FIN 508- MARTINGALE THROUGH MEASURE THEORY
LECTURE 4

GENERALIZED MEASURES AND RADON-NIKODYM'S DERIVATIVE

1. *Generalized measures.* Let $(\Omega, \mathcal{F})$ be a measurable space. If μ_1 and μ_2 are two measures on $\mathcal{F}$, and if one of them is finite so that their difference

$$\mu(E) = \mu_1(E) - \mu_2(E)$$

for $E \in \mathcal{F}$ defines a function (called a signed measure) from $\mathcal{F}$ to $[-\infty, \infty]$, which is, though not a positive measure, countably additive. Such "generalized measures" are interesting and

are arisen naturally in Lebesgue's integration. For example, if f is integrable function on a measure space $(\Omega, \mathcal{F}, \mu)$, then

$$\mu_f(E) = \int_E f d\mu = \int_E f^+ d\mu - \int_E f^- d\mu, \text{ for } E \in \mathcal{F},$$

is an example of "generalized measures". We therefore generalize the definition of measures to the so-called *generalized measures* as the following. A function $\mu : \mathcal{F} \to (-\infty, \infty]$ is called a *generalized measure* (which does not take value $-\infty$) if

1) $\mu(\emptyset) = 0$,
2) μ is countably additive in the sense that

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

for any $A_i \in \mathcal{F}$ which are disjoint. While of course we can define generalized measures μ take values in $[-\infty, \infty)$ instead, but it is not necessary, as in this case $-\mu$ takes values in $(-\infty, \infty]$.

2. *Hahn's decomposition for generalized measures.* Clearly, any signed measure $\mu = \mu_1 - \mu_2$, where μ_i are measures on $(\Omega, \mathcal{F})$ and $\mu_2(\Omega) < \infty$, is a generalized measure. The converse is also true.

Theorem 4.1 *(*Hahn's decomposition*) If μ is a generalized measure on $(\Omega, \mathcal{F})$, then there is a decomposition $\Omega = A^+ \cup A^-$, where $A^+, A^- \in \mathcal{F}$ such that $A^+ \cap A^- = \emptyset$, and*

$$\mu(E \cap A^+) \geq 0, \; \mu(E \cap A^-) \leq 0$$

for every $E \in \mathcal{F}$. Moreover the positive and negative part A^+ and A^- are unique in the sense that if A_i^+ and A_i^- (where $i = 1, 2$) are two pairs satisfying the Hahn's decomposition, then

$$\mu(E \cap A_1^+) = \mu(E \cap A_2^+), \text{ and } \mu(E \cap A_1^-) = \mu(E \cap A_2^-)$$

for every $E \in \mathcal{F}$.

Proof. [*The proof is not examinable.*] The unique sets A^+ and A^- (up to a "null set") are called the *positive* (resp. *negative*) set of the generalized measure μ. Let

$$\lambda = \inf\{\mu(G) : \text{ where } G \in \mathcal{F} \text{ such that } \mu(E \cap G) \leq 0 \text{ for all } E \in \mathcal{F}\}.$$

Choose a sequence $G_n \in \mathcal{F}$ such that $\mu(G_n) \to \lambda$ as $n \to \infty$. Then the candidate for A^- should be the largest possible negative set, that is

$$A^- = \bigcup_{n=1}^{\infty} \left(G_n \setminus \cup_{j=1}^{n-1} G_j\right).$$

In fact, A^- is still a negative set: $\mu(E \cap A^-) \leq 0$ for every $E \in \mathcal{F}$, and therefore $\mu(A^-) = \lambda$ (which yields also that $\lambda > -\infty$). We claim that the pair $A^+ = \Omega \setminus A^-$ and A^- is a decomposition satisfying that $\mu(E \cap A^+) \geq 0$ and $\mu(E \cap A^-) \leq 0$ for every $E \in \mathcal{F}$.

