Chapter 5

Lebesgue's convergence theorems and $\mathcal{L}^p$ spaces

In this chapter we study two important convergence theorems and some of their uses and applications.

5.1 Convergence theorems

Theorem 5.1. *(Lebesgue's Monotone Convergence Theorem) Let $f_n : (\mathbb{X}, \mathcal{F}, \mu) \to (\bar{\mathbb{R}}, \bar{\mathcal{B}})$ for $n \in \mathbb{N}$ be integrable functions such that $f_1 \leq f_2 \leq \cdots$ and $f := \lim_{n\to\infty} f_n = \sup_{n\in\mathbb{N}} f_n$. Then,*

$$f \in \mathcal{L}_{\bar{\mathbb{R}}}(\mu) \iff \sup_{n\in\mathbb{N}} \int_{\mathbb{X}} f_n d\mu < \infty.$$

In this case,

$$\sup_{n\in\mathbb{N}} \int_{\mathbb{X}} f_n d\mu = \int_{\mathbb{X}} \sup_{n\in\mathbb{N}} f_n d\mu = \int_{\mathbb{X}} f d\mu.$$

Proof. Since $f_n \in \mathcal{L}_{\bar{\mathbb{R}}}$ and $f_1 \leq f_2 \leq \cdots$ we have that $0 \leq f_n - f_1 \in \mathcal{L}_{\bar{\mathbb{R}}}$ forms an increasing sequence of nonnegative measurable functions. Hence, by Theorem 4.5

$$0 \leq \sup_{n\in\mathbb{N}} \int_{\mathbb{X}} (f_n - f_1) d\mu = \int_{\mathbb{X}} \sup_{n\in\mathbb{N}} (f_n - f_1) d\mu. \tag{5.1}$$

Now, suppose $f \in \mathcal{L}_{\bar{\mathbb{R}}}$ and note that from the left side of equation (5.1)

$$\sup_{n\in\mathbb{N}} \int_{\mathbb{X}} f_n d\mu - \int_{\mathbb{X}} f_1 d\mu = \int_{\mathbb{X}} (f - f_1) d\mu, \text{ or}$$
$$\begin{aligned}\sup_{n\in\mathbb{N}} \int_{\mathbb{X}} f_n d\mu &= \int_{\mathbb{X}} f_1 d\mu + \int_{\mathbb{X}} (f - f_1) d\mu \\ &= \int_{\mathbb{X}} f_1 d\mu + \int_{\mathbb{X}} f d\mu - \int_{\mathbb{X}} f_1 d\mu = \int_{\mathbb{X}} f d\mu < \infty.\end{aligned}$$

If $\sup_{n\in\mathbb{N}} \int_{\mathbb{X}} f_n d\mu < \infty$, then from equation (5.1) we have $\int_{\mathbb{X}} (f - f_1) d\mu < \infty$ and since f_1 is integrable $f = (f - f_1) + f_1$ is integrable. Therefore,

$$\int f d\mu = \int (f - f_1) d\mu + \int f_1 du = \sup_{n\in\mathbb{N}} \int_{\mathbb{X}} f_n d\mu < \infty.$$

■

We now prove Markov's Inequality, which is useful in many settings is useful in many settings. Our immediate use of this inequality is on the following Theorem 5.3.

Theorem 5.2. *(Markov's Inequality) Let* $(\mathbb{X}, \mathcal{F}, \mu)$ *be a measure space and* $f \in \mathcal{L}_{\bar{\mathbb{R}}}$*. Then, for all* $E \in \mathcal{F}$ *and* $a > 0$

$$\mu\left(\{x : |f(x)| \geq a\} \cap E\right) \leq \frac{1}{a} \int_E |f| d\mu.$$

Proof. Note that, $aI_{\{x:|f(x)|\geq a\}\cap E} = aI_{\{x:|f(x)|\geq a\}} I_E \leq |f(x)| I_E$ and consequently, integrating both sides, $a\mu(\{x : |f(x)| \geq a\} \cap E) \leq \int_E |f| d\mu$. Therefore,

$$\mu(\{x : |f(x)| \geq a\} \cap E) \leq \frac{1}{a} \int_E |f| d\mu.$$

■

Remark 5.1. *Note that if* $E = \mathbb{X}$ *we have* $\mu(\{x : |f(x)| \geq a\}) \leq \frac{1}{a} \int_{\mathbb{X}} |f| d\mu$*. When* $(\mathbb{X}, \mathcal{F}, \mu) = (\Omega, \mathcal{F}, P)$ *a probability space and* $f := X$ *a random variable, we have*

$$P(\{\omega : |X(\omega)| \geq a\}) \leq \frac{1}{a} E_P(|X|).$$

Also, note that if $f(\omega) := (X(\omega) - E_P(X))^2$ we have

$$P(\{\omega : (X(\omega) - E_P(X))^2 \geq a\}) = P(\{\omega : |X(\omega) - E_P(X)| \geq a^{1/2}\}) \leq \frac{1}{a}E_P((X - E_P(X))^2),$$

and letting $b = a^{1/2}$ we have

$$P(\{\omega : |X(\omega) - E_P(X)| \geq b\}) \leq \frac{1}{b^2}E_P((X - E_P(X))^2),$$

which is known as Chebyshev's Inequality.

Recall that for a measure space $(\mathbb{X}, \mathcal{F}, \mu)$, N is a null set if $N \in \mathcal{F}$ and $\mu(N) = 0$. If a certain property $\mathcal{P}(x)$ that depends on $x \in \mathbb{X}$ holds for all $x \in \mathbb{X}$ except $x \in N_{\mathcal{P}} \subset N$, where N is a null set, we say that the property is true almost everywhere (ae) or almost surely (as). Note the set $N_{\mathcal{P}}$ where the property does not hold need not be a measurable set.

Theorem 5.3. *Let $(\mathbb{X}, \mathcal{F}, \mu)$ be a measure space and $f \in \mathcal{L}_{\bar{\mathbb{R}}}$. Then,*

1. *if N is a null set $\int_N f d\mu = 0$,*

2. *$\int_{\mathbb{X}} |f| d\mu = 0 \iff |f| = 0$ ae.*

Proof. 1. For $j \in \mathbb{N}$, let $f_j = \min\{|f|, j\}$ and note $0 \leq f_1 \leq f_2 \leq \cdots$ with $\lim_{j\to\infty} f_j = |f|$. Hence, by Theorem 4.5

$$\begin{aligned} 0 \leq \left|\int_N f d\mu\right| &= \left|\int_{\mathbb{X}} I_N f d\mu\right| \leq \int_{\mathbb{X}} I_N |f| d\mu \\ &= \lim_{j\to\infty} \int_{\mathbb{X}} I_N f_j d\mu = \lim_{j\to\infty} \int_{\mathbb{X}} I_N \min\{|f|, j\} d\mu \leq \lim_{j\to\infty} \int_{\mathbb{X}} j I_N d\mu \\ &= \lim_{j\to\infty} j \int_{\mathbb{X}} I_N d\mu = \lim_{j\to\infty} j\mu(N) = 0. \end{aligned}$$

2. $(\Leftarrow)$ $\int_{\mathbb{X}} |f| d\mu = \int_{\{|f|=0\}} |f| d\mu + \int_{\{|f|\neq 0\}} |f| d\mu = \int_{\{|f|\neq 0\}} |f| d\mu = 0$ by item 1.

$(\Rightarrow)$ Note that by the fact that μ is a measure

$$\begin{aligned} \mu(\{|f| > 0\}) &= \mu\left(\bigcup_{j\in\mathbb{N}} \{|f| \geq 1/j\}\right) \leq \sum_{j\in\mathbb{N}} \mu(\{|f| \geq 1/j\}) \\ &\leq \sum_{j\in\mathbb{N}} j \int_{\mathbb{X}} |f| d\mu = 0 \end{aligned}$$

by Markov's Inequality and the assumption that $\int_{\mathbb{X}} |f| d\mu = 0$. ■

Remark 5.2. *1. If $f, g \geq 0$ are measurable, integrable and $f = g$ μ-ae then $\int_{\mathbb{X}} f d\mu = \int_{\{x: f(x) \neq g(x)\}} f d\mu + \int_{\{x: f(x) = g(x)\}} f d\mu$. But by Theorem 5.3.1, the first integral in this sum is equal to zero. Consequently, $\int_{\mathbb{X}} f d\mu = \int_{\{x: f(x) = g(x)\}} f d\mu = \int_{\{x: f(x) = g(x)\}} g d\mu = \int_{\{x: f(x) \neq g(x)\}} g d\mu + \int_{\{x: f(x) = g(x)\}} g d\mu = \int_{\mathbb{X}} g d\mu$.*

2. If $f \in \mathcal{L}_{\bar{\mathbb{R}}}$ and $f = g$ μ-ae then $g \in \mathcal{L}_{\bar{\mathbb{R}}}$. To see this, note that $f = g$ μ-ae implies $f^+ = g^+$ and $f^- = g^-$ μ-ae. Using the previous remark on f^+ and f^- we have $\int_{\mathbb{X}} f^+ d\mu = \int_{\mathbb{X}} g^+ d\mu$ and $\int_{\mathbb{X}} f^- d\mu = \int_{\mathbb{X}} g^- d\mu$. Hence, $g \in \mathcal{L}_{\bar{\mathbb{R}}}$ and $\int_{\mathbb{X}} f d\mu = \int_{\mathbb{X}} g d\mu$.

