FIN 508- MARTINGALE THROUGH MEASURE THEORY
LECTURE 1

MEASURES AND INTEGRATION

The conventions about the extended real line $[-\infty, \infty]$ will be applied in these notes, where two symbols $-\infty$ and ∞ are added to $\mathbb{R}$, so that $[-\infty, \infty] = \{-\infty\} \cup \mathbb{R} \cup \{\infty\}$. For every $a \in \mathbb{R}$, $-\infty < a < \infty$,

$$a + \infty = \infty + a = \infty, \; a - \infty = -\infty + a = -\infty.$$

$$\frac{a}{-\infty} = \frac{a}{\infty} = 0,$$

but $\frac{\infty}{\infty}$, $\frac{a}{0}$, $\infty - \infty$, $\infty + (-\infty)$ and $(-\infty) + \infty$ are not defined, while $0 \cdot \infty = -\infty \cdot 0 = 0$, $-\infty + (-\infty) = -\infty$ and $\infty + \infty = \infty$.

Let us generalize the notions of measures and outer measures introduced in Part A Integration with modification, for our convenience for this course.

1. Measures. Let Ω be a (sample) space, and $\mathcal{R}$ be a collection of some subsets of Ω. Suppose $\mathcal{R}$ contains an empty set denoted by $\varnothing$. A function $\mu : \mathcal{R} \to [0, \infty]$ is called a *measure* on $\mathcal{R}$ if

1.1) $\mu(\varnothing) = 0$,

1.2) $\mu(A) \leq \mu(B)$ for $A, B \in \mathcal{R}$ such that $A \subseteq B$, and

1.3) μ is *countably additive*:

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$$

for any $A_i \in \mathcal{R}$ $(i = 1, 2, \cdots)$ which are disjoint, such that $\bigcup_{i=1}^{\infty} A_i \in \mathcal{R}$.

2. Outer measures. If the condition 1.3) is replaced by *countable sub-additivity*, then we obtain the definition of outer measures. That is, μ is an *outer measure* on $\mathcal{R}$, if 1) and 2) hold, and μ is a countably sub-additive:

$$\mu(A) \leq \sum_{i=1}^{\infty} \mu(A_i)$$

for any $A_i, A \in \mathcal{R}$ $(i = 1, 2, \cdots)$ such that $A \subset \cup_{i=1}^{\infty} A_i$.

3. Finite measures and σ-finite measures. A measure μ on $\mathcal{R}$ is finite if $\mu(E) < \infty$ for every $E \in \mathcal{R}$. μ is called σ-finite on $\mathcal{R}$ if there is a sequence of subsets $E_i \in \mathcal{R}$ such that $\bigcup_{i=1}^{\infty} E_i = \Omega$ and $\mu(E_i) < \infty$ for every $i = 1, 2, \cdots$. If $\mu(\Omega) = 1$, then μ is called a *probability measure* on $\mathcal{R}$.

4. Ring, algebra, σ-algebras and measurable spaces. We haven't imposed any algebraic structures yet on $\mathcal{R}$. Several notions may be introduced via set-theoretic operations: $\cup$, $\cap$ and complementary operation $\setminus$. A collection $\mathcal{R}$ of subsets of Ω is called a *ring* over Ω if $E_1 \cup E_2 \in \mathcal{R}$ and $E_1 \setminus E_2 \in \mathcal{R}$ for any $E_1, E_2 \in \mathcal{R}$. A ring $\mathcal{R}$ is an *algebra* if the total space $\Omega \in \mathcal{R}$. An algebra $\mathcal{F}$ over Ω is called a σ-algebra (or called a σ-field) if $\bigcup_{i=1}^{\infty} E_i \in \mathcal{F}$ for any $E_i \in \mathcal{F}$. If $\mathcal{F}$ is a σ-algebra over Ω, then $(\Omega, \mathcal{F})$ is called a measurable space.

