Measure Theory

Additional reference:

Kirman, A.P. (1982). Measure theory with applications to economics. Chapter 5 in *Handbook of Mathematical Economics*, ed. K.J. Arrow and M. Intriligator, Vol. 1, 159–209.

1. Lebesgue measure

Henri L. Lebesgue (1875 - 1941)

Émile Borel (1871 - 1956)

- Let I be an interval in $\mathbb{R}$. It may be open: (a, b), closed: $[a, b]$, right-semiclosed: $(a, b]$, or left-semiclosed: $[a, b)$, with $a < b$.
- The length of I is

$$\ell(I) = b - a.$$

- **Definition.** The σ-algebra $\mathcal{B}$ of the Borel sets in $\mathbb{R}$ is the σ-algebra generated by the collection of all the intervals of $\mathbb{R}$. In other words, $\mathcal{B}$ is the smallest σ-algebra containing all the intervals of $\mathbb{R}$.
- *Note:* A single point in $\mathbb{R}$ is a Borel set since

$$\{a\} = \bigcap_{n=1}^{\infty} \left[a, a + \frac{1}{n}\right].$$

- Moreover, $(-\infty, b) = \bigcup_{n=1}^{\infty} (a-n, b)$,
  $(a, \infty) = \bigcup_{n=1}^{\infty} (a, b+n)$, $(-\infty, b] = \bigcup_{n=1}^{\infty} (a-n, b]$,
  $[a, \infty) = \bigcup_{n=1}^{\infty} [a, b+n)$, etc.

- Not all subsets of $\mathbb{R}$ are Borel sets.

- **Definition.** Let B be a Borel set in $\mathbb{R}$, $B \in \mathcal{B}$. The **Lebesgue measure** of the Borel set B is

$$\mu(B) = \inf_{C} \sum_{k} \ell(I_k),$$

  where C is a countable union of intervals in $\mathbb{R}$ covering B, i.e., $B \subset C = \bigcup_{k} I_k$, where $\{I_1, I_2, ...\}$ is a countable collection of intervals in $\mathbb{R}$.

- Let $I \subset \mathbb{R}^n$ be an interval (or rectangle or box), i.e.,

$$I = \left\{ x \in \mathbb{R}^n \;\middle|\; a_1 \underset{(<)}{\leq} x_1 \underset{(<)}{\leq} b_1, a_2 \underset{(<)}{\leq} x_2 \underset{(<)}{\leq} b_2, \ldots, a_n \underset{(<)}{\leq} x_n \underset{(<)}{\leq} b_n \right\}.$$

- The length (or the area or the volume) of I is $\ell(I) = \prod_{i=1}^{n} (b_i - a_i)$.
- **Definition.** The σ-algebra $\mathcal{B}(\mathbb{R}^n)$ of the Borel sets in $\mathbb{R}^n$ is the σ-algebra generated by the collection of all the intervals (or rectangles or boxes) of $\mathbb{R}^n$. In other words, $\mathcal{B}(\mathbb{R}^n)$ is the smallest σ-algebra containing all the intervals of $\mathbb{R}^n$.
- **Definition.** Let B be a Borel set in $\mathbb{R}^n$, $B \in \mathcal{B}(\mathbb{R}^n)$. The **Lebesgue measure** of the Borel set B is

$$\mu(B) = \inf_{C} \sum_{k} \ell(I_k),$$

where C is a countable union of intervals (or rectangles or boxes) in $\mathbb{R}^n$ covering B, i.e., $B \subset C = \bigcup_k I_k$, where $\{I_1, I_2, \ldots\}$ is a countable collection of intervals (or rectangles or boxes) in $\mathbb{R}^n$.

- *Note:* The σ-algebra $\mathcal{B}(\mathbb{R}^n)$ of the Borel sets in $\mathbb{R}^n$ is sometimes also defined as the σ-algebra generated by the collection of all the open (or all the closed) sets of $\mathbb{R}^n$.

- **Proposition.** If $A = \{x_1, x_2, ...\}$ is a countable subset of $\mathbb{R}$, then $\mu(A) = 0$.

- **Proof.**

$$0 \leq \mu(A) \leq \inf_{\varepsilon > 0} \sum_k \ell\left[x_k - \varepsilon, x_k + \varepsilon\right] = 0.$$

- *Example:* Let $\mathbb{Q}$ be the set of rational numbers in $[0, 1]$, then $\mu(\mathbb{Q}) = 0$ since $\mathbb{Q}$ is countable. Moreover, the Lebesgue measure of the set $\mathbb{Q}^c$ of irrational numbers in $[0, 1]$ is $\mu(\mathbb{Q}^c) = 1$.

- *Note:* If $A = \{x_1, x_2, ...\}$ is a countable subset of $\mathbb{R}^n$, then $\mu(A) = 0$.

- The Lebesgue measure is the unique measure on the Borel sets of $\mathbb{R}^n$ with the property that the measure of any interval (or rectangle or box) equals its length (or area or volume).

2. Finite and sigma-finite measures

- **Definition.** A measure on $(\Omega, \mathcal{F})$ is finite if $\mu(\Omega) < \infty$.
- The Lebesgue measure on $(\mathbb{R}, \mathcal{B})$ is not finite.
- A probability measure on $(\Omega, \mathcal{F})$ is finite since $P(\Omega) = 1 < \infty$.
- **Definition.** A measure on $(\Omega, \mathcal{F})$ is σ-finite if $\Omega = \bigcup_i A_i$, where $\{A_1, A_2, ...\}$ is a countable collection of elements of $\mathcal{F}$ with $\mu(A_i) < \infty$ for all i.
- *Examples:*
- (1) The Lebesgue measure on $(\mathbb{R}, \mathcal{B})$ is σ-finite.
- (2) Counting measure: $\mu(A) = \#A$. Obviously, the counting measure on $(\mathbb{R}, \mathcal{B})$ is not σ-finite.
- (3) The Lebesgue measure extended on $\left(\overline{\mathbb{R}}, \mathcal{B}\left(\overline{\mathbb{R}}\right)\right)$ is not σ-finite. Note that $\mathcal{B}\left(\overline{\mathbb{R}}\right)$ is the σ-algebra generated by the intervals in $\overline{\mathbb{R}}$.

3. Lebesgue-Stieltjes measures

- **Definition.** A Lebesgue-Stieltjes measure on $\mathbb{R}^n$ is a measure μ on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ such that $\mu(I) < \infty$ for every bounded interval (or rectangle or box) I in $\mathbb{R}^n$, $n = 1, 2, \ldots$
- The Lebesgue measure is a particular case of Lebesgue-Stieltjes measure.
- A Lebesgue-Stieltjes measure on $\mathbb{R}^n$ is σ-finite (but not necessarily finite).

4. Distribution functions

- **Definition.** A distribution function $F : \mathbb{R} \longrightarrow \mathbb{R}$ is a (weakly) increasing and right-continuous function.
- Increasing: $a < b \Longrightarrow F(a) \leq F(b)$.
- Right-continuous at x_0: $\lim_{x \to x_0^+} F(x) \equiv F\left(x_0^+\right) = F\left(x_0\right)$, where $x > x_0$.
- A function is right-continuous if it is right-continuous at all points of its domain.

- **Proposition.**

- (a) Let μ be a Lebesgue-Stieltjes measure on $\mathbb{R}$. Let F be defined up to an additive constant by $F(b) - F(a) = \mu\,(a, b]$ for all $a < b$. Then, F is a distribution function.

- (b) Let $F : \mathbb{R} \longrightarrow \mathbb{R}$ be a distribution function. Let μ be a measure on $(\mathbb{R}, \mathcal{B})$ such that $\mu\,(a, b] = F(b) - F(a)$, for all $a < b$, and $\mu(B) = \inf\limits_{C} \sum\limits_{k} \mu\left(\hat{I}_k\right)$ for all $B \in \mathcal{B}$, where C is a countable union of right-semiclosed intervals in $\mathbb{R}$ covering B, i.e., $B \subset C = \bigcup\limits_{k} \hat{I}_k$, where $\left\{\hat{I}_1, \hat{I}_2, ...\right\}$ is a countable collection of right-semiclosed intervals in $\mathbb{R}$. Then, μ is a Lebesgue-Stieltjes measure on $\mathbb{R}$.

- *Note:*

$$\lim_{x\to x_0^+} F(x) - F(x_0) = \lim_{x\to x_0^+} \mu\,(x_0, x] = \mu\,(\emptyset) = 0,$$

$$F(x_0) - \lim_{x\to x_0^-} F(x) = \lim_{x\to x_0^-} \mu\,(x, x_0] = \mu\,\{x_0\}\,.$$

- Thus, the distribution function $F : \mathbb{R} \longrightarrow \mathbb{R}$ is discontinuous at x_0 if and only if $\mu\,\{x_0\} > 0$.