3. If f is measurable and $0 \leq g \in \mathcal{L}_{\bar{\mathbb{R}}}$ with $|f| \leq g$ ae, then

$$f^+ \leq |f| \leq g \text{ ae and } f^- \leq |f| \leq g \text{ ae} .$$

Hence, $\int_{\mathbb{X}} f^+ d\mu \leq \int_{\mathbb{X}} g d\mu$, $\int_{\mathbb{X}} f^- d\mu \leq \int_{\mathbb{X}} g d\mu$ and f is integrable.

Theorem 5.4. *Let $f : (\mathbb{X}, \mathcal{F}, \mu) \to (\bar{\mathbb{R}}, \bar{\mathcal{B}})$ be integrable. Then f is real valued almost everywhere.*

Proof. Note that $N := \{x : |f(x)| = \infty\} = \{x : f(x) = \infty\} \cup \{x : f(x) = -\infty\} \in \mathcal{F}$. Also, if $\bigcap_{n \in \mathbb{N}} \{x : |f(x)| \geq n\} := \bigcap_{n \in \mathbb{N}} B_n$ with $B_1 \supset B_2 \supset \cdots$, then $\lim_{n \to \infty} B_n = \bigcap_{n \in \mathbb{N}} B_n = N$. Also, note that by Markov's Inequality and integrability of f

$$\mu(B_1) = \mu(\{x : |f(x)| \geq 1\}) \leq \int_{\mathbb{X}} |f| d\mu < \infty.$$

Hence, by continuity of measures from above, and Markov's Inequality

$$\mu(N) = \lim_{n \to \infty} \mu(B_n) = \lim_{n \to \infty} \mu(\{x : |f(x)| \geq n\}) \leq \lim_{n \to \infty} \frac{1}{n} \int_{\mathbb{X}} |f| d\mu = 0.$$

■

Theorem 5.5. *(Lebesgue's Dominated Convergence Theorem) Let* $(\mathbb{X}, \mathcal{F}, \mu)$ *be a measure space and* $\{f_n\}_{n\in\mathbb{N}}$ *be a sequence of integrable functions such that* $|f_n| \leq g$ *for all* n*, almost everywhere, where* g *is some integrable nonnegative function. If* $\lim_{n\to\infty} f_n(x) = f(x)$ *exists almost everywhere in* $\bar{\mathbb{R}}$*, then* f *is integrable and*

$$\lim_{n\to\infty} \int_{\mathbb{X}} f_n d\mu = \int_{\mathbb{X}} \lim_{n\to\infty} f_n d\mu := \int_{\mathbb{X}} f d\mu.$$

Proof. We start by observing that since the f_n and g are measurable, the set

$$\mathcal{N} = \{x : \lim_{n\to\infty} f_n(x) \text{ does not exist}\} \cup \left(\bigcup_{n\in\mathbb{N}} \{x : |f_n(x)| > g(x)\} \right)$$

is measurable and $\mu(\mathcal{N}) = 0$. Thus, we proceed by taking $\mathcal{N} = \emptyset$ as it does not contribute to any of the integrals in the proof of the theorem. By the point wise limit of the sequence f_n, for any $\epsilon > 0$ there exists $N_{(\epsilon,x)} \in \mathbb{N}$ such that for all $n > N_{(\epsilon,x)}$

$$\begin{aligned} |f| &= |f - f_n + f_n| \leq |f_n| + |f - f_n| \\ &\leq g + |f - f_n| \text{ by } |f_n| < g \\ &\leq g + \epsilon. \end{aligned}$$

Therefore, $\int_{\mathbb{X}} f d\mu < \infty$ provided $g \in \mathcal{L}_{\bar{\mathbb{R}}}(\mu)$. Also, $|f_n| \leq g \iff -g \leq f_n \leq g$. Hence, $f_n + g \geq 0$. By Fatou's Lemma,

$$\begin{aligned} \int \liminf_{n\to\infty} (f_n + g) d\mu = \int (f + g) d\mu &\leq \liminf_{n\to\infty} \int (f_n + g) d\mu \\ &= \liminf_{n\to\infty} \int f_n d\mu + \int g d\mu. \end{aligned}$$

Therefore,

$$\int f d\mu \leq \liminf_{n\to\infty} \int f_n d\mu. \tag{5.2}$$

Also, $g - f_n \geq 0$ and again by Fatou's Lemma,

$$\begin{aligned} 0 \leq \int \liminf_{n\to\infty}(g - f_n)d\mu &= \int g d\mu - \int f d\mu \\ &\leq \liminf_{n\to\infty} \int (g - f_n)d\mu \\ &= \int g d\mu + \liminf_{n\to\infty} - \int f_n d\mu \\ &= \int g d\mu - \limsup_{n\to\infty} \int f_n d\mu. \end{aligned}$$

The second inequality together with the last equality imply that

$$\int f d\mu \geq \limsup_{n\to\infty} \int f_n d\mu. \tag{5.3}$$

Combining (5.2) and (5.3) completes the proof. ■

We now consider a measurable function that is indexed by a parameter $\theta \in (a, b)$ for $a < b$. As such, we define $f(x, \theta) : (\mathbb{X}, \mathcal{F}, \mu) \times (a, b) \to (\mathbb{R}, \mathcal{B})$ where f is measurable for all $\theta \in (a, b)$.

Theorem 5.6. *Let $f(x, \theta) : (\mathbb{X}, \mathcal{F}, \mu) \times (a, b) \to (\mathbb{R}, \mathcal{B})$ where f is measurable and $f \in \mathcal{L}_{\mathbb{R}}$ for all $\theta \in (a, b)$. Also, assume that $f(x, \theta)$ is continuous for every $x \in \mathbb{X}$ and $|f(x, \theta)| \leq g(x)$ for all $(x, \theta) \in \mathbb{X} \times (a, b)$ and some nonnegative integrable function g. Then, the function $h : (a, b) \to \mathbb{R}$ given by*

$$h(\theta) := \int_{\mathbb{X}} f(x, \theta)d\mu$$

is continuous.

Proof. The function h is well defined because of integrability of $f(x, \theta)$. It suffices to show that for any sequence $\{\theta_n\}_{n\in\mathbb{N}} \subset (a, b)$ such that $\theta_n \to \theta$ we have $h(\theta_n) \to h(\theta)$. By continuity of $f(x, \theta)$, for every x, we have $f(x, \theta_n) \to f(x, \theta)$ and $|f(x, \theta_n)| \leq g(x)$. By Lebesgue's Dominated Convergence Theorem,

$$\lim_{n\to\infty} h(\theta_n) = \int_{\mathbb{X}} \lim_{n\to\infty} f(x, \theta_n)d\mu = \int_{\mathbb{X}} f(x, \theta)d\mu = h(\theta).$$

■

Theorem 5.7. *Let $f(x,\theta) : (\mathbb{X}, \mathcal{F}, \mu) \times (a,b) \to (\mathbb{R}, \mathcal{B})$ where f is measurable and $f \in \mathcal{L}_{\mathbb{R}}$ for all $\theta \in (a,b)$. Also, assume that $f(x,\theta)$ is differentiable on (a,b) for every $x \in \mathbb{X}$ and $|\frac{d}{d\theta} f(x,\theta)| \leq g(x)$ for all $(x,\theta) \in \mathbb{X} \times (a,b)$ and some nonnegative integrable function g. Then, the function $h : (a,b) \to \mathbb{R}$ given by*

$$h(\theta) := \int_{\mathbb{X}} f(x,\theta) d\mu$$

is differentiable and its derivative is given by

$$\frac{d}{d\theta} h(\theta) = \int_{\mathbb{X}} \frac{d}{d\theta} f(x,\theta) d\mu.$$

Proof. For θ, $\theta_n \in (a,b)$ with $\theta_n \to \theta$ as $n \to \infty$ and $\theta_n \neq \theta$ for all n

$$\frac{d}{d\theta} f(x,\theta) = \lim_{n\to\infty} \frac{f(x,\theta_n) - f(x,\theta)}{\theta_n - \theta}$$

for all $x \in \mathbb{X}$ and consequently $\frac{d}{d\theta} f(x,\theta)$ is measurable. By the Mean Value Theorem, there exists $\theta_{c,n} = c\theta_n + (1-c)\theta$ with $c \in (0,1)$ such that $f(x,\theta_n) - f(x,\theta) = \frac{d}{d\theta} f(x,\theta_{c,n})(\theta_n - \theta)$. Consequently,

$$\left| \frac{f(x,\theta_n) - f(x,\theta)}{\theta_n - \theta} \right| = \left| \frac{d}{d\theta} f(x,\theta_{c,n}) \right| \leq g(x),$$

and since g is integrable and f is measurable by assumption, we have by Theorem 4.8 that $\frac{f(x,\theta_n) - f(x,\theta)}{\theta_n - \theta}$ is integrable. Thus,

$$\frac{h(\theta_n) - h(\theta)}{\theta_n - \theta} = \int_{\mathbb{X}} \frac{f(x,\theta_n) - f(x,\theta)}{\theta_n - \theta} d\mu.$$

Hence, by the Lebesgue's Dominated Convergence Theorem

$$\lim_{n\to\infty} \frac{h(\theta_n) - h(\theta)}{\theta_n - \theta} = \frac{d}{d\theta} h(\theta) = \int_{\mathbb{X}} \lim_{n\to\infty} \frac{f(x,\theta_n) - f(x,\theta)}{\theta_n - \theta} d\mu = \int_{\mathbb{X}} \frac{d}{d\theta} f(x,\theta) d\mu.$$