We only have to show that $\mu(E \cap A^+) \geq 0$ for every $E \in \mathcal{F}$, that is for any $E \subseteq A^+$, $\mu(E) \geq 0$. Let us argue by a contradiction. Suppose there is an $E_0 \subseteq A^+$ such that $\mu(E_0) < 0$. Then, since $E_0 \cap A^- = \emptyset$, so that

$$\mu(A^- \cup E_0) = \mu(A^-) + \mu(E_0) = \lambda + \mu(E_0) < \lambda$$

which is a contradiction to the definition of λ, and therefore $A^- \cup E_0$ can not be a negative set of μ, so there is a subset $A_1 \subseteq E_0$ such that $\mu(A_1) > 0$. Hence

$$k_1 = \min\left\{n \in \mathbb{N} : \text{ there is } A_1 \subseteq E_0, \mu(A_1) \geq \frac{1}{n}\right\}$$

exists, and we can find an $E_1 \subseteq \mathcal{F}$ such that $E_1 \subseteq E_0$ and $\frac{1}{k_1} \leq \mu(E_1) < \frac{1}{k_1 - 1}$. Clearly

$$\mu(E_0 \setminus E_1) = \mu(E_0) - \mu(E_1) < 0$$

so we can argue as above with $E_0 \setminus E_1$ in place of E_0 and choose $E_2 \subseteq E_0 \setminus E_1$ such that $\mu(E_2) > 0$ and $\frac{1}{k_2} \leq \mu(E_2) < \frac{1}{k_2 - 1}$, where

$$k_2 = \min\left\{n \in \mathbb{N} : \text{ there is } A_1 \subseteq E_0 \setminus E_1, \mu(A_1) \geq \frac{1}{n}\right\}.$$

Repeating the previous procedure we may construct a sequence of E_n inductively, such that $E_n \subseteq E_0 \setminus \cup_{j=1}^{n-1} E_j$ [in particular E_n are disjoint], k_n are non-decreasing, such that $\frac{1}{k_n} \leq \mu(E_n) < \frac{1}{k_n - 1}$, and

$$k_n = \min\left\{n \in \mathbb{N} : \text{ there is } A \subseteq E_0 \setminus \bigcup_{i=1}^{n-1} E_i \text{ such that } \mu(A) \geq \frac{1}{n}\right\}.$$

We claim that $\sum_n \frac{1}{k_n} < \infty$, since, otherwise, we would have

$$\sum_n \mu(E_n) \geq \sum_n \frac{1}{k_n} = \infty.$$

Since $\mu(E_0) < 0$ and

$$\mu(E_0) = \sum_n \mu(E_n) + \mu\left(E_0 \setminus \cup_{n=1}^{\infty} E_n\right)$$

we may deduce that

$$\mu\left(E_0 \setminus \bigcup_{n=1}^{\infty} E_n\right) = -\infty$$

which is a contradiction to the assumption that $\mu(E) > -\infty$ for every $E \in \mathcal{F}$. Therefore it must be hold that $k_n \to \infty$, so that $\mu(E_n) \to 0$, hence any subset of $E_0 \setminus \cup_{n=1}^{\infty} E_n$ has non-positive measure, and

$$\mu\left(E_0 \setminus \bigcup_{n=1}^{\infty} E_n\right) = \mu(E_0) - \sum_{n=1}^{\infty} \mu(E_n) < \lambda$$

which contradicts to the definition of λ. ■

For a different approach, read W. Rudin: Real and Complex Analysis, Third Edition, pages 120-126.

3. Jordan's decomposition for generalized measures. Thus, if μ is a generalized measure over $(\Omega, \mathcal{F})$, and $\Omega = A^+ \cup A^-$ is an Hahn decomposition with respect to μ, then $\mu^+(E) = \mu(E \cap A^+)$ and $\mu^-(E) = -\mu(E \cap A^-)$ (where $E \in \mathcal{F}$) define two measures on $(\Omega, \mathcal{F})$. Moreover, μ^- is a finite measure. By definition, $\mu = \mu^+ - \mu^-$ is thus a *signed measure*, called the *Jordan decomposition* of the generalized measure μ. We may also define $|\mu| = \mu^+ + \mu^-$

which is also a measure on (Ω, μ), called the *total variation measure* of the generalized measure $\mu = \mu^+ - \mu^-$.