- The σ-algebra generated by the right-semiclosed intervals is the σ-algebra of the Borel sets. To see this, observe that

$$(a, b) = \bigcup_{n=1}^{\infty} \left(a, b - \tfrac{1}{n}\right],\ \ (-\infty, b) = \bigcup_{n=1}^{\infty} (a - n, b)\,,$$

$$(a, \infty) = \bigcup_{n=1}^{\infty} (a, b + n)\,,\ \ [a, b] = \{(-\infty, a) \cup (b, \infty)\}^c\,,$$

$$[a, b) = (c, b) \cap [a, d]\,,\ \ (-\infty, b] = \bigcup_{n=1}^{\infty} (a - n, b]\,,$$

$$[a, \infty) = \bigcup_{n=1}^{\infty} [a, b + n)\,,\ \ \{a\} = \bigcap_{n=1}^{\infty} \left[a, a + \tfrac{1}{n}\right],\ \ \text{etc.}$$

- Usually in Statistics, when μ is a finite measure (for instance, when μ is a probability) on $(\mathbb{R}, \mathcal{B})$, we can make the following normalization without loss of generality:

$$\lim_{x\to-\infty} F(x) \equiv F(-\infty) = 0$$

so that

$$F(x) = F(x) - F(-\infty) = \mu\,(-\infty, x]\,.$$

- Recall that the Lebesgue measure is a Lebesgue-Stieltjes measure. The distribution function associated with the Lebesgue measure is $F(x) = x + K$, where K is an arbitrary constant. However, in this case, $F(-\infty) = -\infty$ since the Lebesgue measure is not a finite measure.

- Clearly, the Lebesgue measure satisfies

$$\mu\,(a, b] = F(b) - F(a) = b - a.$$

Lebesgue-Stieltjes measure of intervals:

$$\mu\{a\} = F(a) - \underbrace{F(a^-)}_{\lim_{x \to a^-} F(x)}.$$

$$\mu(a, b] = F(b) - F(a).$$

$$\begin{aligned} \mu[a, b] &= \mu(a, b] + \mu\{a\} = F(b) - F(a) + F(a) - F(a^-) \\ &= F(b) - F(a^-). \end{aligned}$$

$$\begin{aligned} \mu(a, b) &= \mu(a, b] - \mu\{b\} = F(b) - F(a) - F(b) + F(b^-) \\ &= F(b^-) - F(a). \end{aligned}$$

$$\begin{aligned} \mu[a, b) &= \mu(a, b) + \mu\{a\} = F(b^-) - F(a) + F(a) - F(a^-) \\ &= F(b^-) - F(a^-). \end{aligned}$$

- **Definition.** A measure on $(\Omega, \mathcal{F})$ is concentrated on $B \in \mathcal{F}$ if $\mu(B^c) = 0$.
- **Example:** Assume a Lebesgue-Stieltjes measure concentrated on the set of rational numbers in the interval $[a, b]$ such that $0 < \mu[a, b] < \infty$. Then, the corresponding distribution function $F : \mathbb{R} \longrightarrow \mathbb{R}$ is continuous at each irrational point of $[a, b]$, the function F must have discontinuity points on the interval $[a, b]$ and, thus, these discontinuities must occur at rational points.

- **Extension to** $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$:
- Right-semiclosed interval (or rectangle or box) $(a, b]$ in $\mathbb{R}^n$:

$$(a, b] = \{x \in \mathbb{R}^n \,|\, a_1 < x_1 \leq b_1, a_2 < x_2 \leq b_2, ..., a_n < x_n \leq b_n\}.$$

- **Definition.** A distribution function $F : \mathbb{R}^n \longrightarrow \mathbb{R}$ is a (weakly) increasing and right-continuous function.
- Increasing: $a < b \Longrightarrow F(a) \leq F(b)$, where a and b belong to $\mathbb{R}^n$.
- Right-continuous at x_0: $\lim\limits_{x \to x_0^+} F(x) \equiv F\left(x_0^+\right) = F\left(x_0\right)$, where $x > x_0 \in \mathbb{R}^n$.
- **Notation.** Let $x = (x_1, ..., x_n)^\mathsf{T}$ and $y = (y_1, ..., y_n)^\mathsf{T}$ belong to $\mathbb{R}^n$.

$x = y$ means that $x_i = y_i$ for $i = 1, ..., n$.

$x \geqq y$ means that $x_i \geq y_i$ for $i = 1, ..., n$.

$x > y$ means that $x_i \geq y_i$ for $i = 1, ..., n$ and $x \neq y$.

$x \gg y$ means that $x_i > y_i$ for $i = 1, ..., n$.

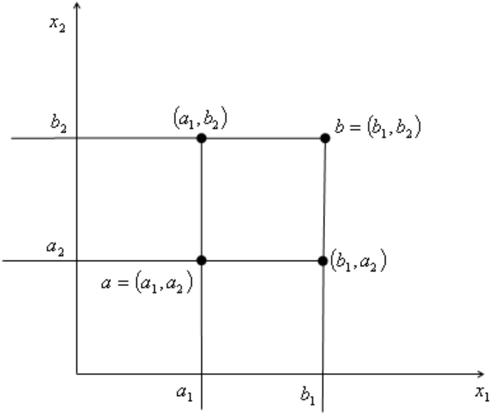
- The same relationships hold for row vectors.

- **From μ to F :**
- Assume that μ is a finite measure on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$. Define the corresponding distribution F as follows:

$$F(x) = \mu(-\infty, x] = \mu\underbrace{\{\omega \in \mathbb{R}^n \,| -\infty < \omega_i \leq x_i,\ i = 1, ..., n\}}_{(-\infty,x]\subset\mathbb{R}^n}.$$

- Note that $F(x) \to 0$ if at least one of the components x_i of $x \in \mathbb{R}^n$ tends to $-\infty$.
- Moreover $F(x) \to \mu(\mathbb{R}^n)$ if all the components x_i, $i = 1, ..., n$, of $x \in \mathbb{R}^n$ tend to ∞.
- Obviously, if μ is a probability on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ then $F(x) \to 1 = \mu(\mathbb{R}^n)$ if all the components x_i, $i = 1, ..., n$, of $x \in \mathbb{R}^n$ tend to ∞.

- **From** F **to** μ :
- $n = 2$



$$\mu(a, b] = F(b_1, b_2) - F(b_1, a_2) - F(a_1, b_2) + F(a_1, a_2).$$

- $n = 3$

$$\begin{aligned}\mu(a,b] &= F(b_1,b_2,b_3) - F(a_1,b_2,b_3) - F(b_1,a_2,b_3) \\ &\quad - F(b_1,b_2,a_3) + F(b_1,a_2,a_3) + F(a_1,b_2,a_3) \\ &\quad + F(a_1,a_2,b_3) - F(a_1,a_2,a_3).\end{aligned}$$

- We extend the formula to all the Borel sets of $\mathbb{R}^n$ as follows:

$$\mu(B) = \inf_C \sum_k \mu\left(\hat{I}_k\right) \text{ for all } B \in \mathcal{B},$$

where C is a countable union of right-semiclosed intervals (or rectangles or boxes) in $\mathbb{R}^n$ covering B, i.e., $B \subset C = \bigcup_k \hat{I}_k$, where $\left\{\hat{I}_1, \hat{I}_2, ...\right\}$ is a countable collection of right-semiclosed intervals (or rectangles or boxes) in $\mathbb{R}^n$.

5. Measurable and Borel measurable functions

- **Definition.** Consider the two measurable spaces $(\Omega, \mathcal{F})$ and $(\Omega', \mathcal{F}')$. The function $f : \Omega \longrightarrow \Omega'$ is measurable if $f^{-1}(A) \in \mathcal{F}$ for each $A \in \mathcal{F}'$.
- Notation for measurable function: $f : (\Omega, \mathcal{F}) \longrightarrow (\Omega', \mathcal{F}')$.
- $f : (\Omega, \mathcal{F}) \longrightarrow (\mathbb{R}, \mathcal{B})$ is a (real-valued) Borel measurable function.
- $f : (\Omega, \mathcal{F}) \longrightarrow \left(\overline{\mathbb{R}}, \mathcal{B}\right)$ is an (extended real-valued) Borel measurable function.
- $f : (\Omega, \mathcal{F}) \longrightarrow (\mathbb{R}^n, \mathcal{B})$ is a (real vector-valued) Borel measurable function.
- $f : (\Omega, \mathcal{F}) \longrightarrow \left(\overline{\mathbb{R}}^n, \mathcal{B}\right)$ is an (extended real vector-valued) Borel measurable function.