■

5.2 $\mathcal{L}^p$ spaces

Definition 5.1. *The collection of measurable functions* $f : (\mathbb{X}, \mathcal{F}, \mu) \to (\mathbb{R}, \mathcal{B})$ *such that* $\int_{\mathbb{X}} |f|^p d\mu < \infty$ *for* $p \in [1, \infty)$ *is denoted by* $\mathcal{L}^p_{\mathbb{R}}$*,* $\mathcal{L}^p_{\mathbb{R}}(\mu)$ *or* $\mathcal{L}^p_{\mathbb{R}}(\mathbb{X}, \mathcal{F}, \mu)$*.*

Let $f, g \in \mathcal{L}^p_{\mathbb{R}}(\mathbb{X}, \mathcal{F}, \mu)$ and define $s : (\mathbb{X}, \mathcal{F}, \mu) \to (\mathbb{R}, \mathcal{B})$ as $s(x) = f(x) + g(x)$ for all $x \in \mathbb{X}$. Then, $|s(x)| \leq |f(x)| + |g(x)| \leq 2 \max\{|f(x)|, |g(x)|\}$ and

$$|s(x)|^p \leq 2^p \max\{|f(x)|, |g(x)|\}^p = 2^p \max\{|f(x)|^p, |g(x)|^p\} \leq 2^p(|f(x)|^p + |g(x)|^p).$$

Consequently, $\int_{\mathbb{X}} |s|^p d\mu \leq 2^p(\int_{\mathbb{X}} |f|^p d\mu + \int_{\mathbb{X}} |g|^p d\mu) < \infty$. Also, if $a \in \mathbb{R}$ and $m : (\mathbb{X}, \mathcal{F}, \mu) \to (\mathbb{R}, \mathcal{B})$ is defined as $m(x) = af(x)$ for all $x \in \mathbb{X}$, m is measurable and we have $|m(x)|^p = |a|^p |f(x)|^p$ and $\int_{\mathbb{X}} |m|^p d\mu = |a|^p \int_{\mathbb{X}} |f|^p d\mu < \infty$. Lastly, if we take $\theta(x) = 0$ for all $x \in \mathbb{X}$ to be the null vector in $\mathcal{L}^p_{\mathbb{R}}(\mathbb{X}, \mathcal{F}, \mu)$, then $\mathcal{L}^p_{\mathbb{R}}(\mathbb{X}, \mathcal{F}, \mu)$ is a vector space.

If $f \in \mathcal{L}^p_{\mathbb{R}}(\mathbb{X}, \mathcal{F}, \mu)$ we define the function $\| \cdot \|_p : \mathcal{L}^p_{\mathbb{R}}(\mathbb{X}, \mathcal{F}, \mu) \to [0, \infty)$ as $\|f\|_p = \left(\int_{\mathbb{X}} |f|^p d\mu\right)^{1/p}$ and prove the following inequality.

Theorem 5.8. *(Hölder's Inequality) If* $1 < p < \infty$*,* $p^{-1} + q^{-1} = 1$*,* $f \in \mathcal{L}^p_{\mathbb{R}}$*,* $g \in \mathcal{L}^q_{\mathbb{R}}$*, then* $fg \in \mathcal{L}_{\mathbb{R}}$ *and* $\int_{\mathbb{X}} |fg| d\mu \leq \|f\|_p \|g\|_q$*.*

Proof. If $\|f\|_p = 0$ then, by Theorem 5.3 $|f| = 0$ ae, so $|fg| = 0$ ae. Hence, $\int |fg| d\mu = 0$ and the inequality holds. Likewise for $\|g\|_q = 0$. So, assume $\|f\|_p, \|g\|_q \neq 0$. Let $x = f/\|f\|_p$, $y = g/\|g\|_q$ and note that $\|x\|_p = 1$ and $\|y\|_q = 1$. It suffices to prove $\int |xy| d\mu \leq 1$.

Now, note that for any $a, b > 0$ and $0 < \alpha < 1$,

$$a^\alpha b^{1-\alpha} \leq \alpha a + (1 - \alpha) b.$$

To see this, divide by b to obtain $(\frac{a}{b})^\alpha \leq \alpha \frac{a}{b} + (1 - \alpha)$. It suffices to show $u^\alpha \leq \alpha u + (1 - \alpha)$, for $u > 0$.

The inequality holds for $u = 1$. Now, $\frac{d}{du} u^\alpha = \alpha u^{\alpha - 1} = \alpha \frac{1}{u^{1-\alpha}}$. Since $\alpha \in (0, 1)$ we have that $u^{1-\alpha} < 1$ if $u < 1$. Consequently, in this case, $u^{\alpha - 1} > 1$ and $\frac{d}{du} u^\alpha > \alpha$. Also, using the same arguments, if $u > 1$ we have that $\frac{d}{du}u^\alpha < \alpha$. By the Mean Value Theorem, for $\lambda \in (0,1)$

$$u^\alpha - 1 = \alpha(\lambda u + (1-\lambda))^{\alpha-1}(u-1) < \alpha(u-1) \implies u^\alpha < 1 - \alpha + \alpha u \text{ if } u > 1.$$

Also,

$$u^\alpha - 1 = \alpha(\lambda u + (1-\lambda))^{\alpha-1}(u-1) < \alpha(u-1) \implies u^\alpha < 1 + \alpha u - \alpha, \text{ if } u < 1.$$

Thus, $u^\alpha \leq \alpha u + (1-\alpha)$ for $u > 0$.

Now, let $\alpha = 1/p$, $a(\omega) = |x(\omega)|^p$, $b(\omega) = |y(\omega)|^q$ and $1 - \alpha = 1/q$. Then,

$$(|x(\omega)|^p)^{1/p}(|y(\omega)|^q)^{1/q} \leq \alpha|x(\omega)|^p + (1-\alpha)|y(\omega)|^q, \text{ or}$$
$$|x(\omega)y(\omega)| \leq \alpha|x(\omega)|^p + (1-\alpha)|y(\omega)|^q.$$

Thus, integrating both sides of the inequality we obtain $\int |xy| d\mu \leq \alpha\|x\|_p + (1-\alpha)\|y\|_q = 1$.
■

Theorem 5.9. *(Minkowski-Riez Inequality) For $1 \leq p < \infty$, if f and g are in $\mathcal{L}^p$ we have $\|f+g\|_p \leq \|f\|_p + \|g\|_p$.*

Proof. By the triangle inequality

$$\begin{aligned}\|f+g\|_p^p &= \int |f+g||f+g|^{p-1} d\mu \leq \int \left(|f||f+g|^{p-1} + |g||f+g|^{p-1}\right) d\mu \\ &= \int |f||f+g|^{p-1} d\mu + \int |g||f+g|^{p-1} d\mu, \text{ and if } p = 1 \text{ the proof is complete.} \\ &\text{If } p > 1, \text{ by Hölder's Inequality} \\ &\leq \|f\|_p \||f+g|^{p-1}\|_q + \|g\|_p \||f+g|^{p-1}\|_q,\end{aligned}$$

where $1/p + 1/q = 1$ which implies $1/q = 1 - 1/p \implies q = \frac{p}{p-1}$. Thus,

$$\|f+g\|_p^p \leq \|f\|_p \||f+g|^{p/q}\|_q + \|g\|_p \||f+g|^{p/q}\|_q = (\|f\|_p + \|g\|_p)\||f+g|^{p/q}\|_q. \tag{5.4}$$

Now,

$$\||f+g|^{p/q}\|_q = \left(\int (|f+g|^{p/q})^q d\mu\right)^{1/q} = \left(\int |f+g|^p d\mu\right)^{1/q}$$
$$= \left(\int |f+g|^p d\mu\right)^{\frac{p-1}{p}} = \|f+g\|_p^{p-1}$$

Using this in inequality (5.4) we obtain $\|f+g\|_p^{p-(p-1)} = \|f+g\|_p \leq \|f\|_p + \|g\|_p$. ■

Remark 5.3. *1. The Minkowski-Riez Inequality and the fact that for $a \in \mathbb{R}$, $\|af\|_p = |a|\|f\|_p$ and $\|f\|_p \geq 0$ shows that $\|\cdot\|_p$ has almost all of the properties of a norm. The exception is that $\|f\|_p = 0$ does not imply that $f(x) = 0$ for all $x \in \mathbb{X}$. It only implies that $f(x) = 0$ almost everywhere.*

2. $f, g \in \mathcal{L}^p_{\mathbb{R}}(\mathbb{X}, \mathcal{F}, \mu)$ are taken to be equivalent if they differ at most on a set of μ-measure zero (null set), i.e., $f \sim g$ if $\{x : f(x) \neq g(x)\}$ is a null set. Then, for every $f \in \mathcal{L}^p_{\mathbb{R}}(\mathbb{X}, \mathcal{F}, \mu)$ we can define an equivalence class (reflexive, symmetric and transitive) of $\mathcal{L}^p_{\mathbb{R}}$ functions induced by f, which will be denoted by $[f]_p$. The space of all equivalence classes $[f]_p$ of functions $f \in \mathcal{L}^p_{\mathbb{R}}$ is denoted by $L^p_{\mathbb{R}}$ with norm $\|[f]_p\|_p := \inf\{\|g\|_p : g \in \mathcal{L}^p_{\mathbb{R}}$ and $g \sim f\}$. $(L^p, \|f_{[p]}\|_p)$ is a norm vector space and in what follows we will dispense with these technicalities and identify $[f]_p$ with f.

A commonly encountered case, treated in the next theorem, has $p = 2$ and $X, Y : (\Omega, \mathcal{F}, P) \to (\mathbb{R}, \mathcal{B})$ being random variables such that $X, Y \in \mathcal{L}^2_{\mathbb{R}}(\Omega, \mathcal{F}, P)$.

