Fubini’s Theorem and Radon-Nikodym Theorem

This is a supplementary material intended for those who are interested in two important theorems in measure theory – Fubini’s Theorem, which is used to exchange the order of integrals, and Radon-Nikodym Theorem, which is often called change of measure theorem.

1 Fubini’s Theorem

In Exercise 5 of Homework 3, we encounter the exchange of expectation and integral. However, it is not in general valid to do so. The natural question to ask is, under what conditions can two integrals exchange?

We first present definitions of product measure and \( \sigma \)-finite measure.

**Definition 1.1.** A measure space \((X, \mathcal{F}, \mu)\) is said to be \( \sigma \)-finite if and only if \(X\) could be written as countable union of sets \(\{A_n\}_{n \geq 1}\) such that for every \(n\), \(\mu(A_n) < \infty\).

**Definition 1.2.** Let \((X, \mathcal{F}_1, \mu_1)\) and \((Y, \mathcal{F}_2, \mu_2)\) be two measure spaces. The product measure \(\mu\) defined on the product space \(X \times Y\) satisfies that for any measurable subsets \(A \subset X\) and \(B \subset Y\), it holds \(\mu(A \times B) = \mu_1(A)\mu_2(B)\). Moreover, the \(\sigma\)-algebra is the product \(\sigma\)-algebra of \(\mathcal{F}_1\) and \(\mathcal{F}_2\).

**Remark.** Such product measure is not unique in general, and the usual technique is to assume all measures to be \( \sigma \)-finite, which means that \(X\) and \(Y\) could be written as a countable union of sets which have finite measures (with respect to \(\mu_1\) and \(\mu_2\)). Carathéodory’s extension theorem then guarantees that such product measure is unique.

**Example 1.1.** Let \(X = Y = \mathbb{R}\) and \(\mu_1, \mu_2\) be one-dimensional Lebesgue measures. It is obvious that \((\mathbb{R}, \mathcal{B}(\mathbb{R}), dx)\) is \(\sigma\)-finite since \(\mathbb{R} = \bigcup_{N \geq 1} [-N, N]\). The product measure \(\mu\) satisfies, for example,
\[
\mu(\{(x, y) : 0 \leq x \leq 2, 1 \leq y \leq 4\}) = \mu_1([0, 2])\mu_2([1, 4]) = (2 - 0)(4 - 1) = 6.
\]
In geometry, the measure of a set is identical to its area.

Here comes the statement of Fubini’s Theorem.

**Theorem 1.1** (Fubini). Suppose \(X, Y\) are \( \sigma \)-finite measure spaces, and \(X \times Y\) is given the product measure. If \(f(x, y)\) is \(X \times Y\) measurable and integrable, which is
\[
\int_{X \times Y} |f(x, y)| d(x, y) < \infty,
\]
then
\[
\int_X \left( \int_Y f(x, y) dy \right) dx = \int_Y \left( \int_X f(x, y) dx \right) dy = \int_{X \times Y} f(x, y) d(x, y),
\]
which implies that the iterated integral and the double integral are equivalent.

There is also another theorem for non-negative functions, which does not require that \(f\) is integrable.
Theorem 1.2 (Tonelli). Suppose $X, Y$ are $\sigma$-finite measure spaces, and $X \times Y$ is given the product measure. If $f(x, y) \geq 0$ is $X \times Y$ measurable, then

$$\int_X \left( \int_Y f(x, y) \, dy \right) \, dx = \int_Y \left( \int_X f(x, y) \, dx \right) \, dy = \int_{X \times Y} f(x, y) \, d(x, y).$$

Remark. The reason that Fubini’s Theorem requires $f$ to be integrable is that it adopts the decomposition $f = f^+ - f^-$ where $f^+ = \max(f, 0)$ and $f^- = \max(-f, 0)$ and applies to both parts Tonelli’s Theorem. The constraint that $f$ is integrable eliminates the case of $\infty - \infty$, which is ill-defined.

The proof of Fubini and Tonelli is rather technical. You may refer to any classical measure theory books for the proof, or alternatively, see the following page.

We would like to use the following example to end this section.

Example 1.2 ($p$-th moment of non-negative functions). Suppose $f \in L^p(\Omega, \mathcal{F}, \mu)$ for $p \geq 1$, where

$$L^p(\Omega, \mathcal{F}, \mu) = \left\{ f \text{ measurable} : \int_{\Omega} |f(x)|^p \mu(dx) < \infty \right\}.$$

For simplicity, we write it as $f \in L^p$. Such space is called the Lebesgue space. It can be proved that $L^p$ is a complete normed space (Banach space) endowed with the norm

$$\|f\|_{L^p} := \left( \int_{\Omega} |f(x)|^p \mu(dx) \right)^{\frac{1}{p}}.$$

This example provides an alternative way to calculate the norm. Indeed,

$$\|f\|_{L^p}^p = \int_{\Omega} |f(x)|^p \mu(dx) = \int_{\Omega} \left( \int_0^\infty 1_{\{f(x) \geq s\}} \, ds \right) \mu(dx) = \int_0^\infty \left( \int_{\Omega} 1_{\{|f(x)| \geq s\}} \mu(dx) \right) \, ds,$$

where the last equality is due to Fubini’s (or Tonelli’s) Theorem. Let $s = t^p$. By the change of variable, we have $ds = pt^{p-1} \, dt$ and

$$\|f\|_{L^p}^p = \int_0^\infty \left( \int_0^t 1_{\{f(x) \geq \tau\}} \mu(dx) \right) pt^{p-1} \, dt = p \int_0^\infty t^{p-1} \mu(\{x : |f(x)| \geq t\}) \, dt.$$