- Recall that we can write $\mathcal{B}\left(\overline{\mathbb{R}}\right)$, $\mathcal{B}\left(\mathbb{R}^n\right)$, $\mathcal{B}\left(\overline{\mathbb{R}}^n\right)$.
- Note that $f:(\Omega,\mathcal{F}) \longrightarrow \left(\overline{\mathbb{R}},\mathcal{B}\right)$ is Borel measurable iff

$$\{\omega \mid f(\omega) < c\} \in \mathcal{F}\ , \quad \text{for all } c \in \overline{\mathbb{R}}.$$

- or iff $\{\omega \mid f(\omega) > c\} \in \mathcal{F}$, or iff $\{\omega \mid f(\omega) \leq c\} \in \mathcal{F}$, or iff $\{\omega \mid f(\omega) \geq c\} \in \mathcal{F}$, for all $c \in \overline{\mathbb{R}}$.

6. Properties of measurable functions.

- **1.** If $f_n : (\Omega, \mathcal{F}) \longrightarrow \left(\overline{\mathbb{R}}, \mathcal{B}\right)$, for all n, and $\lim_{n\to\infty} f_n(\omega) = f(\omega)$ (i.e., $f_n(\omega) \to f(\omega)$) for all $\omega \in \Omega$, then $f : (\Omega, \mathcal{F}) \longrightarrow \left(\overline{\mathbb{R}}, \mathcal{B}\right)$. (The pointwise limit of Borel measurable functions is Borel measurable).
- **2.** If $f_1 : (\Omega, \mathcal{F}) \longrightarrow \left(\overline{\mathbb{R}}, \mathcal{B}\right)$ and $f_2 : (\Omega, \mathcal{F}) \longrightarrow \left(\overline{\mathbb{R}}, \mathcal{B}\right)$ are Borel measurable functions, so are $f_1 + f_2$, $f_1 - f_2$, $f_1 \cdot f_2$, and f_1 / f_2 (assuming these are well defined, i.e., we do not add $+\infty$ and $-\infty$, divide by 0, or divide ∞ by ∞).
- *Note:* $\infty - \infty$, $\dfrac{a}{0}$ for all $a \in \mathbb{R}$, and $\dfrac{\infty}{\infty}$ are not well-defined (or are "indeterminate" or "indefinite") in $\overline{\mathbb{R}}$.

- **3.** If $f : (\Omega_1, \mathcal{F}_1) \longrightarrow (\Omega_2, \mathcal{F}_2)$ and $g : (\Omega_2, \mathcal{F}_2) \longrightarrow (\Omega_3, \mathcal{F}_3)$, then $g(f) : (\Omega_1, \mathcal{F}_1) \longrightarrow (\Omega_3, \mathcal{F}_3)$. (The composition of measurable functions is measurable).
- **4.** $f : (\Omega, \mathcal{F}) \longrightarrow \left(\overline{\mathbb{R}}^n, \mathcal{B}\right)$ if and only if $f_i : (\Omega, \mathcal{F}) \longrightarrow \left(\overline{\mathbb{R}}, \mathcal{B}\right)$, for $i = 1, ..., n$, where $f = (f_1, ..., f_n)$.
- *Note:* $f_i = p_i(f)$, where the projection maps p_i, for $i = 1, ..., n$, satisfy $p_i : \left(\overline{\mathbb{R}}^n, \mathcal{B}\right) \longrightarrow \left(\overline{\mathbb{R}}, \mathcal{B}\right)$.

- **Definition.** The Borel measurable function $f : (\Omega, \mathcal{F}) \longrightarrow \left(\overline{\mathbb{R}}, \mathcal{B}\right)$ is simple if it takes on only finitely many distinct values $\{y_1, y_2, ..., y_n\}$.

- Therefore, $f : (\Omega, \mathcal{F}) \longrightarrow \left(\overline{\mathbb{R}}, \mathcal{B}\right)$ is simple if it can be written as

$$f(\omega) = \sum_{i=1}^{n} y_i \mathbb{I}_{A_i}(\omega), \quad \text{for all } \omega \in \Omega,$$

where $f(\omega) = y_i$ if $\omega \in A_i$ and $\mathbb{I}_{A_i}$ is the indicator function of A_i. That is,

$$\mathbb{I}_A(\omega) = \begin{cases} 1 \text{ if } \omega \in A \\ \\ 0 \text{ if } \omega \notin A. \end{cases}$$

- Note that $\{A_i\}_{i=1}^{n}$ is a collection of sets in $\mathcal{F}$ that form a partition of Ω.

- **5.** Every non-negative Borel measurable function $f : (\Omega, \mathcal{F}) \longrightarrow (\overline{\mathbb{R}}, \mathcal{B})$ is the limit of an increasing sequence of non-negative, finite-valued, simple functions f_n.

- **6.** Every Borel measurable function $f : (\Omega, \mathcal{F}) \longrightarrow (\overline{\mathbb{R}}, \mathcal{B})$ is the limit of a sequence of finite-valued simple functions f_n, with $|f_n| \leq |f|$ for all n.

7. Integral with respect to a measure and its properties

Arithmetic of $\overline{\mathbb{R}}$: if $a \in \mathbb{R}$, then

- $a + \infty = \infty$,
- $a - \infty = -\infty$,
- $a/\infty = a/\left(-\infty\right) = 0$,
- $a \cdot \infty = \infty$ if $a > 0$,
- $a \cdot \infty = -\infty$ if $a < 0$,
- $0 \cdot \infty = 0 \cdot \left(-\infty\right) = 0$,
- $\infty + \infty = \infty$,
- $-\infty - \infty = -\infty$.
- Recall that $\infty - \infty$, $\dfrac{a}{0}$ for all $a \in \mathbb{R}$, and $\dfrac{\infty}{\infty}$ are not well-defined (or are "indeterminate" or "indefinite") in $\overline{\mathbb{R}}$.

- **Definition of integral with respect to a measure μ on $(\Omega, \mathcal{F})$.**
- If $h : (\Omega, \mathcal{F}) \longrightarrow (\overline{\mathbb{R}}, \mathcal{B})$ is a simple Borel measurable function, we define

$$\int_{\Omega} h d\mu = \sum_{i=1}^{n} y_i \mu(A_i), \quad \text{where } h(\omega) = y_i \text{ if } \omega \in A_i, \tag{1}$$

provided $+\infty$ and $-\infty$ do not appear in the sum.

- If $h : (\Omega, \mathcal{F}) \longrightarrow (\overline{\mathbb{R}}, \mathcal{B})$ is a non-negative Borel measurable function, we define

$$\int_{\Omega} h d\mu = \sup_{s} \int_{\Omega} s d\mu, \tag{2}$$

where s is a simple Borel measurable function with $0 \leq s \leq h$.

- *Note:* the integral in (1) may be equal to $\pm\infty$, whereas that in (2) may be equal to $+\infty$

- If $h : (\Omega, \mathcal{F}) \longrightarrow (\overline{\mathbb{R}}, \mathcal{B})$ is a Borel measurable function, let

$$h^+ = \max(h, 0) \geq 0, \quad \text{(the positive part of } h\text{)}$$

$$h^- = \max(-h, 0) \geq 0, \quad \text{(the negative part of } h\text{)}$$

and observe that $h = h^+ - h^-$.

- We define the integral of h w.r.t. μ as follows:

$$\int_\Omega h d\mu = \int_\Omega h^+ d\mu - \int_\Omega h^- d\mu,$$

provided it is not of the form $\infty - \infty$.

- We could write $\int_\Omega h(\omega) d\mu(\omega)$ or $\int_\Omega h(\omega) \mu(d\omega)$.