Theorem 5.10. *Let $X, Y : (\Omega, \mathcal{F}, P) \to (\mathbb{R}, \mathcal{B})$ be random variables such that $X, Y \in \mathcal{L}^2(\Omega, \mathcal{F}, P)$.*

1. *$XY \in \mathcal{L}(\Omega, \mathcal{F}, P)$ and $|\int_\Omega XY dP| \leq \left(\int_\Omega X^2 dP\right)^{1/2} \left(\int_\Omega Y^2 dP\right)^{1/2}$,*
2. *If $X \in \mathcal{L}^2(\Omega, \mathcal{F}, P)$ then $X \in \mathcal{L}(\Omega, \mathcal{F}, P)$ and $\left(\int_\Omega X dP\right)^2 \leq \int_\Omega X^2 dP$,*
3. *$\mathcal{L}^2(\Omega, \mathcal{F}, P)$ is a vector space.*

Proof. 1. This is just a special case of Hölder's Inequality with $p = q = 2$. 3. follows from the comments after Definition 5.1. 2. Let $X \in \mathcal{L}^2$ and note that $I_\Omega \in \mathcal{L}^2$ with $\int_\Omega I_\Omega dP = \int_\Omega dP$. Then,

$$\left|\int_\Omega X I_\Omega dP\right| \leq \left(\int_\Omega X^2 dP\right)^{1/2} \left(\int_\Omega dP\right)^{1/2}.$$

Since $\int_\Omega dP = 1$, we have

$$\left|\int_\Omega X dP\right| \leq \left(\int_\Omega X^2 dP\right)^{1/2} \text{ or } \left(\int_\Omega X dP\right)^2 \leq \int_\Omega X^2 dP.$$

■

Remark 5.4. *If $X \in \mathcal{L}^2$ we define $V_P(X) = \int_\Omega (X - E_P(X))^2 dP = \int_\Omega X^2 dP - (\int_\Omega X dP)^2$ and call it the variance of X (under P).*

Theorem 5.11. *Let X be a random variable defined on the probability space $(\Omega, \mathcal{F}, P)$ taking values in $(\mathbb{R}, \mathcal{B})$ and $h : (\mathbb{R}, \mathcal{B}) \to (\mathbb{R}, \mathcal{B})$ be measurable.*

1. *$f := h \circ X$ is integrable in $(\Omega, \mathcal{F}, P)$ if, and only if, h is integrable in $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$, where P_X is as defined in Example 3.2.*

2. *$E_P(h(X)) := \int_\Omega f dP = \int_\mathbb{R} h dP_X$.*

Proof. First, let h be a non-negative simple function. Then we have that $f(\omega) = \sum_{j=0}^m y_j I_{A_j}(\omega)$ where $A_j \in \mathcal{F}$. Consequently,

$$\begin{aligned} I_P(f) = \int_\Omega f dP &= \sum_{j=0}^m y_j P(A_j) = \sum_{j=0}^m y_j P(X^{-1}(B_j)) \text{ where } B_j = \{x \in \mathbb{R} : h(x) = y_j\} \\ &= \sum_{j=0}^m y_j (P \circ X^{-1})(B_j) = \sum_{j=0}^m y_j P_X(B_j) = \int_\mathbb{R} h dP_X = I_{P_X}(h). \end{aligned}$$

Second, let $h \geq 0$, not necessarily simple. Then, by Theorem 4.4 there exists a sequence of increasing non-negative simple function ϕ_n such that $\phi_n \to h$ as $n \to \infty$. Hence, if we define $f_n(\omega) = \phi_n(X(\omega)) = (\phi_n \circ X)(\omega)$, it is a sequence of increasing simple function that converges to f.

$$\begin{aligned}\int_\Omega f dP = \int_\Omega (h \circ X) dP &= \int_\Omega \lim_{n\to\infty} (\phi_n \circ X) dP \\ &= \lim_{n\to\infty} \int_\Omega (\phi_n \circ X) dP \text{ by Beppo-Levi's Theorem} \\ &= \lim_{n\to\infty} \int_{\mathbb{R}} \phi_n dP_X \text{ by the first part of the argument for simple functions} \\ &= \int_{\mathbb{R}} h dP_X, \text{ by Beppo-Levi's Theorem.}\end{aligned}$$

This proves 2. for simple and non-negative h. If h takes values in $\mathbb{R}$, consider $|h|$ and let ϕ_n be a sequence of increasing non-negative simple function such that $\phi_n \to |h|$ as $n \to \infty$. Then, we have from above that

$$\int_\Omega |f| dP = \int_{\mathbb{R}} |h| dP_X.$$

But from Remark 4.5, if $|h|$ is integrable in $(\mathbb{R}, \mathcal{B}, P_X)$ then h is integrable in $(\mathbb{R}, \mathcal{B}, P_X)$, establishing 1. Now, for arbitrary h we can prove the rest of part 2 by applying the same arguments to h^+ and h^- and using the fact that $h = h^+ - h^-$. ■

Clearly, taking $h(x) = x$ in the previous theorem gives $E_P(X) := \int_\Omega X dP = \int_{\mathbb{R}} x dP_X(x)$ where in the last integral we emphasize that the "variable" in integration is taking values in $\mathbb{R}$. In this proof, there is no requirement that $P(\Omega) = 1$. Hence, we can take $(\Omega, \mathcal{F}, P)$ to be an arbitrary measure space.

Definition 5.2. *The density of a probability measure P_X associated with a random variable X defined on a probability space $(\Omega, \mathcal{F}, P)$ is a non-negative Borel measurable function f_X that satisfies*

$$P_X((-\infty, a]) = \int_{(-\infty,a]} f_X d\lambda = \int_{\mathbb{R}} I_{(-\infty,a]} f_X d\lambda$$

where λ is Lebesgue measure on $\mathbb{R}$.

Theorem 5.12. *f_X is a density $\iff$ $\int_{\mathbb{R}} f_X d\lambda = 1$, f_X is unique almost everywhere.*

Proof. ($\implies$) f_X a density implies $F_X(a) = P_X((-\infty, a]) = \int_{(-\infty,a]} f_X d\lambda$. $\lim_{a\to\infty} P_X((-\infty, a]) = 1 = \lim_{a\to\infty} \int_{(-\infty,a]} f_X d\lambda$, where the first equality follows from Definition 2.4 and continuity of probability measures.

($\impliedby$) Suppose f_X is a non-negative Borel measurable function such that $\int_{\mathbb{R}} f_X d\lambda = 1$. For all $A \in \mathcal{B}$, we put

$$P_X(A) = \int_A f_X d\lambda = \int_{\mathbb{R}} I_A f_X d\lambda.$$

By Theorem 4.11, P_X is a measure on $\mathcal{B}$ with $P_X(\mathbb{R}) = 1$, by assumption. Taking $A = (-\infty, a]$,

$$P_X((-\infty, a]) = \int_{(-\infty,a]} f_X d\lambda$$

and f_X is a density for F_X.

Now, suppose g_X is another density for F_X. Then, $P_X(A) = \int_A g_X d\lambda = \int_{\mathbb{R}} g_X I_A d\lambda$. Let $A_n = \{x : g_X(x) \geq f_X(x) + 1/n\}$. For all $n \in \mathbb{N}$, $\int_{A_n} g_X d\lambda \geq \int_{A_n} (f_X + \frac{1}{n}) d\lambda = \int_{A_n} f_X d\lambda + \frac{1}{n}\lambda(A_n)$. Since $\int_{A_n} f_X d\lambda = \int_{A_n} g_X d\lambda$ it must be that $\lambda(A_n) = 0$.

Note that $A_1 \subset A_2 \subset \cdots$. $\lim_{n\to\infty} A_n = \cup_{n=1}^{\infty} A_n = A = \{x : g_X(x) > f_X(x)\}$ and $\lambda(A) = \lim_{n\to\infty} \lambda(A_n) = 0$. Similarly, we have $\lambda(B) = 0$ for $B = \{x : g_X(x) < f_X(x)\}$. So, $\lambda(\{x : g_X = f_X\}) = 1$. ■

Theorem 5.13. *Let $X : (\Omega, \mathcal{F}, P) \to (\mathbb{R}, \mathcal{B})$ be a random variable with density f_X and $h : (\mathbb{R}, \mathcal{B}) \to (\mathbb{R}, \mathcal{B})$ be a measurable function such that $\int_\Omega |h \circ X| dP < \infty$, i.e., $f = h \circ X$ is integrable. Then,*

$$\int_\Omega (h \circ X) dP = \int_{\mathbb{R}} h dP_X = \int_{\mathbb{R}} h(x) f_X(x) d\lambda(x)$$

Proof. Homework. ■

5.3 Stieltjes measure

Consider the measurable space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ and a measure $\mu : \mathcal{B}(\mathbb{R}) \to [0, \infty]$ such that for every $n \in \mathbb{N}$, $\mu([-n, n)) < \infty$. Define $F_\mu : \mathbb{R} \to \mathbb{R}$ as

$$F_\mu(x) = \begin{cases} \mu([0, x)) & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -\mu([x, 0)) & \text{if } x < 0 \end{cases} .$$

Theorem 5.14. $F_\mu : \mathbb{R} \to \mathbb{R}$ *is monotonically increasing and left-continuous.*

Proof. Given that $\mu([-n, n)) < \infty$, F_μ takes values in $\mathbb{R}$. First, we show that for all $x < x'$, $F_\mu(x) \leq F_\mu(x')$. There are three cases to be considered

1. $(0 \leq x < x')$: if $0 < x < x'$, $F_\mu(x') - F_\mu(x) = \mu([0, x')) - \mu([0, x))$. Since $[0, x') = [0, x) \cup [x, x')$, σ-additivity of μ gives $\mu([0, x')) = \mu([0, x)) + \mu([x, x'))$ or $\mu([x, x')) = \mu([0, x')) - \mu([0, x)) = F_\mu(x') - F_\mu(x) \geq 0$. If $x = 0$, $F_\mu(x') - F_\mu(0) = \mu([0, x')) \geq 0$.