If ρ is a function defined on (a, b), which has finite total variation, that is,

$$\sup_D \sum_{j=1}^n |\rho(t_j) - \rho(t_{j-1})| < \infty$$

where the sup takes over all possible finite partitions $D : a < t_0 < t_1 < \cdots < t_n < b$. Then

$$\rho_{\mathrm{TV}}(t) \equiv \sup_{D_t} \sum_{j=1}^n |\rho(t_j) - \rho(t_{j-1})|$$

defines an increasing function, where the sup runs over all finite partitions $D_t : a < t_0 < t_1 < \cdots < t_n = t$, for every $t \in (a, b)$. $\rho_N(t) \equiv \rho_{\mathrm{TV}}(t) - \rho(t)$ is also increasing. In particular, ρ is a difference of two increasing functions, so that ρ has left and right limits at every $t \in (a, b)$. Moreover, if ρ is right continuous at t, then so is ρ_{TV}. Therefore if ρ is right continuous and has finite total variation, then $\rho = \rho_1 - \rho_2$ a difference of two right continuous and increasing functions. $m_\rho \equiv m_{\rho_1} - m_{\rho_2}$ is a signed measure. In this case the total variation measure $|m_\rho| = m_{\rho_{\mathrm{TV}}}$.

4. Lebesgue's integrals w.r.t. a generalized measure. The usual concepts about measures may be applied to generalized measures via Jordan's decomposition. For example, we say a generalized measure μ is σ-finite if $|\mu|$ is σ-finite, which is equivalent to say both μ^+ and μ^- are σ-finite. The theory of Lebesgue's integration may be applied to a generalized measure $\mu = \mu^+ - \mu^-$ on $(\Omega, \mathcal{F})$ too. For example, an $\mathcal{F}$-measurable function $f : \Omega \to [-\infty, \infty]$ is μ-integrable if and only if, by definition, f is integrable against the total variation measure $|\mu| = \mu^+ + \mu^-$ (which is equivalent to say f is integrable with respect both measures μ^+ and μ^-), and in this case

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

5. Absolute continuity and Radon-Nikodym's theorem. Next we turn to an important concept about two generalized measures: the concept of absolute continuity.

Definition 4.2 *Let ν and μ be two measures on a measurable space $(\Omega, \mathcal{F})$, then we say ν is* absolutely continuous *with respect to μ, written as $\nu \ll \mu$, if $E \in \mathcal{F}$ and $\mu(E) = 0$ implies that $\nu(E) = 0$. That is, any μ-null set is also a ν-null set.*

Theorem 4.3 (Radon-Nikodym's derivative) *If μ and ν are two σ-finite measures on $(\Omega, \mathcal{F})$, such that $\nu \ll \mu$, then there is a* non-negative *$\mathcal{F}$-measurable function ρ such that*

$$\nu(E) = \int_E \rho d\mu \text{ for every } E \in \mathcal{F}.$$

Moreover ρ is unique up to μ-almost everywhere. ρ is called the Radon-Nikodym derivative of ν with respect to μ, denoted by $\frac{d\nu}{d\mu}$.

Proof. [*The proof is not examinable.*] Let us outline the proof of this important theorem for the case where ν and μ are two *finite measures*: $\mu(\Omega) < \infty$ and $\nu(\Omega) < \infty$. In this case, let $\mathscr{L}$ denote the collection of all non-negative measurable functions h such that

$$\mu[h : E] = \int_E h d\mu \le \nu(E) \text{ for every } E \in \mathcal{F}.$$

Then, $\mathscr{L}$ is a non-empty class. Now consider $\lambda = \sup_{h\in\mathscr{L}} \int_\Omega h d\mu$. Then, clearly $\lambda \geq 0$ and $\lambda \leq \nu(\Omega) < \infty$. Choose a sequence of functions $h_n \in \mathscr{L}$ such that $\int_\Omega h_n d\mu \to \lambda$. Let $\rho = \sup_n h_n$. We claim that ρ is the Radon-Nikodym derivative. To this end, set $\rho_n = \max\{h_1, \cdots, h_n\}$ for every n. For every n, we may choose a decomposition $\Omega = \cup_{i=1}^n E_i^{(n)}$ where $E_i^{(n)} \in \mathcal{F}$ which are disjoint, and $\rho_n = h_i$ on $E_i^{(n)}$ for $i = 1, \cdots, n$. Thus, for every $E \in \mathcal{F}$, we have