- **Definition:** The function h is said to be μ-integrable (or integrable w.r.t. μ) if $\int_\Omega h d\mu$ is finite.
- If $\int_\Omega h d\mu$ is of the form $\infty - \infty$, we say that the integral $\int_\Omega h d\mu$ does not exist. Therefore, the condition of μ-integrability requires the previous existence of the integral.
- From now on, when we write an integral, it should be understood that the integral exists.
- Sometimes $\int_\Omega h d\mu$ is called "abstract Lebesgue integral".
- We define

$$\int_B h(\omega) d\mu(\omega) = \int_\Omega h(\omega) \mathbb{I}_B(\omega) d\mu(\omega), \quad \text{where } B \in \mathcal{F}.$$

- Let $f : (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)) \longrightarrow \left(\overline{\mathbb{R}}, \mathcal{B}\right)$. The Lebesgue integral of f is the integral of f with respect to Lebesgue measure and can be written as

$$\int_B f(x_1, ..., x_n) \, d(x_1, x_2, .., x_n), \quad \text{for } B \in \mathcal{B}(\mathbb{R}^n).$$

- **Properties of the integral** (for extended real-valued Borel measurable functions)**:**

- **1.**
$$\int_{\Omega} chd\mu = c\int_{\Omega} hd\mu, \quad \text{for any constant } c.$$

- **2.** If $g \geq h$, then
$$\int_{\Omega} gd\mu \geq \int_{\Omega} hd\mu.$$

- **3.**
$$\left|\int_{\Omega} hd\mu\right| \leq \int_{\Omega} |h|\, d\mu.$$

- **4.** If $\int_\Omega h d\mu$ exists, so does $\int_B h d\mu$ for all $B \in \mathcal{F}$.
- **5.** Define $\lambda(B) = \int_B h d\mu$ for all $B \in \mathcal{F}$. Then, λ is a signed measure on $(\Omega, \mathcal{F})$.
- **Proof.** Let $\{B_i\}$ be a countable collection of disjoint measurable sets, then

$$\lambda\left(\bigcup_i B_i\right) = \int_{\bigcup_i B_i} h d\mu = \sum_i \left(\int_{B_i} h d\mu\right) = \sum_i \lambda(B_i).$$

Moreover, $\lambda(\emptyset) = \int_{\{\emptyset\}} h d\mu = 0$. *Q.E.D.*

- **6.**
$$\int_\Omega \left(\sum_i f_i\right) d\mu = \sum_i \left(\int_\Omega f_i d\mu\right).$$
- **7.** h is integrable if and only if $|h|$ is integrable.
- Note that $|h| = h^+ + h^-$.
- **8.** If $|g| \leq h$ and h is integrable, then g is integrable.

- **Definition:** Let f and g be Borel measurable functions, then $f = g$ almost everywhere with respect to μ (or a.e. $[\mu]$) if

$$\mu\{f \neq g\} \equiv \mu\{\omega \in \Omega \mid f(\omega) \neq g(\omega)\} = 0.$$

- **9.** If $f = 0$ a.e. with respect to μ, then $\int_\Omega f d\mu = 0$.
- **10.** If $f = g$ a.e. $[\mu]$, then

$$\int_\Omega f d\mu = \int_\Omega g d\mu.$$

- **11.** If $h \geq 0$ and $\int_\Omega h d\mu = 0$, then $h = 0$ a.e. (with respect to μ).
- **12.** If h is integrable (with respect to μ), then h is finite a.e. (with respect to μ).
- **13.**

$$\mu(B) = \int_B 1 d\mu \equiv \int_B d\mu \equiv \int_\Omega \mathbb{I}_B(\omega) d\mu(\omega) \text{ for all } B \in \mathcal{F}.$$

8. Convergence of integrals

- **Definition (convergence almost everywhere)**. Let $f_1, f_2, ...,$ and f be extended real-valued Borel measurable functions on $(\Omega, \mathcal{F}, \mu)$. The sequence of functions $\{f_n\}$ converges to f almost everywhere with respect to the measure μ (or $f = \lim_{n\to\infty} f_n$ a.e. $[\mu]$ or $f_n \longrightarrow f$ a.e. $[\mu]$) if

$$\mu \left\{ \omega \in \Omega \;\middle|\; \lim_{n\to\infty} f_n(\omega) \neq f(\omega) \right\} = 0.$$

- **Fatou's lemma.** Let $f_1, f_2, \ldots$ be a sequence of non-negative a.e. $[\mu]$, extended real-valued Borel measurable functions on $(\Omega, \mathcal{F}, \mu)$ and

$$f = \liminf_{n\to\infty} f_n \equiv \lim_{n\to\infty} \left(\inf_{k\geq n} f_k \right) \quad \text{a.e. } [\mu].$$

Then the function f is Borel measurable and

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

where the integrals may be infinite.

- **Monotone Convergence Theorem (Levi).** Let $f_1, f_2, \ldots$ be a non-decreasing a.e. $[\mu]$ sequence of extended real-valued Borel measurable functions on $(\Omega, \mathcal{F}, \mu)$ and $f = \lim_{n\to\infty} f_n$ a.e. $[\mu]$ (or $f_n \longrightarrow f$ a.e. $[\mu]$). Then f is Borel measurable and

$$\lim_{n\to\infty} \int_\Omega f_n d\mu = \int_\Omega f d\mu \quad \left(\text{or } \int_\Omega f_n d\mu \longrightarrow \int_\Omega f d\mu \right),$$

where the integrals may be infinite.

- **Dominated Convergence Theorem (Lebesgue).** Let $f_1, f_2, \ldots, f$, and g be extended real-valued Borel measurable functions on $(\Omega, \mathcal{F}, \mu)$ with $|f_n| \leq g$ a.e. $[\mu]$ for all n, where g is μ-integrable, and $f = \lim_{n\to\infty} f_n$ a.e. $[\mu]$ (or $f_n \longrightarrow f$ a.e. $[\mu]$). Then f is μ-integrable and

$$\lim_{n\to\infty} \int_\Omega f_n d\mu = \int_\Omega f d\mu \quad \left(\text{or } \int_\Omega f_n d\mu \longrightarrow \int_\Omega f d\mu \right).$$

- Thus, under the assumptions of the previous two theorems, "the limit of integrals is equal to the integral of the limit."

9. The Lebesgue and the Lebesgue-Stieltjes integrals and their relation with the Riemann and the Riemann-Stieltjes integrals.

- Let $f : (\mathbb{R}, \mathcal{B}) \longrightarrow (\mathbb{R}, \mathcal{B})$
- Riemann integral of the function f on $[a, b]$:

$$\int_a^b f(x)dx.$$

- Lebesgue integral of the function f on $[a, b]$:

$$\int_{[a,b]} f(x)dx.$$

- *Note:*

$$\int_{[a,b]} f(x)dx = \int_{(a,b)} f(x)dx = \int_{(a,b]} f(x)dx = \int_{[a,b)} f(x)dx.$$

- Riemann-Stieltjes (or just Stieltjes) integral of the function f on $(a, b]$:

$$\int_a^b f(x) dF(x),$$

where F is a distribution function (i.e., an increasing and right-continuous function) on $\mathbb{R}$.

- Lebesgue-Stieltjes integral of the function f on $(a, b]$:

$$\int_{(a,b]} f dF \quad \text{or} \quad \int_{(a,b]} f(x) dF(x),$$

where we are integrating with respect to the Lebesgue-Stieltjes measure μ on $(a, b]$ associated with the distribution function F.

- Note that

$$\int_{[a,b]} f dF = \int_{(a,b]} f dF + f(a)\mu\{a\} = \int_{(a,b]} f dF + f(a)\left[F(a) - F(a^-)\right].$$

- See the handout for the relation between these integrals.
- The function defined on $[0,1]$ by $f(x) = 1$ if x is irrational, and $f(x) = 0$ if x is rational, is the standard example of a function that is Lebesgue integrable but not Riemann integrable. Its Lebesgue integral is $\int_{[0,1]} f dx = 1$.

10. Absolute continuity of measures

- **Definition.** Let $(\Omega, \mathcal{F}, \mu)$ be a measure space, $h : (\Omega, \mathcal{F}) \longrightarrow (\overline{\mathbb{R}}, \mathcal{B})$ a Borel measurable function such that $\int_\Omega h d\mu$ exists. Then the countably additive set function (or signed measure),

$$\lambda(A) = \int_A h d\mu, \quad \text{for all } A \in \mathcal{F},$$

is called the indefinite integral of h (with respect to μ).

- If μ is Lebesgue measure and $A = [a, x]$, then

$$\lambda [a, x] = \int_{[a,x]} h(t) dt$$

is the familiar indefinite integral of calculus.

- **Definition.** If μ is a measure on $(\Omega, \mathcal{F})$ and λ is a signed-measure on $(\Omega, \mathcal{F})$, we say that λ is absolutely continuous with respect to μ, $\lambda \ll \mu$, if $\mu(A) = 0$ implies that $\lambda(A) = 0$.

- **Proposition.** If λ is an indefinite integral of the Borel measurable function h w.r.t. μ, i.e.,

$$\lambda(A) = \int_A h d\mu, \quad \text{for all } A \in \mathcal{F},$$

  then $\lambda \ll \mu$.

11. Radon-Nikodym theorem.

- **Radon-Nikodym Theorem.** Let μ be a σ-finite measure on $(\Omega, \mathcal{F})$ and λ be a signed-measure on $(\Omega, \mathcal{F})$. Assume that $\lambda \ll \mu$. Then, there exists a function $g : (\Omega, \mathcal{F}) \longrightarrow (\overline{\mathbb{R}}, \mathcal{B})$ such that

$$\lambda(A) = \int_A g d\mu, \quad \text{for all } A \in \mathcal{F}.$$

  This Borel measurable function g is called the Radon-Nikodym derivative of λ w.r.t. μ, and is written $g = d\lambda / d\mu$. Finally, if h is another such function, then $g = h$ a.e. $[\mu]$.