2. $(x < 0 \leq x')$: If $x' > 0$, $F_\mu(x') - F_\mu(x) = \mu([0, x')) + \mu([x, 0)) \geq 0$. If $x' = 0$, $F_\mu(0) - F_\mu(x) = \mu([x, 0)) \geq 0$.

3. $(x < x' < 0)$: $F_\mu(x') - F_\mu(x) = -\mu([x', 0)) + \mu([x, 0))$. Since $[x, 0) = [x, x') \cup [x', 0)$, σ-additivity of μ gives $\mu([x, 0)) = \mu([x, x')) + \mu([x', 0))$ or $\mu([x, 0)) - \mu([x', 0)) = F_\mu(x') - F_\mu(x) = \mu([x, x')) \geq 0$.

Second, we must show that for $h_1 > 0$ and $h_1 \geq h_2 \geq h_3 \geq \cdots$ with $h_n \downarrow 0$ as $n \to \infty$, $\lim_{n\to\infty} F_\mu(x - h_n) = F_\mu(x)$ for all $x \in \mathbb{R}$. There are three cases to consider.

1. $(x > 0)$: Choose $h_1 \in (0, x)$ and define $A_n = [0, x - h_n)$. Then, $A_1 \subset A_2 \subset \cdots$ and $\lim_{n\to\infty} A_n = \bigcup_{n \in \mathbb{N}} A_n = [0, x)$. By continuity of measure from below,

$$\lim_{n\to\infty} F_\mu(x - h_n) = \lim_{n\to\infty} \mu([0, x - h_n)) = \mu([0, x)) = F_\mu(x).$$

2. ($x = 0$): Define $A_n = [-h_n, 0)$. Then, $A_1 \supset A_2 \supset \cdots$ and $\lim_{n\to\infty} A_n = \bigcap_{n\in\mathbb{N}} A_n = \emptyset$. By continuity of measures from above, and given that $\mu([-h_1, 0)) < \infty$,

$$\lim_{n\to\infty} F_\mu(-h_n) = \lim_{n\to\infty} \mu([-h_n, 0)) = \mu(\emptyset) = 0 = F_\mu(0).$$

3. ($x < 0$): Define $A_n = [x-h_n, 0)$. Then, $A_1 \supset A_2 \supset \cdots$ and $\lim_{n\to\infty} A_n = \cap_{n=1}^{\infty} A_n = [x, 0)$. By continuity of measures from above and given that $\mu([x-h_1, 0)) < \infty$,

$$\lim_{n\to\infty} F_\mu(x-h_n) = \lim_{n\to\infty} -\mu([x-h_n, 0)) = -\mu([x, 0)) = F_\mu(x).$$

■

Remark 5.5. *Note that F_μ is not right-continuous. Let $x \geq 0$, $h_n > 0$ and $A_n = [0, x+h_n)$. Then $A_1 \supset A_2 \supset \cdots$ and $\lim_{n\to\infty} A_n = \cap_{n=1}^{\infty} A_n = [0, x] = [0, x) \cup \{x\}$. Hence, $\lim_{n\to\infty} F_\mu(x+h_n) = \mu([0, x]) = F_\mu(x) + \mu(\{x\})$. Also, if $x < 0$, $0 < h_n < -x$ and $A_n = [x+h_n, 0)$. Then, $A_1 \subset A_2 \subset \cdots$ and $\lim_{n\to\infty} A_n = \bigcup_{n\in\mathbb{N}} A_n = (x, 0) = [x, 0) - \{x\}$. Hence, $\lim_{n\to\infty} F_\mu(x+h_n) = -\mu((x, 0)) = F_\mu(x) + \mu(\{x\})$. Hence, unless $\mu(\{x\}) = 0$ we have $\lim_{n\to\infty} F_\mu(x+h_n) \neq F_\mu(x)$. In fact, for any $x \in \mathbb{R}$, a point of continuity of F_μ,*

$$\begin{aligned}\mu(\{x\}) &= \mu\left(\bigcap_{n\in\mathbb{N}} \left[x, x+\frac{1}{n}\right)\right) = \lim_{n\to\infty} \mu\left(\left[x, x+\frac{1}{n}\right)\right) \\ &= \lim_{n\to\infty} F_\mu\left(x+\frac{1}{n}\right) - F_\mu(x) = 0 \text{ by right continuity of } F_\mu.\end{aligned}$$

Thus, F_μ is continuous at x if, and only if, $\mu(\{x\}) = 0$. Monotonically increasing, left-continuous functions are called Stieltjes functions.[1]

Remark 5.6. *It follows directly from the proof of Theorem 5.14 and Remark 5.5 that if $\mu((-n, n]) < \infty$. for every $n \in \mathbb{N}$ and we define $F_\mu : \mathbb{R} \to \mathbb{R}$ as*

$$F_\mu(x) = \begin{cases} \mu((0, x]) & \text{if } x > 0 \\ 0 & \text{if } x = 0 \\ -\mu((x, 0]) & \text{if } x < 0 \end{cases},$$

then $F_\mu(x)$ is monotonically increasing and right-continuous.

[1] In honor of Dutch mathematician Thomas Stieltjes (1856-1894).

Theorem 5.15. *Let F_μ be defined as in Theorem 5.14 and $\nu_{F_\mu}(([a,b)) = F_\mu(b) - F_\mu(a)$ for all $a \le b$, $a,b \in \mathbb{R}$. Then, ν_{F_μ} extends uniquely to a measure on $\mathcal{B}(\mathbb{R})$ and $\nu_{F_\mu} = \mu$.*

Proof. 1. Recall that $\mathcal{S} = \{[a,b) : a \le b,\, a,b \in \mathbb{R}\}$ is a semi-ring (if $a = b$, $[a,a) = \emptyset$). Given F_μ, we define $\nu_{F_\mu} : \mathcal{S} \to [0,\infty)$ as $\nu_{F_\mu}([a,b)) = F_\mu(b) - F_\mu(a)$ for all $a \le b$. Since F_μ is monotonically increasing, $F_\mu(b) - F_\mu(a) \ge 0$ and $\nu_{F_\mu}([a,a) = \emptyset) = F_\mu(a) - F_\mu(a) = 0$. Also, ν_{F_μ} is finitely additive since for $a < c < b$, we have that $[a,b) = [a,c) \cup [c,b)$ and $\nu_{F_\mu}([a,b)) = F_\mu(b) - F_\mu(a) = F_\mu(c) - F_\mu(a) + F_\mu(b) - F_\mu(c) = \nu_{F_\mu}([a,c)) + \nu_{F_\mu}([c,b))$. We now show that ν_{F_μ} is σ-additive, i.e., for $[a_n,b_n)$, $n \in \mathbb{N}$ a disjoint collection such that $[a,b) = \underset{n\in\mathbb{N}}{\cup}[a_n,b_n)$, we have $\nu_{F_\mu}([a,b)) = \sum\limits_{n\in\mathbb{N}} \nu_{F_\mu}([a_n,b_n))$. Fix ϵ_n, $\epsilon > 0$ and note that $(a_n - \epsilon_n, b_n) \supset [a_n,b_n)$. Hence, $\underset{n\in\mathbb{N}}{\cup}(a_n - \epsilon_n, b_n) \supset \underset{n\in\mathbb{N}}{\cup}[a_n,b_n) = [a,b) \supset [a, b-\epsilon]$. Since $\underset{n\in\mathbb{N}}{\cup}(a_n - \epsilon_n, b_n)$ is an open cover for the compact set $[a, b-\epsilon]$, by the Heine-Borel Theorem, there exists $N \in \mathbb{N}$ such that

$$\cup_{n=1}^{N}\, [a_n - \epsilon_n, b_n) \supset \cup_{n=1}^{N}(a_n - \epsilon_n, b_n) \supset [a, b-\epsilon] \supset [a, b-\epsilon). \tag{5.5}$$

Now, since $\cup_{n\in\mathbb{N}}[a_n,b_n) = [a,b)$ we have $\cup_{n=1}^{N}[a_n,b_n) \subset [a,b)$ and

$$\nu_{F_\mu}([a,b)) \ge \nu_{F_\mu}\left(\cup_{n=1}^{N}[a_n,b_n)\right) = \sum_{n=1}^{N} \nu_{F_\mu}\left([a_n,b_n)\right) \text{ by finite additivity.}$$

Hence, we have

$$\begin{aligned}
0 &\le \nu_{F_\mu}([a,b)) - \sum_{n=1}^{N} \nu_{F_\mu}\left([a_n,b_n)\right) \\
&= \nu_{F_\mu}([a,b-\epsilon)) + \nu_{F_\mu}([b-\epsilon,b)) - \sum_{n=1}^{N} \left(\nu_{F_\mu}([a_n - \epsilon_n, b_n)) - \nu_{F_\mu}([a_n - \epsilon_n, a_n))\right) \\
&= \nu_{F_\mu}([a,b-\epsilon)) - \sum_{n=1}^{N} \nu_{F_\mu}([a_n - \epsilon_n, b_n)) \text{ this term} < 0 \text{ by (5.5)} \\
&+ \nu_{F_\mu}([b-\epsilon,b)) + \sum_{n=1}^{N} \nu_{F_\mu}([a_n - \epsilon_n, a_n)) \\
&\le \nu_{F_\mu}([b-\epsilon,b)) + \sum_{n=1}^{N} \nu_{F_\mu}([a_n - \epsilon_n, a_n)) = F_\mu(b) - F_\mu(b-\epsilon) + \sum_{n=1}^{N}(F_\mu(a_n) - F_\mu(a_n - \epsilon_n)).
\end{aligned}$$

By left-continuity of F_μ, we can choose ϵ such that $F_\mu(b) - F_\mu(b-\epsilon) < \eta/2$ and ϵ_n such that $F_\mu(a_n) - F_\mu(a_n - \epsilon_n) < 2^{-n}\,\eta/2$. Hence,

$$0 \leq \nu_{F_\mu}([a,b)) - \sum_{n=1}^{N} \nu_{F_\mu}\left([a_n, b_n)\right) \leq \frac{\eta}{2}\left(1 + \sum_{n=1}^{N} 2^{-n}\right).$$

Letting $N \to \infty$ we have that $\nu_{F_\mu}([a,b)) = \sum_{n=1}^{\infty} \nu_{F_\mu}\left([a_n, b_n)\right)$.