If $\mathcal{A}$ is a non-empty collection of some subsets of Ω, then there is a unique σ-algebra over Ω, denoted by $\sigma\{\mathcal{A}\}$, which possesses the following properties: (1) $\mathcal{A} \subseteq \sigma\{\mathcal{A}\}$, and (2) if $\mathcal{F}$ is a σ-algebra over Ω containing $\mathcal{A}$, then $\sigma\{\mathcal{A}\} \subseteq \mathcal{F}$. In fact

$$\sigma\{\mathcal{A}\} = \bigcap \{\mathcal{F} : \mathcal{F} \text{ is a } \sigma\text{-algebra containing } \mathcal{A}\}.$$

$\sigma\{\mathcal{A}\}$ is the smallest σ-algebra containing $\mathcal{A}$, called the σ-algebra generated by $\mathcal{A}$.

5. *Measure spaces and probability spaces.* If $(\Omega, \mathcal{F})$ is a measurable space and μ is a measure on $\mathcal{F}$, then $(\Omega, \mathcal{F}, \mu)$ is called a *measure space.* If $\mu(\Omega) = 1$ then $(\Omega, \mathcal{F}, \mu)$ is called a probability space. In this case Ω is called a sample space (of fundamental events), an element A in the σ-algebra $\mathcal{F}$ is called an event, and $\mu(A)$ is called the probability that the event A occurs. A probability measure μ is usually denoted by a blackboard letter $\mathbb{P}$.

6. *Measurable functions.* $\mathscr{B}(\mathbb{R}^n)$ denotes the Borel σ-algebra on $\mathbb{R}^n$, which is the smallest σ-algebra containing open subsets. A function $f : \Omega \to [-\infty, \infty]$ is measurable with respect to a σ-field $\mathcal{F}$, or simply called $\mathcal{F}$-measurable, if

$$f^{-1}(G) = \{f \in G\} \equiv \{\omega \in \Omega : f(\omega) \in G\}$$

belongs to $\mathcal{F}$ for every $G \in \mathscr{B}(\mathbb{R})$, and both $f^{-1}(\infty)$ and $f^{-1}(-\infty)$ belong to $\mathcal{F}$ as well.

7. *Structure of measurable functions.* A simple (measurable) function φ on $(\Omega, \mathcal{F})$ is a (real valued) function on Ω which can be written as $\varphi = \sum_{k=1}^{n} c_k 1_{E_k}$ for some n, some constants c_k and some $E_k \in \mathcal{F}$. A function $f : \Omega \to [0, \infty]$ is $\mathcal{F}$-measurable, if and only if there is an increasing sequence of non-negative, $\mathcal{F}$-measurable simple functions φ_n such that $\varphi_n \uparrow f$ everywhere on Ω.

8. *Definition of Lebesgue's integrals.* Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. The Lebesgue theory of integration, developed in Part A Integration, may be applied to the measure μ. Let us recall quickly the procedure of defining Lebesgue's integrals. First define integrals for simple function, namely, if $\varphi = \sum_{j=1}^{m} c_j 1_{E_j}$ is a non-negative ($\mathcal{F}$-measurable) *simple function* on Ω, where $c_i \geq 0$ and $E_i \in \mathcal{F}$ for $i = 1, \cdots, m$, then $\int_E \varphi d\mu = \sum_{i=1}^{m} c_i \mu(E_i)$. If $f : \Omega \to [0, \infty]$ is a non-negative $\mathcal{F}$-measurable function, then

$$\int_\Omega f d\mu = \sup \left\{ \int_E \varphi d\mu : \varphi \leq f \text{ where } \varphi = \sum_{i=1}^{m} c_i 1_{E_i} \text{ and } c_i \geq 0, E_i \in \mathcal{F} \right\}.$$

9. *Integrable functions.* If f is non-negative measurable and if $\int_\Omega f d\mu < \infty$, then we say f is (Lebesgue) integrable on Ω with respect to the measure μ, denoted by $f \in L^1(\Omega, \mathcal{F}, \mu)$, $f \in L^1(\Omega, \mu)$, $L^1(\Omega)$ or simply by $f \in L^1$ if the measure space in question is clear. If $f : \Omega \to [-\infty, \infty]$ is $\mathcal{F}$-measurable, so are $f^+ = f \vee 0$, $f^- = (-f) \vee 0$ and $|f| = f^+ - f^-$. If both f^+ and f^- are integrable, then we say f is integrable, denoted by $f \in L^1(\Omega, \mathcal{F}, \mu)$ etc., and define its (Lebesgue) integral by