- Therefore, if μ is σ-finite, $\lambda \ll \mu$ if and only if λ is an indefinite integral of some Borel measurable function h w.r.t. μ.

- If the signed-measure λ is finite then the function g should be finite a.e. $[\mu]$. Therefore, in this case, we can assume without loss of generality that the existing function g is finite, $g : (\Omega, \mathcal{F}) \longrightarrow (\mathbb{R}, \mathcal{B})$.

- *Note:* The assumption that μ is σ-finite is crucial.
- **Example:** Let μ be the counting measure on $(\mathbb{R}, \mathcal{B})$ and λ be the Lebesgue measure on $(\mathbb{R}, \mathcal{B})$.
- Obviously, $\mu(B) = 0$ implies that $\lambda(B) = 0$, i.e., $\lambda \ll \mu$.
- Assume that

$$\lambda(A) = \int_A g d\mu, \quad \text{for all } A \in \mathcal{B},$$

  for some function $g : (\mathbb{R}, \mathcal{B}) \longrightarrow \left(\overline{\mathbb{R}}, \mathcal{B}\right)$. Take $A = \{x\}$ so that $\lambda\{x\} = g(x)$. Since $\lambda\{x\} = 0$ for all $x \in \mathbb{R}$, it holds that $g(x) = 0$ for all $x \in \mathbb{R}$. Therefore, we obtain that $\lambda(A) = 0$ for all $A \in \mathcal{B}$, which is a contradiction.
- In this example the Radon-Nikodym theorem does not apply since the counting measure μ on $(\mathbb{R}, \mathcal{B})$ is not σ-finite.

12. Absolutely continuous functions

- **Definition.** The function $f : [a, b] \longrightarrow \mathbb{R}$ is absolutely continuous on $[a, b]$ if there exists a Borel measurable function $g : ([a, b], \mathcal{B}) \longrightarrow (\mathbb{R}, \mathcal{B})$, which is integrable w.r.t. Lebesgue measure, such that

$$f(x) - f(a) = \int_{[a,x]} g(t)dt, \quad \text{for } a \leq x \leq b.$$

- **Note:** In the previous definition we could have $g : ([a, b], \mathcal{B}) \longrightarrow (\overline{\mathbb{R}}, \mathcal{B})$, but g should be finite a.e. on $[a, b]$ w.r.t. Lebesgue measure since f is a real-valued function.

- **Theorem.** Let the function f be absolutely continuous on $[a, b]$, i.e.,

$$f(x) - f(a) = \int_{[a,x]} g(t)dt, \quad \text{for } a \leq x \leq b, \qquad (\star)$$

for some Borel measurable function g defined on $[a, b]$, then $f' = g$ a.e. w.r.t. Lebesgue measure.

- Therefore, f is absolutely continuous on $[a, b]$ if and only if it is the integral of its derivative (a.e. w.r.t. Lebesgue measure).
- Clearly, if f is absolutely continuous on $[a, b]$, then

$$\lim_{h\to 0} f(x+h) - f(x) = \lim_{h\to 0} \int_{[a,x+h]} g(t)dt - \int_{[a,x]} g(t)dt$$

$$= \lim_{h\to 0} \int_{[x,x+h]} g(t)dt = \int_{\{x\}} g(t)dt = 0 \text{ for all } x \in [a, b].$$

- Therefore, absolute continuity implies continuity but the converse is not true (counterexample: the Cantor function).
- *Note:* In $(\star)$ if g is continuous at x then f is differentiable at x and $f'(x) = g(x)$.

13. Product measures

- **Definition:** Consider the measurable spaces $(\Omega_i, \mathcal{F}_i)$, $i = 1, ..., n$, and let
$$\Omega = \Omega_1 \times \Omega_2 \times ... \times \Omega_n = \prod_{i=1}^{n} \Omega_i.$$
A measurable rectangle in the product space Ω is a set
$$A = A_1 \times A_2 \times ... \times A_n = \prod_{i=1}^{n} A_i,$$
where $A_i \in \mathcal{F}_i$, for $i = 1, ..., n$.

- **Definition:** The σ-algebra on Ω generated by the measurable rectangles, $\sigma\left(\left\{\prod_{i=1}^{n} A_i \,\middle|\, A_i \in \mathcal{F}_i\right\}\right)$, is called the product σ-algebra and is denoted by

$$\mathcal{F} = \mathcal{F}_1 \otimes \mathcal{F}_2 \otimes ... \otimes \mathcal{F}_n = \bigotimes_{i=1}^{n} \mathcal{F}_i.$$

- Note that the product σ-algebra is <u>not</u> the Cartesian product of the $\mathcal{F}_i$'s, $\mathcal{F}_1 \times \mathcal{F}_2 \times ... \times \mathcal{F}_n = \prod_{i=1}^{n} \mathcal{F}_i$. The Cartesian product of the $\mathcal{F}_i$'s is the set of measurable rectangles. Therefore, $\bigotimes_{i=1}^{n} \mathcal{F}_i = \sigma\left(\prod_{i=1}^{n} \mathcal{F}_i\right)$.

- The σ-algebra $\mathcal{B}(\mathbb{R}^n)$ of the Borel sets of $\mathbb{R}^n$ is the product σ-algebra generated by the measurable rectangles $A = A_1 \times A_2 \times ... \times A_n$, where A_i is a Borel set of $\mathbb{R}$ for $i = 1, ..., n$, that is, $\mathcal{B}(\mathbb{R}^n) = \bigotimes_{i=1}^{n} \mathcal{B}_i = \sigma\left(\prod_{i=1}^{n} \mathcal{B}_i\right).$

- **Assumption for the next two theorems.** Let $(\Omega_i, \mathcal{F}_i, \mu_i)$ be a measure space where μ_i is σ-finite for $i = 1, 2$ and let $\Omega = \Omega_1 \times \Omega_2$ and $\mathcal{F} = \mathcal{F}_1 \otimes \mathcal{F}_2$.

Product Measure Theorem. There is a unique measure μ on $(\Omega, \mathcal{F})$ such that

$$\mu\left(A_1 \times A_2\right) = \mu_1(A_1) \cdot \mu_2(A_2), \quad \text{for all } A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2.$$

Furthermore, μ is σ-finite and, if μ_1 and μ_2 are probability measures, then μ is also a probability measure. The measure μ is called the product measure of μ_1 and μ_2 $(\mu = \mu_1 \times \mu_2)$ and is given by

$$\mu(B) = \int_{\Omega_1} \mu_2\left(B\left(\omega_1\right)\right) d\mu_1\left(\omega_1\right) = \int_{\Omega_2} \mu_1\left(B\left(\omega_2\right)\right) d\mu_2\left(\omega_2\right),$$

$$\text{for all } B \in \mathcal{F},$$

where

$$B\left(\omega_1\right) = \left\{\omega_2 \in \Omega_2 \,\middle|\, (\omega_1, \omega_2) \in B\right\}, \quad (\text{section of } \omega_1),$$

$$B\left(\omega_2\right) = \left\{\omega_1 \in \Omega_1 \,\middle|\, (\omega_1, \omega_2) \in B\right\}, \quad (\text{section of } \omega_2).$$

14. Fubini's theorem

- Recall that $\Omega = \Omega_1 \times \Omega_2$, $\mathcal{F} = \mathcal{F}_1 \otimes \mathcal{F}_2$, and $\mu = \mu_1 \times \mu_2$.
- **Fubini's Theorem.** If $f : (\Omega, \mathcal{F}) \longrightarrow \left(\overline{\mathbb{R}}, \mathcal{B}\right)$ is a Borel measurable function such that $\int_\Omega f d\mu$ exists, then

$$\int_\Omega f(\omega_1, \omega_2)\, d\mu(\omega_1, \omega_2) = \int_{\Omega_1} \int_{\Omega_2} f(\omega_1, \omega_2)\, d\mu_2(\omega_2)\, d\mu_1(\omega_1)$$

$$= \int_{\Omega_2} \int_{\Omega_1} f(\omega_1, \omega_2)\, d\mu_1(\omega_1)\, d\mu_2(\omega_2).$$