Since ν_{F_μ} is a pre-measure on a semi-ring, by Carathéodory's Theorem, it has an extension to $\sigma(\mathcal{S}) = \mathcal{B}(\mathbb{R})$. Furthermore, since for $n \in \mathbb{N}$, $[-n, n) \uparrow \mathbb{R}$ and $\nu_{F_\mu}([-n,n)) = F_\mu(n) - F_\mu(-n) = \mu([0,n)) + \mu([-n,0))) < \infty$, this extension is unique.

To verify that $\nu_{F_\mu} = \mu$, it suffices to verify that $\nu_{F_\mu} = \mu$ on $\mathcal{S}$, since ν_{F_μ} extends uniquely to $\mathcal{B}(\mathbb{R})$. There are three cases:

Case 1 ($0 \leq a < b$): $\nu_{F_\mu}([a,b)) = F_\mu(b) - F_\mu(a) = \mu([0,b)) - \mu([0,a)) = \mu([0,a)) + \mu([a,b)) - \mu([0,a)) = \mu([a,b))$, since $[0,b) = [0,a) \cup [a,b)$,

Case 2 ($a < 0 < b$): $\nu_{F_\mu}([a,b)) = F_\mu(b) - F_\mu(a) = \mu([0,b)) + \mu([a,0)) = \mu([a,b))$, since $[a,b) = [a,0) \cup [0,b)$,

Case 3 ($a < b \leq 0$): $\nu_{F_\mu}([a,b)) = F_\mu(b) - F_\mu(a) = -\mu([b,0)) + \mu([a,0)) = \mu([a,b))$, since $[a,b) = [a,0) - [b,0)$, which completes the proof. ■

Example 5.1. *1. Let $\mu := \lambda$ the Lebesgue measure. Then,* $F_\lambda(x) = \begin{cases} x & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$ *and* $\nu_{F_\lambda}([a,b)) = F_\lambda(b) - F_\lambda(a) = b - a.$

2. Let $\mu := \delta_0$ the Dirac measure. Then, $F_{\delta_0}(x) = \begin{cases} \delta_0([0,x)) = 1 & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases}$ *and* $\nu_{F_{\delta_0}}([a,b)) = F_{\delta_0}(b) - F_{\delta_0}(a) = \begin{cases} 1 & \text{if } 0 \in [a,b) \\ 0 & \text{otherwise.} \end{cases}$

Remark 5.7. *Since $\mu = \nu_{F_\mu}$, for f measurable we can write $\int_{\mathbb{R}} f d\mu = \int_{\mathbb{R}} f d\nu_{F_\mu}$ or simply $\int_{\mathbb{R}} f d\mu = \int_{\mathbb{R}} f dF_\mu$.*

5.4 Abstract and Riemann integrals

The proper Riemann integral is defined for bounded functions over a compact interval $[a, b] \subset \mathbb{R}$ with $-\infty < a < b < \infty$. We start with the following definition.

Definition 5.3. *1. Let $[a, b] \subset \mathbb{R}$ and $a = t_0 < t_1 < \cdots < t_{k(P)-1} < t_{k(P)} = b$ where $k(P) \in \mathbb{N}$. The set $P = \{t_i\}_{i=0}^{k(P)}$ is called a partition of $[a, b]$.*

2. If P and P' are partitions of $[a, b]$ and $P \subset P'$, we say that P' is finer than P or a refinement of P.

3. The mesh of P is $m(P) := \max_{1 \le i \le k(P)} (t_i - t_{i-1})$.

If $f : [a, b] \to \mathbb{R}$ is bounded and $P = \{t_j\}_{j=0}^{k(P)}$ is a partition of $[a, b]$, let

$$m_j := \inf_{t_{j-1} \le x \le t_j} f(x) \text{ and } M_j := \sup_{t_{j-1} \le x \le t_j} f(x) \text{ for } j = 1, \ldots, k(P). \tag{5.6}$$

Then, the lower and upper Darboux sums are given by

$$S_P[f] := \sum_{j=1}^{k(P)} m_j (t_j - t_{j-1}) \text{ and } S^P[f] := \sum_{j=1}^{k(P)} M_j (t_j - t_{j-1}). \tag{5.7}$$

It is easy to show that if P' is finer than P, then

$$S_P[f] \le S_{P'}[f] \le S^{P'}[f] \le S^P[f].$$

Furthermore, since f is bounded on $[a, b]$, there exists $0 < C < \infty$ such that for all $x \in [a, b]$ we have that $|f(x)| \le C$. Consequently, for any P, $|S_P[f]| \le \sum_{j=1}^{k(P)} |m_j|(t_j - t_{j-1}) \le C \sum_{j=1}^{k(P)} (t_j - t_{j-1}) = C(b - a)$, and in similar fashion we conclude that $|S^P[f]| \le C(b - a)$. Since lower and upper Darboux sums are bounded collections of real numbers, they have suprema and infima. Hence, we provide the following definitions.

Definition 5.4. *Let $f : [a,b] \to \mathbb{R}$ be bounded. The lower and upper Riemann integrals of f over $[a,b]$ are defined as*

$$\underline{\int_a^b} f := \sup_P S_P[f] \text{ and } \overline{\int_a^b} f := \inf_P S^P[f], \tag{5.8}$$

where the supremum and the infimum are taken over all finite partitions of $[a,b]$.

Definition 5.5. *A bounded function $f : [a,b] \to \mathbb{R}$ is said to be Riemann integrable if its upper and lower integrals are the same. The common value is denoted by*

$$\int_a^b f(x)dx := \overline{\int_a^b} f = \underline{\int_a^b} f$$

and called the Riemann integral of f over the interval $[a,b]$. The collection of all Riemann integrable functions on $[a,b]$ will be denoted $\mathcal{R}[a,b]$.

We will show if f is measurable, $f \in \mathcal{R}[a,b] \implies f \in \mathcal{L}(\lambda)$ and $\int_{[a,b]} f d\lambda = \int_a^b f(x)dx$. To this end, we provide some auxiliary definitions and results.

Lemma 5.1. *Let $P = \{t_0, \cdots, t_{k(P)}\}$ be a partition of $[a,b]$, $\sigma_f^P(x) := \sum_{i=1}^{k(P)} m_i I_{[t_{i-1},t_i)}$ and $\Sigma_f^P(x) := \sum_{i=1}^{k(P)} M_i I_{[t_{i-1},t_i)}$. Then, if $P_1 \subset P_2 \subset \cdots$ the sequences $\{\sigma_f^{P_j}(x)\}_{j\in\mathbb{N}}$ and $\{\Sigma_f^{P_j}(x)\}_{j\in\mathbb{N}}$ are respectively monotonically increasing and decreasing in $[a,b)$. Furthermore, for all $x \in [a,b)$, $\sigma_f^P(x) \leq f(x) \leq \Sigma_f^P(x)$.*

Proof. First, we note that $[a,b) = \bigcup_{i=1}^{k(P)} [t_{i-1}, t_i)$. Consider $P' = P \cup \{t\}$ and, without loss of generality, assume $t \in [t_0, t_1)$. Then,

$$\sigma_f^{P'}(x) := \sum_{i=2}^{k(P)} m_i I_{[t_{i-1},t_i)} + \inf_{t_0 \leq x < t} f(x) I_{[t_0,t)} + \inf_{t \leq x < t_1} f(x) I_{[t,t_1)}.$$

Since $m_1 \leq \inf_{t_0 \leq x < t_1} f(x) = \min\left\{\inf_{t_0 \leq x < t} f(x), \inf_{t \leq x < t_1} f(x)\right\}$,

$$\sigma_f^{P'}(x) \geq \sum_{i=2}^{k(P)} m_i I_{[t_{i-1},t_i)} + m_1(I_{[t_0,t)} + I_{[t,t_1)}) = \sum_{i=2}^{k(P)} m_i I_{[t_{i-1},t_i)} + m_1(I_{[t_0,t)\cup[t,t_1)}) = \sigma_f^P(x).$$

Similarly, noting that $M_1 \geq \sup_{t_0 \leq x < t_1} f(x) = \max\left\{\sup_{t_0 \leq x < t} f(x), \sup_{t \leq x < t_1} f(x)\right\}$, we have $\Sigma_f^{P'}(x) \leq \Sigma_f^P(x)$. Hence, associated with partitions $P_1 \subset P_2 \subset \cdots$ there exist increasing and decreasing monotone sequences of simple functions given by $\{\sigma_f^{P_j}(x)\}_{j \in \mathbb{N}}$ and $\{\Sigma_f^{P_j}(x)\}_{j \in \mathbb{N}}$. Note that by definition, if $x \in [a,b)^c$, $\sigma_f^{P_j}(x) = \Sigma_f^{P_j}(x) = 0$.

