ∎

Corollary 14. *If $f, g \in \mathcal{L}_E(\nu)$ and if*

$$\int_A f d\nu = \int_A g d\nu.$$

for $A \subset E$, then $f = g$ a.e. on E.

Theorem 15. ***Lebesgue Dominated Convergence Theorem***
If $\{f_n\}$ is a sequence of measurable functions, and if

$$\lim_{n\to\infty} f_n = f \qquad \textit{almost everywhere}$$

and $|f_n| \leq g$ for all n, for some $g \in \mathcal{L}_E(\nu)$, then

$$\lim_{n\to\infty} \int_E f_n d\nu = \int_E f d\nu$$

Proof. $\{f_n\}$ and f are measurable functions. By using Fatou's Lemma (Theorem 10) on non-negative sequence $\{g+f_n\}$

$$\int_E (g+f)\,d\nu \leq \liminf_{n\to\infty}\left\{\int_E (g+f_n)\,d\nu\right\}$$

so that

$$\int_E f d\nu \leq \liminf_{n\to\infty}\left\{\int_E f_n d\nu\right\}. \tag{4}$$

Similarly, by applying the result to $\{g-f_n\}$, we have that

$$\int_E (g-f)\,d\nu \leq \liminf_{n\to\infty}\left\{\int_E (g-f_n)\,d\nu\right\} \qquad \therefore \qquad -\int_E f d\nu \leq \liminf_{n\to\infty}\left\{-\int_E f_n d\nu\right\}$$

Multiplying through by -1, we have by properties of $\limsup$ and $\liminf$ that

$$\int_E f d\nu \geq \limsup_{n\to\infty}\left\{\int_E f_n d\nu\right\} \tag{5}$$

and hence combining (4) and (5), we have by definition

$$\lim_{n\to\infty}\int_E f_n d\nu = \int_E f d\nu$$

∎

Corollary 16. *If $\{f_n\}$ is a uniformly bounded sequence (bounded above and below by a pair of real constants) of measurable functions such that*

$$\lim_{n\to\infty} f_n = f \qquad \textit{almost everywhere}$$

and if $\upsilon(E) < \infty$, then

$$\lim_{n\to\infty}\int_E f_n d\nu = \int_E f d\nu.$$

LEBESGUE-STIELTJES INTEGRALS ON $\mathbb{R}$.

Rather than considering a general sample space Ω, we now consider the specific case when $\Omega \equiv \mathbb{R}$, with corresponding sigma-algebra which is the Borel sigma-algebra. In this case, the measure υ will often be expressed in terms of (or be generated by) an increasing **real** function F on E. Let E be a set in the Borel sigma-algebra. Then for measurable function g, we can express the integral as

$$\int_E g d\nu = \int_E g dF \qquad \text{or} \qquad \int_E g d\nu = \int_E g(x) dF(x)$$

with special cases

$$\int_a^b g\;dF = \int_{(a,b]} g\;dF \qquad \text{and} \qquad \int_{-\infty}^{\infty} g\;dF = \int_{\mathbb{R}} g\;dF$$

7. MEASURE, INTEGRATION AND PROBABILITY DISTRIBUTIONS

In the measure-theoretic framework, **random variables** are merely measurable functions with respect to the probability space $(\Omega, \mathcal{E}, \mathbb{P})$, that is, for random variable $X = X(\omega)$, $\omega \in \Omega$, with domain E, the inverse image of Borel set B, $X^{-1}(B) \equiv \{\omega \in E : X(\omega) \in B\}$, is an element of σ-algebra $\mathcal{E}$. The **expectation** of X can be written in any of the following ways

$$\int_E X(\omega)\,\mathbb{P} \qquad \int_E X(\omega)\,\mathbb{P}(d\omega) \qquad \int_E X(\omega)\,d\mathbb{P}(\omega)$$

7.1 Lebesgue-Stieltjes Integration

If $\mathbb{P}$ is a probability measure on $\mathcal{B}$, then there is a unique corresponding real function F defined for $x \in \mathbb{R}$ by $F(x) = \mathbb{P}((-\infty, x])$, termed the **distribution function**. Conversely, if F is a distribution function, then F defines a measure μ_F on the Borel sets of $\mathbb{R}$, $\mathcal{B}$: we define μ_F on $\mathcal{B}$ via sets $(a, b])$ by

$$\mu_F((a,b]) = F(b) - F(a)$$

and then **extend** to $\mathcal{B}$ by using union operations. The probability space $(\mathbb{R}, \mathcal{B}, \mu_F)$ is then **completed** by considering and including null sets (under μ_F). Let $\mathcal{L}_F$ denote the smallest σ-algebra containing $\mathcal{B}$ and all μ_F-null sets. Thus the triple $(\mathbb{R}, \mathcal{L}_F, \mu_F)$ is the **completed** probability space.