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

If $f : \Omega \to \mathbb{C}$ is a complex, $\mathcal{F}$-measurable function: $f = u + \sqrt{-1}v$, then f is integrable if both real part u and imaginary part v are integrable against the measure μ, and in this case, the Lebesgue integral of f is defined by

$$\int_\Omega f d\mu = \int_\Omega u d\mu + \sqrt{-1} \int_\Omega v d\mu.$$

$L^1(\Omega, \mathcal{F}, \mu)$ denotes the vector space of all $\mathcal{F}$-measuable (real or complex valued) integrable function on $(\Omega, \mathcal{F}, \mu)$.

The convergence theorems are applicable to a measure space $(\Omega, \mathcal{F}, \mu)$, and they may be stated as the following.

10. Monotone Convergence Theorem (MCT, due to Lebesgue and Levi). Suppose $f_n : \Omega \to [0, \infty]$ are non-negative, measurable, and suppose $f_{n+1} \geq f_n$ almost everywhere on Ω for all n, then

$$\int_\Omega \lim_{n\to\infty} f_n d\mu = \lim_{n\to\infty} \int_\Omega f_n d\mu = \sup_n \int_\Omega f_n d\mu.$$

In particular, if $\left\{\int_\Omega f_n d\mu\right\}$ is bounded above, then $\lim_{n\to\infty} f_n$ is integrable.

11. Series version of MCT (due to Lebesgue and Levi). This is very useful and is handy in applications. If a_n are *non-negative* and *measurable*, then

$$\int_\Omega \sum_n a_n d\mu = \sum_n \int_\Omega a_n d\mu.$$

12. Fatou's Lemma. Suppose $f_n : \Omega \to [0, \infty]$ are non-negative and measurable, then

$$\int_\Omega \liminf_{n\to\infty} f_n d\mu \leq \liminf_{n\to\infty} \int_\Omega f_n d\mu.$$

13. Lebesgue's Dominated Convergence Theorem (DCT). Suppose $f_n : \Omega \to [-\infty, \infty]$ (or $f_n : \Omega \to \mathbb{C}$) are measurable, $f_n \to f$ almost everywhere, and suppose there is an integrable (control) function g such that $|f_n| \leq g$ almost everywhere for all n, then f_n are integrable and

$$\lim_{n\to\infty} \int_\Omega f_n d\mu = \int_\Omega f d\mu.$$

14. Reverse Fatou's Lemma. Suppose f_n and g are integrable, and $f_n \leq g$ almost surely for $n = 1, 2, \cdots$. Then $g - f_n$are non-negative, and $\liminf(g - f_n) = g - \limsup f_n$. Applying Fatou's lemma to $g - f_n$ we obtain

$$\begin{aligned}\int_\Omega \left[g - \limsup_{n\to\infty} f_n\right] d\mu &\leq \liminf_{n\to\infty} \left[\int_\Omega g - \int_\Omega f_n d\mu\right] \\ &= \int_\Omega g d\mu - \limsup_{n\to\infty} \int_\Omega f_n d\mu\end{aligned}$$

which in particular yields that

$$\int_\Omega g d\mu - \limsup_{n\to\infty} \int_\Omega f_n d\mu \geq 0$$

so that $\limsup_{n\to\infty} \int_\Omega f_n d\mu \leq \int_\Omega g d\mu$. If $\limsup_{n\to\infty} \int_\Omega f_n d\mu > -\infty$, then

$$\int_\Omega g d\mu - \limsup_{n\to\infty} \int_\Omega f_n d\mu < \infty$$

so that $g - \limsup_{n\to\infty} f_n$ is integrable, and $\limsup_{n\to\infty} \int_\Omega f_n d\mu \leq \int_\Omega \limsup_{n\to\infty} f_n d\mu$. Let us state what we have proved as the following.