- Observe that $\mu(B) = \int_\Omega \mathbb{I}_B \, d\mu$. Therefore, if we make $f = \mathbb{I}_B$ in Fubini's theorem, we obtain the Product Measure theorem since

$$\mu(B) = \int_\Omega \mathbb{I}_B(\omega_1, \omega_2)\, d\mu(\omega_1, \omega_2)$$

$$= \int_{\Omega_1} \int_{\Omega_2} \mathbb{I}_B(\omega_1, \omega_2)\, d\mu_2(\omega_2)\, d\mu_1(\omega_1)$$

$$= \int_{\Omega_1} \int_{\Omega_2} \mathbb{I}_{B(\omega_1)}(\omega_2)\, d\mu_2(\omega_2)\, d\mu_1(\omega_1) = \int_{\Omega_1} \mu_2(B(\omega_1))\, d\mu_1(\omega_1).$$

- Similarly,

$$\mu(B) = \int_\Omega \mathbb{I}_B(\omega_1, \omega_2)\, d\mu(\omega_1, \omega_2)$$

$$= \int_{\Omega_2} \int_{\Omega_1} \mathbb{I}_B(\omega_1, \omega_2)\, d\mu_1(\omega_1)\, d\mu_2(\omega_2)$$

$$= \int_{\Omega_2} \int_{\Omega_1} \mathbb{I}_{B(\omega_2)}(\omega_1)\, d\mu_1(\omega_1)\, d\mu_2(\omega_2) = \int_{\Omega_2} \mu_1(B(\omega_2))\, d\mu_2(\omega_2).$$

- Note also that

$$\int_B f(\omega_1,\omega_2)\, d\mu(\omega_1,\omega_2) = \int_\Omega f(\omega_1,\omega_2)\, \mathbb{I}_B(\omega_1,\omega_2) d\mu(\omega_1,\omega_2)$$

$$= \int_{\Omega_1}\int_{\Omega_2} f(\omega_1,\omega_2)\, \mathbb{I}_B(\omega_1,\omega_2) d\mu_2(\omega_2)\, d\mu_1(\omega_1)$$

$$= \int_{\Omega_1}\int_{\Omega_2} f(\omega_1,\omega_2)\, \mathbb{I}_{B(\omega_1)}(\omega_2) d\mu_2(\omega_2)\, d\mu_1(\omega_1)$$

or

$$\int_B f(\omega_1,\omega_2)\, d\mu(\omega_1,\omega_2) = \int_\Omega f(\omega_1,\omega_2)\, \mathbb{I}_B(\omega_1,\omega_2) d\mu(\omega_1,\omega_2)$$

$$= \int_{\Omega_2}\int_{\Omega_1} f(\omega_1,\omega_2)\, \mathbb{I}_B(\omega_1,\omega_2) d\mu_1(\omega_1)\, d\mu_2(\omega_2)$$

$$= \int_{\Omega_2}\int_{\Omega_1} f(\omega_1,\omega_2)\, \mathbb{I}_{B(\omega_2)}(\omega_1) d\mu_1(\omega_1)\, d\mu_2(\omega_2).$$

- Both the Product Measure theorem and Fubini's theorem can be extended to n factors with $\Omega = \Omega_1 \times \Omega_2 \times ... \times \Omega_n \equiv \prod_{i=1}^{n} \Omega_i$,

  $\mathcal{F} = \mathcal{F}_1 \otimes \mathcal{F}_2 \otimes ... \otimes \mathcal{F}_n \equiv \bigotimes_{i=1}^{n} \mathcal{F} i$, and

  $\mu = \mu_1 \times \mu_2 \times ... \times \mu_n \equiv \prod_{i=1}^{n} \mu_i$.

- The Lebesgue measure on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ is the product measure of n Lebesgue measures on $(\mathbb{R}, \mathcal{B})$. Thus,

$$\int_{\mathbb{R}^2} f(x_1, x_2)\, d(x_1, x_2) = \int_{\mathbb{R}} \int_{\mathbb{R}} f(x_1, x_2)\, dx_2 dx_1$$

$$= \int_{\mathbb{R}} \int_{\mathbb{R}} f(x_1, x_2)\, dx_1 dx_2.$$

- See the handout for an application of Fubini's theorem.

Riemann and Lebesgue Integrals

Let $[a, b]$ be a bounded closed interval of real numbers and let f be a bounded real-valued function defined on $[a, b]$, fixed throughout the discussion. If $P = \{(x_{i-1}, x_i]\}_{i=1}^{n}$ is a subinterval partition of $(a, b]$ with $x_{i-1} < x_i$ for $i = 1, ..., n$, $x_0 = a$, and $x_n = b$, we may construct the upper and lower sums of f relative to P (also called upper and lower Darboux sums) as follows.

Let

$$M_i = \sup \{f(y) \,|y \in (x_{i-1}, x_i]\}, \quad i = 1, ..., n,$$

$$m_i = \inf \{f(y) \,|y \in (x_{i-1}, x_i]\}, \quad i = 1, ..., n,$$

and define step functions α and β on $[a, b]$, called the upper and lower functions corresponding to P, by

$$\alpha(x) = M_i \quad \text{if} \quad x \in (x_{i-1}, x_i], \quad i = 1, ..., n,$$

$$\beta(x) = m_i \quad \text{if} \quad x \in (x_{i-1}, x_i], \quad i = 1, ..., n,$$

and both $\alpha(a)$ and $\beta(a)$ may be chosen arbitrarily. The upper and lower Darboux sums of f relative to P are given by

$$U(P) = \sum_{i=1}^{n} M_i (x_i - x_{i-1}),$$

$$L(P) = \sum_{i=1}^{n} m_i (x_i - x_{i-1}).$$

Consider the measure space $(\Omega, \mathcal{F}, \mu)$, where $\Omega = [a, b]$, $\mathcal{F} = \mathcal{B}([a, b])$, the Borel sets of $[a, b]$, and $\mu =$ Lebesgue measure. Since α and β are simple functions, we have that the Darboux sums are equal to the following Lebesgue integrals:

$$U(P) = \int_{[a,b]} \alpha dx,$$

$$L(P) = \int_{[a,b]} \beta dx.$$

Observe that the arbitrary values of $\alpha(a)$ and $\beta(a)$ do not affect the values of the Darboux sums.

Now let $\{P_k\}_{k=1}^{\infty}$ be a sequence of subinterval partitions of $(a, b]$ such that P_{k+1} is a refinement of P_k for each k (i.e., the intervals in P_{k+1} are subintervals of the intervals in P_k) and the length of the largest subinterval of P_k approaches 0 as $k \to \infty$. If α_k and β_k are the upper and lower functions corresponding to P_k, then

$$\alpha_1 \geq \alpha_2 \geq ... \geq f \geq ... \geq \beta_2 \geq \beta_1.$$

Thus, α_k and β_k tend to limit functions α and β. If $|f|$ is bounded by B, then all $|\alpha_k|$ and $|\beta_k|$ are bounded by B as well, and the function that is constant at B is Lebesgue integrable on $[a, b]$ since

$$\mu\,[a, b] = b - a < \infty$$

and, hence,

$$\int_{[a,b]} B dx = B\,(b - a) < \infty.$$

By the dominated convergence theorem,

$$\lim_{k\to\infty} U(P_k) = \lim_{k\to\infty} \int_{[a,b]} \alpha_k dx = \int_{[a,b]} \alpha dx,$$

$$\lim_{k\to\infty} L(P_k) = \lim_{k\to\infty} \int_{[a,b]} \beta_k dx = \int_{[a,b]} \beta dx.$$

The integrals $\int_{[a,b]} \alpha dx$ and $\int_{[a,b]} \beta dx$ are called, respectively, the upper and lower Darboux integrals of f for the sequence $\{P_k\}_{k=1}^{\infty}$ of subinterval partitions.

Note that

$$f \text{ is continuous at } x \quad \text{if and only if} \quad \alpha(x) = f(x) = \beta(x),$$

as follows from a standard $\varepsilon - \delta$ argument.

If $\lim_{k\to\infty} U(P_k) = \lim_{k\to\infty} L(P_k) =$ a finite number r, independent of the particular sequence of subinterval partitions, f is said to be Riemann integrable on $[a, b]$, and the number $r \equiv \int_a^b f dx$ is said to be (the value of) the Riemann integral of f on $[a, b]$. The above argument shows that f is Riemann integrable iff

$$\int_{[a,b]} \alpha dx = \int_{[a,b]} \beta dx,$$

independent of the particular sequence of subinterval partitions. Moreover, if f is Riemann integrable, then

$$\int_a^b f dx = \int_{[a,b]} \alpha dx = \int_{[a,b]} \beta dx = \int_{[a,b]} f dx.$$

From the previous discussion we have the following:

Proposition 1. Let f be a bounded real-valued function on $[a, b]$.

(a) The function f is Riemann integrable on $[a, b]$ if and only if f is continuous almost everywhere on $[a, b]$ with respect to Lebesgue measure.