For $Z$ being a non-negative random variable, its $p$-th moment could be calculated in the same way:

$$\mathbb{E}[Z^p] = \mathbb{E} \left[ \int_0^\infty 1_{\{Z^p \geq s\}} \, ds \right] = p \int_0^\infty t^{p-1} \mathbb{P}(\{Z \geq t\}) \, dt.$$

When $p = 1$, it coincides with Exercise 5 of Homework 3.

Exercise. Using a similar argument of the previous example, calculate $\mathbb{E}[\phi(Z)]$ for $\phi$ strictly increasing and having continuous derivative, where $Z \geq 0$ almost surely.

\footnote{https://patternsofideas.github.io/posts/fubini/}
2 Radon-Nikodym Theorem

Before stating the main theorem, we first need the definition of absolute continuity, which coincides with the notion that a random variable $X$ is continuous.

**Definition 2.1.** Let $\mu, \nu$ be two measures defined on the same space $(X, \mathcal{F})$. We say that $\mu$ is absolutely continuous with respect to $\nu$ if for any $A \in \mathcal{F}$ such that $\nu(A) = 0$, it holds $\mu(A) = 0$. It is denoted as $\mu \ll \nu$.

**Remark.** i) The notion defined in the lectures that $X$ is a continuous random variable is equivalent to that the measure induced by $X$ is absolutely continuous with respect to the Lebesgue measure. The measure induced by a random variable $X$ is defined as, for every Borel set $B$, 

$$\mu_X(B) := (X_\# \mathbb{P})(B) = \mathbb{P}(X^{-1}(B)) = \mathbb{P}(\{X \in B\}).$$

It is also called the pushforward measure of $\mathbb{P}$. The absolute continuity then implies that if $|B| = 0$, we have $\mu_X(B) = 0$, which coincides with the notion that $X$ is a continuous random variable.

ii) Absolute continuity has different equivalent statements. One version is that $\mu \ll \nu$ if and only if for every $\varepsilon > 0$, there exists $\delta > 0$ such that for every measurable set $A$, if $\mu(A) < \delta$ then $\nu(A) < \varepsilon$.

iii) We say that a function $f : \mathbb{R} \to \mathbb{R}$ is absolutely continuous if for every $\varepsilon > 0$, there exists $\delta > 0$ such that whenever a finite sequence of pairwise disjoint intervals $(x_n, y_n)$ satisfies $\sum_n (y_n - x_n) < \delta$, it holds $\sum_n |f(y_n) - f(x_n)| < \varepsilon$. It is stronger than uniform continuity but much weaker than differentiability.

Now, we state the Radon-Nikodym Theorem used for change of measure.

**Theorem 2.1.** Let $\mu, \nu$ be two $\sigma$-finite measures defined on the same space $(X, \mathcal{F})$. If $\mu \ll \nu$, there exists a unique non-negative $f \in L^1(X, \mathcal{F}, \mu)$ such that for every $A \in \mathcal{F}$,

$$\nu(A) = \int_A f \, d\mu.$$

Such $f$ is often called the Radon-Nikodym derivative and denoted as $\frac{d\nu}{d\mu}$. Moreover, for any non-negative measurable function $g$,

$$\int_X g \, d\nu = \int_X gf \, d\mu.$$

There is an elegant proof by John von Neumann, which adopts the Riesz Representation Theorem of a functional defined on the Hilbert space $L^2$. The proof could be found everywhere, for example, in the following page^{2}.

**Example 2.1** (Continuous random variables). If $X$ is a continuous random variable defined on the space $(\Omega, \mathcal{F}, \mathbb{P})$, then $\mu_X$ is absolutely continuous with respect to the Lebesgue measure according to the definition. Therefore, by Radon-Nikodym Theorem, there exists $p_X \in L^1$ such that for every Borel set $B$,

$$\mu_X(B) = (X_\# \mathbb{P})(B) = \int_B p_X(x) \, dx.$$
Moreover, for any measurable function \( g \) such that \( g \cdot p_X \) is integrable,

\[
E[g(X)] = \int_{\Omega} g \circ X(\omega) \, dP(\omega) = \int_{\mathbb{R}} g(x) \, d(X_#P)(x) = \int_{\mathbb{R}} g(x) p_X(x) \, dx.
\]

This is the formula calculating the expectation of a function of a continuous random variable \( X \).

If we take \( B = \{ x : x \leq t \} \), it follows that \( F_X(t) = \mu_X(B) \) and

\[
F_X(t) = \int_{-\infty}^{t} p_X(x) \, dx.
\]

This shows that if \( X \) is a continuous random variable, its cumulative distribution function is absolutely continuous. Conversely, if a cumulative distribution function is absolutely continuous, its corresponding random variable is continuous.