Theorem 1.1 *(*Reverse Fatou's Lemma*) Suppose* f_n *and* g *are integrable, and* $f_n \leq g$ *almost surely for* $n = 1, 2, \cdots$, *and suppose* $\limsup_{n\to\infty} \int_\Omega f_n d\mu > -\infty$, *then* $\limsup_{n\to\infty} f_n$ *is integrable and*

$$\int_\Omega \limsup_{n\to\infty} f_n d\mu \geq \limsup_{n\to\infty} \int_\Omega f_n d\mu.$$

15. Notations. If $f \in L^1(\Omega, \mathcal{F}, \mu)$ or if f is non-negative and measurable, then we also use $\mathbb{E}^\mu(f)$, $\mu(f)$ or $\mathbb{E}(f)$ to denote Lebesgue integral $\int_\Omega f d\mu$. If $A \in \mathcal{F}$, then $(A, A \cap \mathcal{F}, \mu)$ is a measure space too. In this case $\int_A f d\mu$ concides with $\int_\Omega f 1_A d\mu$, which will be denoted by $\mathbb{E}^\mu [f : A]$ or by $\mathbb{E}[f : A]$ if the measure in question is clear.

16. The L^p*space* for $p \in [1, \infty]$ can be defined over a measure space. When dealing with L^p-spaces, we identify an $\mathcal{F}$-measurable function f on $(\Omega, \mathcal{F}, \mu)$ with its equivalent class of all $\mathcal{F}$-measurable functions which are equal to f almost surely on Ω. Then $L^p(\Omega, \mathcal{F}, \mu)$ is the vector space of all $\mathcal{F}$-measurable functions f such that $|f|^p$ is μ-integrable, equipped with the L^p-norm: if $p \in [1, \infty)$, then

$$\|f\|_p = \left(\int_\Omega |f|^p d\mu\right)^{\frac{1}{p}} = (\mathbb{E}|f|^p)^{\frac{1}{p}}.$$

If $p = \infty$, then

$$\|f\|_\infty = \inf\{K : |f| \leq K \text{ on } \Omega \setminus N \text{ for some } N \in \mathcal{F} \text{ such that } \mu(N) = 0\}$$

which is called the μ-essential supremum of $|f|$.

17. Convergence in L^p-spaces. $L^p(\Omega, \mathcal{F}, \mu)$ are Banach spaces. $f \to ||f||_p$ is a norm on $L^p(\Omega, \mathcal{F}, \mu)$, and $L^p(\Omega, \mathcal{F}, \mu)$ is a complete metric space under the induced distance $(f, g) \to ||f - g||_p$. We say a sequence f_n converges to f in $L^p(\Omega, \mathcal{F}, \mu)$ if f_n and f belong to $L^p(\Omega, \mathcal{F}, \mu)$ and $||f_n - f||_p \to 0$, which is equivalent to that $\int_\Omega |f_n - f|^p d\mu \to 0$.

Let us give a short discussion about the convergence in L^1-space, and we will come back to this topic by introducing the notion of uniform integrability. The following simple fact about L^1-convergence, it is quite useful though, and its proof is a good exercise about DCT.

Theorem 1.2 *(*Scheffe's Lemma*) Suppose* f_n *and* f *are integrable, and* $f_n \to f$ *almost surely. Then* $f_n \to f$ *in* $L^1(\Omega, \mathcal{F}, \mu)$ *if and only if* $\mathbb{E}^\mu[|f_n|] \to \mathbb{E}^\mu[|f|]$.

Proof. "Only if" part is easy. In fact, if $f_n \to f$ in L^1, then, by the triangle inequality,

$$||f_n| - |f|| \leq |f_n - f|$$

so that

$$0 \leq \left|\int_\Omega |f_n| d\mu - \int_\Omega |f| d\mu\right| = \left|\int_\Omega (|f_n| - |f|)\right| \leq \int_\Omega |f_n - f| d\mu \to 0$$

which implies that $\int_\Omega |f_n| d\mu \to \int_\Omega |f| d\mu$.