$$\int_E \rho_n d\mu = \sum_{i=1}^n \int_{E_i\cap E} h_i d\mu \leq \sum_{i=1}^n \nu(E_i \cap E) = \nu(E)$$

that is, $\rho_n \in \mathscr{L}$. By definition, $\rho_n \uparrow \rho$, so by MCT, $\rho = \lim \rho_n \in L^1(\Omega, \mu)$, and by our construction, $\int_\Omega \rho d\mu = \lambda$ and $\rho \in \mathscr{L}$, i.e. $\int_E \rho d\mu \leq \nu(E)$ for every $E \in \mathcal{F}$. In particular, $\rho < \infty$ μ-almost everywhere, hence ν-almost everywhere as $\nu \ll \mu$. Therefore, we may assume that ρ is finite everywhere.

We next show that $\nu(E) = \int_E \rho d\mu$ for every $E \in \mathcal{F}$. To this end consider the generalized measure

$$m(E) = \nu(E) - \int_E \rho d\mu$$

where $E \in \mathcal{F}$. Since $\rho \in \mathscr{L}$, m is a measure, and we want to show that $m = 0$. Suppose there is $E_0 \in \mathcal{F}$ such that $m(E_0) > 0$, thus

$$\nu(E_0) > \int_{E_0} \rho d\mu.$$

Hence, there must exist $\varepsilon > 0$, such that $\nu(E_0) > \varepsilon\mu(E_0)$. Applying Hahn's decomposition to the generalized measure $\nu - \varepsilon\mu$, there is an positive set A^+ with respect to $\nu - \varepsilon\mu$, so that

$$\nu(A^+ \cap E) - \varepsilon\mu(A^+ \cap E) \geq 0$$

and

$$\nu(A^+) - \varepsilon\mu(A^+) > 0.$$

Since $\nu \ll \mu$, the last inequality yields that $\mu(A^+) > 0$. Now consider $\varphi = \rho + \varepsilon 1_{A^+}$. Then for every $E \in \mathcal{F}$, we have

$$\begin{aligned}\int_E \varphi d\mu &= \int_{E\cap A^+} (\rho + \varepsilon 1_{A^+})\, d\mu + \int_{E\setminus A^+} \rho d\mu \\ &\leq (\nu - m)\,(E \cap A^+) + \varepsilon\mu(E \cap A^+) + \nu(E \setminus A^+) \\ &\leq \nu(E \cap A^+) + \nu(E \setminus A^+) \\ &= \nu(E)\end{aligned}$$

so that $\varphi \in \mathscr{L}$. On the other hand

$$\int_\Omega \varphi d\mu = \int_\Omega \rho d\mu + \varepsilon \int_\Omega 1_A d\mu = \lambda + \varepsilon\mu(A) > \lambda$$

a contradiction to the definition of λ. ∎

6. An integral formula. The following theorem follows from a routine computation.

Theorem 4.4 *Suppose μ and ν are two σ-finite measures on $(\Omega, \mathcal{F})$, such that $\nu \ll \mu$. Let f be an $\mathcal{F}$-measurable function. Then f is integrable with respect to ν if and only if $f\frac{d\nu}{d\mu}$ is integrable with respect to μ, and*