Furthermore, since $m_i \leq f(x) \leq M_i$ for any $x \in [t_{i-1}, t_i)$ we have that for any P and for all $x \in \mathbb{R}$

$$I_{[t_i-1,t_i)}(x) m_i \leq f(x) I_{[t_i-1,t_i)}(x) \leq M_i I_{[t_i-1,t_i)}(x).$$

Note that if $x \notin [t_{i-1}, t_i)$, $I_{[t_i-1,t_i)}(x) = 0$ and both inequalities hold with equality. Hence, we have

$$\sigma_f^P(x) \leq f(x) \sum_{i=1}^{k(P)} I_{[t_{i-1},t_i)} \leq \Sigma_f^P(x), \text{ for all } x \in \mathbb{R}. \tag{5.9}$$

But since $\{[t_{i-1}, t_i)\}_{i=1}^{k(P)}$ is a collection of disjoint sets, $\sum_{i=1}^{k(P)} I_{[t_{i-1},t_i)} = 1$ if $x \in [a,b)$, and $\sum_{i=1}^{k(P)} I_{[t_{i-1},t_i)} = 0$ if $x \in [a,b)^c$. Thus,

$$\sigma_f^P(x) \leq f(x) \leq \Sigma_f^P(x) \text{ for all } x \in [a,b),$$

and all members of the inequalities are zero if $x \in [a,b)^c$.■

Remark 5.8. 1. *We note that*

$$|\sigma_f^P(x)| \leq \sum_{i=1}^{k(P)} |m_i| I_{[t_{i-1},t_i)} = \begin{cases} |m_i|, & \text{if } x \in [t_{i-1}, t_i) \text{ for some } i \\ 0, & \text{if } x \in [a,b)^c \end{cases},$$

and consequently $|\sigma_f^P(x)| \leq C$. *In similar fashion, we have that* $|\Sigma_f^P(x)| \leq C$. *Since, bounded increasing (decreasing) sequences converge do their supremum (infimum), for* $P_1 \subset P_2 \subset \cdots$ *we put*

$$\sigma_f(x) := \lim_{i \to \infty} \sigma_f^{P_i}(x) = \sup_{i \in \mathbb{N}} \sigma_f^{P_i}(x) \text{ and } \Sigma_f(x) := \lim_{i \to \infty} \Sigma_f^{P_i}(x) = \inf_{i \in \mathbb{N}} \Sigma_f^{P_i}(x).$$

2. *Since* $[t_{i-1}, t_i)$ *is a measurable set for any* i, *i.e.,* $[t_{i-1}, t_i) \in \mathcal{B}_{\mathbb{R}}$, *both* σ_f^P *and* Σ_f^P *are measurable functions for any* P. *In addition,*

$$\int_{[a,b)} \sigma_f^P d\lambda = \int_{\mathbb{R}} \sum_{i=1}^{k(P)} m_i I_{[t_{i-1},t_i)} I_{[a,b)} d\lambda = \sum_{i=1}^{k(P)} m_i \int_{\mathbb{R}} I_{[t_{i-1},t_i)} d\lambda = \sum_{i=1}^{k(P)} m_i (t_i - t_{i-1}) = S_P[f].$$

Similarly, $\int_{[a,b)} \Sigma_f^P d\lambda = S^P[f]$.

Theorem 5.16. *Let f be measurable and $f \in \mathcal{R}[a,b]$. Then, $f \in \mathcal{L}$ and $\int_{[a,b]} f d\lambda = \int_a^b f(x)dx$.*

Proof. Since $f \in \mathcal{R}[a,b]$, there exists $P_1 \subset P_2 \subset \cdots$ such that

$$\underline{\int_a^b} f := \sup\{S_{P_i}[f]\}_{i\in\mathbb{N}} = \lim_{i\to\infty} S_{P_i}[f] = \lim_{i\to\infty} S^{P_i}[f] = \inf\{S^{P_i}[f]\}_{i\in\mathbb{N}} := \overline{\int_a^b} f.$$

Now, from Lemma 5.1 and Remark 5.8.1,

$$\sigma_f(x) := \lim_{i\to\infty} \sigma_f^{P_i}(x) \leq f(x) \leq \lim_{i\to\infty} \Sigma_f^{P_i}(x) := \Sigma_f(x).$$

By Lebesgue's monotone convergence theorem, given that $\sup\limits_{i\in\mathbb{N}} \int \sigma_f^{P_i} d\lambda = \sup\{S_{P_i}[f]\}_{i\in\mathbb{N}} = \underline{\int_a^b} f < \infty$ and from Remark 5.8, we have

$$\underline{\int_a^b} f = \sup_{i\in\mathbb{N}} \int \sigma_f^{P_i} d\lambda = \int \sup_{i\in\mathbb{N}} \sigma_f^{P_i} d\lambda = \int \sigma_f d\lambda. \tag{5.10}$$

Following the same argument, we have

$$\overline{\int_a^b} f = \int \Sigma_f d\lambda. \tag{5.11}$$

By Riemann integrability of f and (5.10) and (5.11), we have that $\int \sigma_f d\lambda = \int \Sigma_f d\lambda$. Consequently,

$$\int (\Sigma_f - \sigma_f) d\lambda = 0.$$

Now, note $\Sigma_f - \sigma_f \geq 0$, hence by Theorem 5.3.2, $\Sigma_f = \sigma_f$ λ-almost everywhere. Given that

$$\{x : f(x) \neq \Sigma_f(x)\} \cup \{x : f(x) \neq \sigma_f(x)\} \subset \{x : \sigma_f(x) < \Sigma_f(x)\} \in \mathcal{N}_\lambda$$

and that f, σ_f and Σ_f are measurable, $\{x : f(x) \neq \Sigma_f(x)\}$, $\{x : f(x) \neq \sigma_f(x)\} \in \mathcal{B}$. Hence, we have that $f = \sigma_f = \Sigma_f(x)$ almost everywhere, and since σ_f (and Σ_f) is integrable so is f. Furthermore, by Remark 5.2.2 $\int f d\lambda = \int \sigma_f d\lambda = \int_a^b f(x)dx$.■

An extension of the proper Riemann integral to intervals of the form $[a, \infty)$ for $a \in \mathbb{R}$ is called the improper Riemann integral. As we will verify later, there are functions that are improperly Riemann integrable but are not Lebesgue integrable. Hence, although the Lebesgue integral extends proper Riemann integration, it does not include improper Riemann integration.

Definition 5.6. *If $f \in \mathcal{R}[a, b]$ for all $b \in (a, \infty)$ and if $\lim\limits_{b\to\infty} \int_a^b f(x)dx$ exists and is finite, we say that f is improperly Riemann integrable on $[a, \infty)$ and we write $f \in \mathcal{R}[a, \infty)$. This limit is denoted by $\int_a^\infty f(x)dx$. Similarly, if $f \in \mathcal{R}[a, b]$ for all $a \in (-\infty, b]$ and if $\lim\limits_{a\to-\infty} \int_a^b f(x)dx$ exists and is finite, we say that f is improperly Riemann integrable on $(-\infty, b]$ and we write $f \in \mathcal{R}(-\infty, b]$.*

Theorem 5.17. *Let $f : [0, \infty) \to \mathbb{R}$ be measurable and $f \in \mathcal{R}[0, N]$ for all $N \in \mathbb{N}$. Then,*

$$f \in \mathcal{L} \iff \lim_{N\to\infty} \int_0^N |f(x)|dx < \infty.$$

In this case, $\int_0^\infty f(x)dx = \int_{[0,\infty)} f d\lambda$.

Proof. ($\Longleftarrow$) $f \in \mathcal{R}[0, N] \implies f^+,\ f^- \in \mathcal{R}[0, N]$. Furthermore, since f is measurable, so are f^+, f^-. Then, by Theorem 5.16

$$\int_0^N f^+(x)dx = \int_{[0,N]} f^+ d\lambda = \int f^+ I_{[0,N]} d\lambda \text{ and } \int_0^N f^-(x)dx = \int_{[0,N]} f^- d\lambda = \int f^- I_{[0,N]} d\lambda.$$

Since, $\int_0^N |f(x)|dx = \int_0^N (f^+(x) + f^-(x))dx = \int_0^N f^+(x)dx + \int_0^N f^-(x)dx$, we have that

$$\lim_{N\to\infty} \int_0^N |f(x)|dx = \lim_{N\to\infty} \int_0^N f^+(x)dx + \lim_{N\to\infty} \int_0^N f^-(x)dx < \infty,$$

which implies that $\lim\limits_{N\to\infty} \int_0^N f^+(x)dx,\ \lim\limits_{N\to\infty} \int_0^N f^-(x)dx < \infty$. Consequently,

$$\lim_{N\to\infty} \int f^+ I_{[0,N]} d\lambda < \infty \text{ and } \lim_{N\to\infty} \int f^- I_{[0,N]} d\lambda < \infty.$$

Now, $\{f^+ I_{[0,N]}\}_{N\in\mathbb{N}}$ and $\{f^- I_{[0,N]}\}_{N\in\mathbb{N}}$ form increasing sequences of integrable function with $\lim_{N\to\infty} f^+ I_{[0,N]} = f^+$ and $\lim_{N\to\infty} f^- I_{[0,N]} = f^-$. Then, by the monotone convergence theorem,

$$\lim_{N\to\infty} \int f^+ I_{[0,N]} d\lambda = \int f^+ d\lambda \text{ and } \lim_{N\to\infty} \int f^- I_{[0,N]} d\lambda = \int f^- d\lambda.$$

Since $f^+, f^- \in \mathcal{L} \iff f \in \mathcal{L}$, the first part of the proof is complete.