If $g : \mathbb{R} \longrightarrow \mathbb{R}$ is a measurable function on the probability space $(\mathbb{R}, \mathcal{L}_F, \mu_F)$, then the **Lebesgue-Stieltjes integral** of g is

$$\int g\,d\mu_F = \int g(x)\,dF = \int g(x)\,dF(x).$$

7.2 The Radon-Nikodym Theorem and Change of Measure

- *σ-finite Measure :* a measure μ defined on a σ-algebra $\mathcal{C}$ of subsets of a set C is called **finite** if $\mu(C)$ is a **finite** real number. The measure μ is called σ-finite if C is the countable union of measurable sets of finite measure. A set in a measure space has σ-finite measure if it is a union of sets with finite measure.
- *Absolute Continuity :* If μ and ν are two measures on $(C, \mathcal{C})$, then ν is **absolutely continuous** with respect to μ, denoted $\nu \ll \mu$, if and only if for all $E \in \mathcal{C}$, $\mu(E) = 0 \Longrightarrow \nu(E) = 0$.
- *The Radon-Nikodym Theorem :* For a measure space $(C, \mathcal{C})$, if measure ν is absolutely continuous with respect to a σ-finite measure μ, then there exists a measurable function f defined on C and taking values in $[0, \infty)$, such that
$$\nu(A) = \int_A f\,d\mu$$
for any measurable set A. The function f is unique almost everywhere wrt μ; it is termed the **Radon-Nikodym Derivative** of **density** of ν with respect to μ, and is often expressed as
$$f = \frac{d\nu}{d\mu}$$
- *Change of Measure :* If μ_F is absolutely continuous wrt σ-finite measure μ, then we can rewrite
$$\int g\,d\mu_F = \int g\,\frac{d\mu_F}{d\mu}\,d\mu = \int gf\,d\mu$$
In practice, μ is either **counting** or **Lebesgue** measure.

7.3 Expectations

We can construct expectations for random variables using an identical method of construction as for integral with respect to measure; for a probability space $(\Omega, \mathcal{E}, \mathbb{P})$

1. A random variable $X : \Omega \longrightarrow \mathbb{R}$ is called **simple** if it only takes finitely many distinct values; simple random variables can be written
$$X = \sum_{i=1}^{n} x_i I_{A_i}$$
for some partition $A_1, \ldots, A_n$ of Ω. The **expectation** of X is
$$\mathbb{E}[X] = \sum_{i=1}^{n} x_i \mathbb{P}(A_i)$$

2. Any non-negative random variable $X : \Omega \longrightarrow [0, \infty)$ is the **limit of some increasing sequence of simple variables**, $\{X_n\}$. That is, $X_n(\omega) \uparrow X(\omega)$, for all $\omega \in \Omega$. The expectation of X is
$$\mathbb{E}[X] = \lim_{n \longrightarrow \infty} \mathbb{E}[X_n]$$
and the limit may be infinite.

3. Any random variable $X : \Omega \longrightarrow \mathbb{R}$ can be written as $X = X^+ - X^-$, where
$$X^+(\omega) = \max\{X(\omega), 0\} \qquad X^-(\omega) = \max\{-X(\omega), 0\} = -\min\{X(\omega), 0\}$$
The expectation of X is
$$\mathbb{E}[X] = \mathbb{E}[X^+] - \mathbb{E}[X^-]$$
if at least one of the two expectations on the right hand side is finite.

4. Thus the expectation of any random variable, $\mathbb{E}[X]$ is well-defined for every variable X such that
$$\mathbb{E}[|X|] = \mathbb{E}[X^+ + X^-] < \infty$$

Expectation defined in this fashion obeys the following rules: if $\{X_n\}$ is a sequence of rvs with $X_n(\omega) \longrightarrow X(\omega)$ for all $\omega \in \Omega$, then

(i) *Monotone Convergence:* If $X_n(\omega) \geq 0$ and $X_n(\omega) \leq X_{n+1}(\omega)$ for all n and ω, then
$$\mathbb{E}[X_n] \longrightarrow \mathbb{E}[X]$$

(ii) *Dominated Convergence:* If $|X_n(\omega)| \leq Y(\omega)$ for all n and ω, and $\mathbb{E}[|Y|] < \infty$, then
$$\mathbb{E}[X_n] \longrightarrow \mathbb{E}[X]$$

(iii) *Bounded Convergence:* If $|X_n(\omega)| \leq c$, for some c, and for all n and ω, then
$$\mathbb{E}[X_n] \longrightarrow \mathbb{E}[X]$$

These results hold even if $X_n(\omega) \longrightarrow X(\omega)$ for all ω except possibly ω in sets of probability zero (termed **null events**), that is, if $X_n(\omega) \longrightarrow X(\omega)$ **almost everywhere.**

One further result is of use in expectation calculations:

Fatou's Lemma : If $\{X_n\}$ is a sequence of rvs with $X_n(\omega) \geq Y(\omega)$ almost everywhere for all n and for some Y with $\mathbb{E}[Y] < \infty$, then
$$\mathbb{E}[\liminf_{n \longrightarrow \infty} X_n] \leq \liminf_{n \longrightarrow \infty} \mathbb{E}[X_n]$$