(b) If f is Riemann integrable on $[a, b]$, then f is integrable with respect to Lebesgue measure on $[a, b]$, and the two integrals are equal, $\int_{[a,b]} f dx = \int_a^b f dx$.

The previous proposition gives us in part (a) the precise criterion for Riemann integrability. From its part (b) we see that Lebesgue integration is more general than Riemann integration as Riemann integrability implies Lebesgue integrability (but the converse is not necessarily true).

Riemann-Stieltjes (or Stieltjes) and Lebesgue-Stieltjes Integrals

Let f be a bounded real-valued function defined on the bounded interval $[a, b]$ and F is a distribution function (i.e., an increasing and right-continuous function) on $\mathbb{R}$ that induces the corresponding Lebesgue-Stieltjes measure μ on the Borel sets of $(a, b]$. Define M_i, m_i, α, and β as before and take

$$U(P) = \sum_{i=1}^{n} M_i \left(F\left(x_i\right) - F\left(x_{i-1}\right)\right) = \int_{(a,b]} \alpha d\mu \equiv \int_{(a,b]} \alpha dF,$$

$$L(P) = \sum_{i=1}^{n} m_i \left(F\left(x_i\right) - F\left(x_{i-1}\right)\right) = \int_{(a,b]} \beta d\mu \equiv \int_{(a,b]} \beta dF.$$

Let $\{P_k\}_{k=1}^{\infty}$ be a sequence of subinterval partitions of $(a, b]$ such that the length of the largest subinterval of P_k approaches 0 as $k \to \infty$ and P_{k+1} is a refinement of P_k for each k, and α_k and β_k be the upper and lower functions corresponding to P_k. Then,

$$\lim_{k\to\infty} U(P_k) = \int_{(a,b]} \alpha dF,$$

$$\lim_{k\to\infty} L(P_k) = \int_{(a,b]} \beta dF,$$

where $\alpha = \lim_{k\to\infty} \alpha_k$ and $\beta = \lim_{k\to\infty} \beta_k$. If $U(P_k)$ and $L(P_k)$ approach the same limit s independent of the particular sequence of subinterval partitions, this number $s \equiv \int_a^b f dF$ is called the Riemann-Stieltjes integral of f with respect to F on $(a, b]$, and f is said to be Riemann-Stieltjes integrable with respect to F on $(a, b]$.

Proposition 2. Let f be a bounded real-valued function on $[a, b]$ and F be a distribution function on $\mathbb{R}$.

(a) The function f is Riemann-Stieltjes integrable with respect to F on $(a, b]$ if and only if f is continuous almost everywhere on $(a, b]$ with respect to the Lebesgue-Stieltjes measure μ on $(a, b]$ induced by F.

(b) If f is Riemann-Stieltjes integrable with respect to F on $(a, b]$, then f is integrable with respect to the Lebesgue-Stieltjes measure μ on $(a, b]$ induced by F, and the two integrals are equal, $\int_{(a,b]} f dF = \int_a^b f dF$.

An application of Fubini's theorem

Let $\Omega_i = \mathbb{R}$, $i = 1, 2$, and $\Omega = \Omega_1 \times \Omega_2$. Therefore, $\Omega = \mathbb{R}^2$.
Let $\mu_i =$ Lebesgue measure on $(\mathbb{R}, \mathcal{B})$, $i = 1, 2$, where $\mathcal{B}$ is the $\sigma-$algebra of the Borel sets in $\mathbb{R}$.
Let $\mathcal{B} \otimes \mathcal{B} = \mathcal{B}(\mathbb{R}^2)$.
Let $\mu = \mu_1 \times \mu_2$. Therefore, μ is the Lebesgue measure on $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$.
Define the Borel measurable function $f : (\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2)) \longrightarrow (\mathbb{R}, \mathcal{B})$,

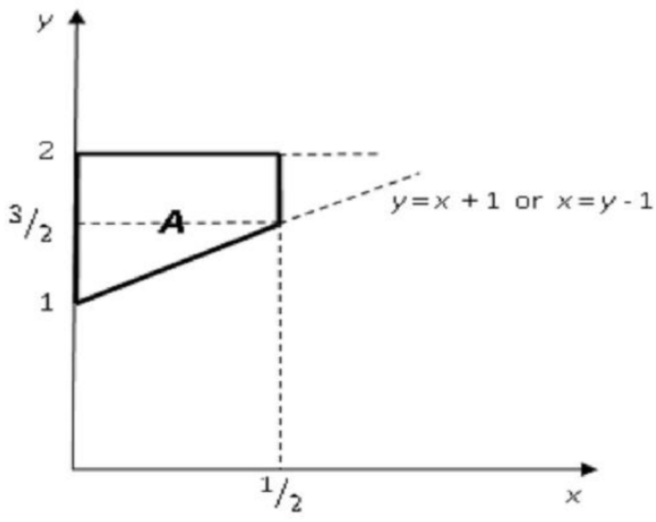
$$f(x, y) = \frac{3}{5}x(y + x).$$

We want to compute the following Lebesgue integral:

$$\int_A f d\mu = \int_A f(x, y) d(x, y),$$

where the Borel set A is given by

$$A = \left\{(x, y) \in \mathbb{R}^2 \mid 0 \leq x < 1/2\,, x + 1 < y \leq 2\right\}.$$



Then,

$$\int_A f(x, y) d(x, y) = \int_{\mathbb{R}^2} f(x, y) \mathbb{I}_A(x, y) d(x, y)$$

$$= \int_{\mathbb{R}} \int_{\mathbb{R}} f(x, y) \mathbb{I}_A(x, y) dy dx = \int_{\mathbb{R}} \int_{\mathbb{R}} f(x, y) \mathbb{I}_{A(x)}(y) dy dx,$$

where the second equality follows from Fubini's theorem and

$$A(x) = \begin{cases} \{y \in \mathbb{R} \mid y \in (x + 1, 2]\} & \text{for } x \in [0, 1/2) \\ \emptyset & \text{for } x \notin [0, 1/2)\,. \end{cases}$$

Hence,

$$\int_{\mathbb{R}}\int_{\mathbb{R}} f(x,y)\mathbb{I}_{A(x)}(y)dydx = \int_{[0,1/2)}\int_{(x+1,2]} f(x,y)dydx + \underbrace{\int_{[0,1/2)^c}\int_{\{\emptyset\}} f(x,y)dydx}_{=0}$$

$$= \int_0^{1/2}\int_{x+1}^{2} \frac{3}{5}x(y+x)dydx = \frac{63}{640},$$

where the second equality follows from the continuity of the function f, which allows us to use Riemann integration to compute the Lebesgue integral.

Alternatively, we could integrate first with respect to the variable x and then with respect to the variable y according to Fubini's theorem. Thus,

$$\int_A f(x,y)d(x,y) = \int_{\mathbb{R}^2} f(x,y)\mathbb{I}_A(x,y)d(x,y)$$

$$= \int_{\mathbb{R}}\int_{\mathbb{R}} f(x,y)\mathbb{I}_A(x,y)dxdy = \int_{\mathbb{R}}\int_{\mathbb{R}} f(x,y)\mathbb{I}_{A(y)}(x)dxdy,$$

where

$$A(y) = \begin{cases} \{x \in \mathbb{R} \mid x \in [0, y-1)\} & \text{for } y \in (1, 3/2] \\ \{x \in \mathbb{R} \mid x \in [0, 1/2)\} & \text{for } y \in (3/2, 2] \\ \emptyset & \text{for } y \notin (1, 2]\,. \end{cases}$$

Hence,

$$\int_{\mathbb{R}}\int_{\mathbb{R}} f(x,y)\mathbb{I}_{A(y)}(x)dxdy =$$

$$\int_{(1,3/2]}\int_{[0,y-1)} f(x,y)dxdy + \int_{(3/2,2]}\int_{[0,1/2)} f(x,y)dxdy + \underbrace{\int_{(1,2]^c}\int_{\{\emptyset\}} f(x,y)dxdy}_{=0}$$

$$= \int_1^{3/2}\int_0^{y-1} \frac{3}{5}x(y+x)dxdy + \int_{3/2}^{2}\int_0^{1/2} \frac{3}{5}x(y+x)dxdy = \frac{63}{640}.$$

Exercises. Probability and Statistics. IDEA.
Measure Theory

1. Let Ω be a countably infinite set, and let $\mathcal{F}$ consist of all subsets of Ω $\left(\mathcal{F} = 2^{\Omega}\right)$. Define $\mu(A) = 0$ if A is finite, $\mu(A) = \infty$ if A is infinite.

   (a) Show that μ is finitely additive but not countably additive.

   (b) Show that Ω is the limit of an increasing sequence of sets A_n with $\mu(A_n) = 0$ for all n, but $\mu(\Omega) = \infty$.