Proof of "If" part. Assume that $f_n \to f$ almost surely and $\int_\Omega |f_n| d\mu \to \int_\Omega |f| d\mu$. We want to show that $f_n \to f$ in L^1. To this end, we decompose the sample space Ω into two components for each n: $A_n = \{f_n f \geq 0\}$, $B_n = \{f_n f < 0\}$. Then

$$|f_n - f| = ||f_n| - |f|| \quad \text{on } A_n$$

and, by the triangle inequality,

$$|f_n - f| = ||f_n| + |f|| \leq ||f_n| - |f|| + 2|f| \quad \text{on } B_n.$$

Hence

$$\begin{aligned}
\int_\Omega |f_n - f| d\mu &= \int_{A_n} |f_n - f|\, d\mu + \int_{B_n} |f_n - f|\, d\mu \\
&\leq \int_{A_n} ||f_n| - |f||\, d\mu + \int_{B_n} [||f_n| - |f|| + 2|f|]\, d\mu \\
&= \int_\Omega ||f_n| - |f||\, d\mu + 2\int_{B_n} |f| d\mu \\
&= \int_\Omega ||f_n| - |f||\, d\mu + 2\int_\Omega 1_{B_n}|f| d\mu.
\end{aligned}$$

The first term on the right-hand side of the previous inequality many be rewritten as the following

$$\begin{aligned}
\int_\Omega ||f_n| - |f||\, d\mu &= \int_\Omega (|f_n| - |f|)^+\, d\mu + \int_\Omega (|f_n| - |f|)^-\, d\mu \\
&= \int_\Omega (|f_n| - |f|)\, d\mu + 2\int_\Omega (|f_n| - |f|)^-\, d\mu
\end{aligned}$$

where we have used the identity

$$|g| = g^+ + g^- = g^+ - g^- + 2g^- = g + 2g^-.$$

Putting together we obtain the following estimate for the L^1-norm of $f_n - f$:

$$\begin{aligned}
\int_\Omega |f_n - f| d\mu &\leq \int_\Omega ||f_n| - |f||\, d\mu + 2\int_\Omega 1_{B_n}|f| d\mu \\
&= \int_\Omega (|f_n| - |f|)\, d\mu + 2\int_\Omega (|f_n| - |f|)^-\, d\mu + 2\int_\Omega 1_{B_n}|f| d\mu. \qquad (1.1)
\end{aligned}$$

We next want to let $n \to \infty$ in the inequality above. The first term on the right-hand side tends to zero as $n \to \infty$ by assumption. In fact

$$\int_\Omega (|f_n| - |f|)\, d\mu = \int_\Omega |f_n| d\mu - \int_\Omega |f| d\mu \to 0$$

as $n \to \infty$. For the second term, we observe that

$$(|f_n| - |f|)^- = 0 \qquad \text{on } \{|f_n| \geq |f|\}$$

and

$$(|f_n| - |f|)^- = |f_n| - |f| \leq |f_n| \leq |f| \qquad \text{on } \{|f_n| < |f|\}$$

so that

$$(|f_n| - |f|)^- \leq |f|$$

for all n, $|f|$ is integrable, and $(|f_n| - |f|)^- \to 0$ almost surely, thus by the Dominated Convergence Theorem we conclude that

$$\int_\Omega (|f_n| - |f|)^-\, d\mu \to 0.$$

To show the last term on the right-hand side of (1.1) $\int_{B_n} |f| d\mu$ tends to zero, we notice that $|f|1_{B_n} \to 0$. While it is clear that $|f|1_{B_n} = 0$ on $\{|f| = 0\}$ for all n. If $|f(x)| > 0$, and $f_n(x) \to f(x)$, then there is N (depending on x in general) such that $|f_n(x) - f(x)| < \frac{1}{2} f(x)$ so that $f_n(x) f(x) > 0$ for all $n > N$, hence $x \notin B_n$ for $n > N$. Thus $1_{B_n}(x) = 0$ for all $n > N$. Hence $|f|1_{B_n}(x) = 0$ for all $n > N$. Since $f_n \to f$ almost surely, we thus can conclude that $|f|1_{B_n} \to 0$ almost everywhere as $n \to \infty$. a $|f|1_{B_n}$ is controlled by the integral function $|f|$, so by DCT we have $\int_{B_n} |f| d\mu = \int_\Omega |f| 1_{B_n} d\mu \to 0$. Therefore, by Sandwich lemma, it follows from (1.1) that $\lim_{n\to\infty} \int_\Omega |f_n - f| d\mu = 0$. ■