$(\implies)$ $f \in \mathcal{R}[0,N] \implies |f| \in \mathcal{R}[0,N]$, and since f is measurable, by Theorem 5.16 $\int_0^N |f(x)|dx = \int_{[0,N]} |f| d\lambda = \int_{\mathbb{R}} |f| I_{[0,N]} d\lambda = \int_{\mathbb{R}} (f^+ I_{[0,N]} + f^- I_{[0,N]}) d\lambda$. By Theorem 4.10, for any $a \in \mathbb{R}$, $f \in \mathcal{L}$ implies $f^+, f^-, f I_{[0,a]}, f^+ I_{[0,a]}, f^- I_{[0,a]} \in \mathcal{L}$. Hence,

$$\int_0^N |f(x)|dx = \int_{\mathbb{R}} f^+ I_{[0,N]} d\lambda + \int_{\mathbb{R}} f^- I_{[0,N]} d\lambda.$$

By the monotone convergence theorem,

$$\lim_{N\to\infty} \int f^+ I_{[0,N]} d\lambda = \int_{[0,\infty)} f^+ d\lambda \text{ and } \lim_{N\to\infty} \int f^- I_{[0,N]} d\lambda = \int_{[0,\infty)} f^- d\lambda.$$

Hence,

$$\lim_{N\to\infty} \int_0^N |f(x)|dx = \int_{[0,\infty)} f^+ d\lambda + \int_{[0,\infty)} f^- d\lambda = \int_{[0,\infty)} |f| d\lambda < \infty.$$

Finally, note that since $f \in \mathcal{R}[0,N]$ for every $N \in \mathbb{N}$, $f, f^+, f^- \in \mathcal{R}[0,a_n]$ for all $a_n \in \mathbb{R}$ such that $a_n \uparrow \infty$. Hence, using the same arguments we have

$$\int_0^\infty f(x)dx := \lim_{n\to\infty} \int_0^{a_n} f(x)dx = \int_{[0,\infty)} |f| d\lambda < \infty.$$

■

We now give an example of a function that is improperly Riemann integrable, but is not Lebesgue integrable.

Example 5.2. *Let*

$$f(x) = \begin{cases} \frac{\sin x}{x} & \textit{for } x \in (0,\infty) \\ 1 & \textit{for } x = 0 \end{cases}$$

and for any $a > 0$ *consider* $\int_0^a f(x)dx$*. Note that, for any* $a > 0$ *there exists* $N_a \in \mathbb{N}$ *such that* $N_a\pi \le a < (N_a+1)\pi$ *and* $\int_0^a f(x)dx = \int_{N_a\pi}^a f(x)dx + \sum_{n=0}^{N_a-1}\int_{n\pi}^{(n+1)\pi} f(x)dx$*. Hence, by the triangle inequality, the fact that* $|\sin x| < 1$ *and given that* $a < (N_a+1)\pi$

$$\left|\int_0^a f(x)dx - \sum_{n=0}^{N_a-1}\int_{n\pi}^{(n+1)\pi} f(x)dx\right| \le \int_{N_a\pi}^a |f(x)|dx \le \int_{N_a\pi}^a \frac{1}{|x|}|dx \le \frac{1}{N_a\pi}(a-N_a\pi) < \frac{1}{N_a}.$$

Hence, as $N_a \to \infty$ *as* $a \to \infty$ *we can write* $\lim\limits_{a\to\infty}\int_0^a f(x)dx = \lim\limits_{N_a\to\infty}\sum_{n=0}^{N_a-1} a_n$ *where* $a_n = \int_{n\pi}^{(n+1)\pi} f(x)dx$*. Letting* $x = n\pi + y$ *and changing variables in integration, we have* $a_n = \int_0^\pi \frac{\sin(n\pi+y)}{i\pi+y}dy$*. Note that if* n *is even* $\frac{\sin(n\pi+y)}{i\pi+y} = \frac{\sin(y)}{i\pi+y}$ *and if* n *is odd* $\frac{\sin(n\pi+y)}{i\pi+y} = -\frac{\sin(y)}{i\pi+y}$*. Thus,*

$$\frac{\sin(n\pi+y)}{i\pi+y} = (-1)^n\frac{\sin(y)}{i\pi+y},\ \textit{and}\ a_n = (-1)^n\int_0^\pi \frac{\sin y}{n\pi+y}dy\ \textit{for}\ n = 0,1,\cdots,N-1.$$

Then, $|a_n| = \int_0^\pi \frac{\sin y}{n\pi+y}dy \le \int_0^\pi \frac{\sin y}{ny+y}dy = \frac{1}{n+1}\int_0^\pi \frac{\sin y}{y}dy$*. Also, since* $n < n+1$*,*

$$|a_n| \ge |a_{n=1}| := \int_0^\pi \frac{\sin y}{(n+1)\pi+y}dy \ge \int_0^\pi \frac{\sin y}{(n+1)\pi+\pi}dy = \frac{1}{(2+n)\pi}\int_0^\pi \sin y dy = \frac{2}{(2+n)\pi}.$$

Note that $\sin y, y \to 0$ *as* $y \downarrow 0$*. By L'Hôpital's rule,* $\lim\limits_{y\downarrow 0}\frac{\sin y}{y} = \lim\limits_{y\downarrow 0}\cos y = 1$*. Hence,* $f(x)$ *is continuous on* $[0,\pi]$*, and consequently* $f \in \mathcal{R}[0,\pi]$*. Then, setting* $C = \int_0^\pi \sin y dy$*, we have*

$$\frac{2}{\pi}\frac{1}{2+n} \le |a_{n+1}| \le |a_n| \le \frac{1}{n+1}C.$$

Now, $\sum_{n=0}^{N_a-1} a_n = \sum_{n=0}^{N_a-1}(-1)^n\int_0^\pi \frac{\sin y}{n\pi+y}dy = \sum_{n=0}^{N_a-1}(-1)^n|a_n|$*. Since* $|a_n| > 0$*,* $\sum_{n=0}^{\infty}(-1)^n|a_n|$ *is called an alternating series. This is a convergent series if* $|a_n|$ *is decreasing and* $\underset{n\to\infty}{|a_n|} = 0$*.*[2] *Hence,* $\lim\limits_{N_a\to\infty}\sum_{n=0}^{N_a-1} a_n$ *is finite and we can write* $\int_0^\infty \frac{\sin y}{y}dy = C < \infty$*, establishing the improper integrability of* $f(x)$ *over* $[0,\infty)$*.*

Now,

$$\int_{[0,a]} |f|d\lambda = \sum_{n=0}^{N_a-1}\int_{[n\pi,(n+1)\pi]}\left|\frac{\sin y}{y}\right|d\lambda + \int_{[N_a\pi,a]}\left|\frac{\sin y}{y}\right|d\lambda.$$

[2] See Leibniz convergence test (Apostol, 1974, p. 188).

As above,

$$\sum_{n=0}^{N_a-1} \int_{[n\pi,(n+1)\pi]} \left|\frac{\sin y}{y}\right| d\lambda = \sum_{n=0}^{N_a-1} \int_0^{\pi} \frac{\sin y}{n\pi + y} dy \geq \sum_{n=0}^{N_a-1} \frac{2}{\pi}\left(\frac{1}{2+n}\right),$$

which diverges. Hence, the Lebesgue integral does not exist.

5.5 Exercises

1. Prove Theorem 4.2.

2. Prove Theorem 4.10.

3. Use Markov's inequality to prove the following for $a > 0$ and $g : (0, \infty) \to (0, \infty)$ that is increasing:

$$P(|X(\omega)| \geq a) \leq \frac{1}{g(a)} \int g(|X|)dP$$

4. Let X be a random variable defined in the probability space $(\Omega, \mathcal{F}, P)$ with $E(X^2) < \infty$. Consider a function $f : \mathbb{R} \to \mathbb{R}$. What restrictions are needed on f to guarantee that $f(X)$ is a random variable with $E(f(X)^2) < \infty$?

5. Let $X : (\Omega, \mathcal{F}, P) \to (\mathbb{R}, \mathcal{B})$ be a random variable. Show that if $V(X) := E\left((X - E(X))\right)^2 = 0$ then X is a constant with probability 1.

6. Consider the following statement:*f is continuous almost everywhere if, and only if, it is almost everywhere equal to an everywhere continuous function.* Is this true or false? Explain, with precise mathematical arguments.

7. Adapt the proof of Lebesgue's Dominated Convergence Theorem in your notes to show that any sequence $\{f_n\}_{n\in\mathbb{N}}$ of measurable functions such that $\lim_{n\to\infty} f_n(x) = f(x)$ and $|f_n| \leq g$ for some g with g^p nonnegative and integrable satisfies

$$\lim_{n\to\infty} \int |f_n - f|^p d\mu = 0.$$

8. Let λ be the one-dimensional Lebesgue measure for the Borel sets of $\mathbb{R}$. Show that for every integrable function f, the function

$$g(x) = \int_{(0,x)} f(t) d\lambda, \text{ for } x > 0$$

is continuous.

9. Show that if X is a random variable with $E(|X|^p) < \infty$ then $|X|$ is almost everywhere real valued.

10. Suppose $X : (\Omega, \mathcal{F}, P) \to (\mathbb{R}, \mathcal{B})$ is a random variable with $E(|X|) < \infty$. Let $N \in \mathcal{F}$ be such that $P(N) = 0$ and define

$$Y(\omega) = \begin{cases} X(\omega) & \text{if } \omega \notin N \\ c & \text{if } \omega \in N \end{cases},$$

where $c \in \mathbb{R}$. Is Y integrable? Is $E(X) = E(Y)$?