$$\int_\Omega f d\nu = \int_\Omega f \frac{d\nu}{d\mu} d\mu.$$

FIN 508- MARTINGALE THROUGH MEASURE THEORY
LECTURE 4

7. Conditional expectations. This is perhaps the most important concept in probability theory. Let $(\Omega, \mathcal{F}, \mu)$ be a measure space, and let $f : \Omega \to [0, \infty]$ be $\mathcal{F}$-measurable. For every $A \in \mathcal{F}$, define $\mu_f(A) = \int_\Omega f 1_A d\mu = \int_A f d\mu$. Then μ_f is a measure defined on $\mathcal{F}$. In fact, if A_n is a sequence of disjoint $\mathcal{F}$-measurable subsets, then $f 1_{\bigcup_{n=1}^\infty A_n} = \sum_{n=1}^\infty f 1_{A_n}$, thus, by MCT (series version) we have

$$\mu_f\left(\bigcup_{n=1}^{\infty} A_n\right) = \int_\Omega f 1_{\bigcup_{n=1}^\infty A_n} d\mu = \sum_{n=1}^{\infty} \int_\Omega f 1_{A_n} d\mu = \sum_{n=1}^{\infty} \mu_f(A_n)$$

so μ_f is a measure on $(\Omega, \mathcal{F})$.

μ_f possesses an important property – if $A \in \mathcal{F}$ is a μ-null set, i.e. $\mu(A) = 0$, then A is also a μ_f-null set: $\mu_f(A) = 0$ [which of course follows from that the integral of function on a null set is zero on any measure space]. That is to say the measure μ_f is absolutely continuous with respect μ, that is, $\mu_f \ll \mu$. Conversely is also true, which is the context of Randon-Nikydom's theorem.

Suppose $(\Omega, \mathcal{F}, \mu)$ is a measure space, and $\mathcal{G}$ is a sub σ-algebra of $\mathcal{F}$. Suppose μ is σ-finite on $\mathcal{G}$, so that there is a sequence $G_n \in \mathcal{G}$, $G_n \uparrow \Omega$ and $\mu(G_n) < \infty$ for every n. Let f be $\mathcal{F}$-measurable and non-negative such that f is σ-integrable on $\mathcal{G}$, that is, there are $G_n \in \mathcal{G}$ such that $G_n \uparrow \Omega$ and $\int_{G_n} f d\mu < \infty$ for every n. Then $\mu_f \ll \mu$ as measures on $(\Omega, \mathcal{G})$, and both μ_f and μ are σ-finite measure on $(\Omega, \mathcal{G})$, therefore, by applying Randon-Nikydom's theorem to μ and μ_f on $(\Omega, \mathcal{G})$, there is a $\mathcal{G}$-measurable and non-negative function ρ (unique up to μ-almost surely) such that $\mu_f(A) = \int_A \rho d\mu$ for every $A \in \mathcal{G}$ [that is, ρ is the Random-Nikydom's derivative of μ_f with respect to μ on $\mathcal{G}$, so denoted by $\rho = \left.\frac{d\mu_f}{d\mu}\right|_{\mathcal{G}}$]. $\left.\frac{d\mu_f}{d\mu}\right|_{\mathcal{G}}$ is called the *conditional expectation of f given $\mathcal{G}$*, denoted by $\mathbb{E}^\mu[f|\mathcal{G}]$ or simply by $\mathbb{E}[f|\mathcal{G}]$ if the measure μ involved is clear. The conditional expectation possesses the following properties:

1) $\mathbb{E}[f|\mathcal{G}]$ is $\mathcal{G}$-measurable,
2) for every $A \in \mathcal{G}$ we have

$$\mathbb{E}[f : A] = \mathbb{E}[\mathbb{E}(f|\mathcal{G}) : A]$$

that is

$$\mathbb{E}[f 1_A] = \mathbb{E}[1_A \mathbb{E}[f|\mathcal{G}]].$$

In particular, $\mathbb{E}[f] = \mathbb{E}[\mathbb{E}[f|\mathcal{G}]]$, so that, if f is integrable, so is its conditional expectation $\mathbb{E}[f|\mathcal{G}]$, which allows us to define the conditional expectation of an integrable function f by

$$\mathbb{E}[f|\mathcal{G}] = \mathbb{E}\left[f^+|\mathcal{G}\right] - \mathbb{E}\left[f^-|\mathcal{G}\right].$$