2. Let μ be a Lebesgue-Stieltjes measure on $\mathbb{R}$ corresponding to a continuous distribution function.

   (a) If A is a countable subset of $\mathbb{R}$, show that $\mu(A) = 0$.

   (b) If $\mu(A) > 0$, must A include an open interval?

   (c) If $\mu(A) > 0$ and $\mu(A^c) = 0$, must A be dense in $\mathbb{R}$? *Note*: A set is dense in $\mathbb{R}$ if its closure is $\mathbb{R}$.

   (d) Do the answers to (b) and (c) change if μ is restricted to be Lebesgue measure?

3. The following statement is true:

   "A distribution function on $\mathbb{R}$ is monotone and thus has only countably many points of discontinuity."

   Is this also true for a distribution function on $\mathbb{R}^n$, $n > 1$?

4. If $f_1, f_2, \ldots$ are extended real-valued Borel measurable functions on $(\Omega, \mathcal{F})$, $n = 1, 2, \ldots$, show that $\sup_n f_n$ and $\inf_n f_n$ are Borel measurable.

5. Let $f = f(x, y)$ be a real-valued function of two real variables, defined for $a < y < b$, $c < x < d$. Assume that for each x, $f(x, \cdot)$ is a Borel measurable function of y, and that there is a Borel measurable function $g : (a, b) \to \mathbb{R}$ such that $|f(x, y)| \leq g(y)$ for all x, y, and $\int_{(a,b)} g(y)dy < \infty$. If $x_0 \in (c, d)$ and $\lim_{x \to x_0} f(x, y)$ exists for all $y \in (a, b)$, show that

$$\lim_{x \to x_0} \int_{(a,b)} f(x, y)dy = \int_{(a,b)} \left[\lim_{x \to x_0} f(x, y)\right] dy.$$

6. The function defined on $[0, 1]$ by $f(x) = 1$ if x is irrational, and $f(x) = 0$ if x is rational, is the standard example of a function that is Lebesgue integrable (it is 1 a.e.) but not Riemann integrable. But what is wrong with the following reasoning:

   "If we consider the behavior of f on the irrationals, f assumes the constant value 1 and is therefore continuous. Since the rationals have Lebesgue measure 0, f is therefore continuous almost everywhere and hence is Riemann integrable."

7. Give an example of a sequence of functions f_n on $[a, b]$ such that each f_n is Riemann integrable, $|f_n| \leq 1$ for all n, $f_n \to f$ everywhere, but f is not Riemann integrable.

8. Prove the Borel-Cantinelli Lemma:

   If $A_1, A_2, ... \in \mathcal{F}$ and $\sum_{n=1}^{\infty} \mu(A_n) < \infty$, then $\mu\left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k\right) = 0$.

   Note that $\omega \in \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$ iff $\omega \in A_n$ for infinitely many n.

9. Give an example of a measure μ and a nonnegative finite-valued Borel measurable function g such that the measure λ defined by $\lambda(A) = \int_A g \, d\mu$ is not σ-finite.

10. Let $f, f_1, f_2, ...$ be Borel measurable functions on $(\Omega, \mathcal{F}, \mu)$. We say that f_n converges to f in measure (or in μ-measure) if for every $\varepsilon > 0$,

$$\lim_{n\to\infty} \mu \{\omega \mid |f_n(\omega) - f(\omega)| \geq \varepsilon\} = 0.$$

    If f_n converges both to f and g in measure, show that $f = g$ a.e. with respect to μ.

11. Let $f_n : (\mathbb{R}, \mathcal{B}) \longrightarrow (\mathbb{R}, \mathcal{B})$, $n = 1, 2, ...,$ be

$$f_n(x) = \begin{cases} 1 & \text{for } x \in [0,1] \\ 0 & \text{otherwise} \end{cases}$$

    if n is odd and

$$f_n(x) = \begin{cases} 1 & \text{for } x \in (1,2] \\ 0 & \text{otherwise} \end{cases}$$

    if n is even.

    Does the following equality hold:

$$\liminf_{n\to\infty} \int_{\mathbb{R}} f_n dx = \int_{\mathbb{R}} \liminf_{n\to\infty} f_n dx \ ?$$

12. Let $f_n : (\mathbb{R}, \mathcal{B}) \longrightarrow (\mathbb{R}, \mathcal{B})$, $n = 1, 2, ...,$ be

$$f_n(x) = \begin{cases} n & \text{for } x \in (0, 1/n] \\ 0 & \text{otherwise} \end{cases}$$

    Does the following equality hold:

$$\lim_{n\to\infty} \int_{\mathbb{R}} f_n dx = \int_{\mathbb{R}} \lim_{n\to\infty} f_n dx \ ?$$

13. Let $\Omega_1 = \Omega_2 =$ the set of positive integers, $\mathcal{F}_1 = \mathcal{F}_2 =$ all subsets, $\mu_1 = \mu_2 =$ counting measure, $f(n,n) = n$, $f(n,n+1) = -n$, $n = 1,2,...,$ $f(i,j) = 0$ if $j \neq i$ or $j \neq i+1$. Show that $\int_{\Omega_1}\int_{\Omega_2} f \; d\mu_2 \; d\mu_1 = 0$ and $\int_{\Omega_2}\int_{\Omega_1} f \; d\mu_1 \; d\mu_2 = \infty$. *Note*: Fubini's theorem fails since the integral of f with respect to $\mu_1 \times \mu_2$ does not exist.

14. Let $\Omega_1 = \Omega_2 = \mathbb{R}$, $\mathcal{F}_1 = \mathcal{F}_2 = \mathcal{B}$, $\mu_1 =$ Lebesgue measure, $\mu_2 =$ counting measure. Let $A = \{(\omega_1, \omega_2) \,|\omega_1 = \omega_2\} \in \mathcal{F}_1 \otimes \mathcal{F}_2$. Show that

$$\int_{\Omega_1}\int_{\Omega_2} \mathbb{I}_A \; d\mu_2 \; d\mu_1 = \int_{\Omega_1} \mu_2\left(A(\omega_1)\right) d\mu_1(\omega_1) = \infty,$$

but

$$\int_{\Omega_2}\int_{\Omega_1} \mathbb{I}_A \; d\mu_1 \; d\mu_2 = \int_{\Omega_2} \mu_1\left(A(\omega_2)\right) d\mu_2(\omega_2) = 0.$$

Recall that $A(\omega_1) = \{\omega_2 \in \Omega_2 \,|\, (\omega_1, \omega_2) \in A\}$ and, symmetrically, $A(\omega_2) = \{\omega_1 \in \Omega_1 \,|\, (\omega_1, \omega_2) \in A\}$. *Note*: Fubini's theorem fails since μ_2 is not $\sigma-$finite.

Probability and Statistics. IDEA. Answers.

Measure Thery.

1. (a) If $A_1, ..., A_n$ are disjoint subsets of Ω, then

$$\mu\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mu(A_i) = \begin{cases} 0 & \text{if all } A_i \text{ are finite,} \\ \infty & \text{if some } A_i \text{ is infinite.} \end{cases}$$

Thus μ is finitely additive. If $\Omega = \{x_1, x_2, ...\}$ and $A_i = \{x_i\}$, then $\infty = \mu(\bigcup_{n=1}^{\infty} A_n) \neq \sum_{n=1}^{\infty} \mu(A_n) = 0$, hence μ is not countably additive.

(b) Take $A_n = \{x_1, ..., x_n\}$.

— — — — — — — — — — — —

2. (a) If $A = \{x_1, x_2, ...\}$, then $\mu(A) = \sum_n \mu\{x_n\} = 0$.

(b) No. Let A be the set of irrationals.

(c) No. For example, if $F(x) = c_1$ for $x \leq a$ and $F(x) = c_2$ for $x \geq b$, we may take $A = [a, b]$.

(d) The answer to (b) is unchanged, but in (c), A must be dense in $\mathbb{R}$. Consider any point $x \in \mathbb{R}$ and an arbitrary open interval $I_x = (a, b)$ centered at x. We have $\mu(I_x) = b - a > 0$ and, hence, I_x cannot be a subset of A^c since $\mu(A^c) = 0$. Therefore, $A \cap I_x$ is not empty. This means that x belongs to the closure of A. Since this is true for all $x \in \mathbb{R}$, A is dense in $\mathbb{R}$.

— — — — — — — — — — — —

3. No. For example, let $F(x, y) = F_1(x) \cdot F_2(y)$, where F_1 and F_2 are distribution functions on $\mathbb{R}$. If F_1 is discontinuous at x_0, then F is discontinuous at (x_0, y) whenever $F_2(y) > 0$. For an explicit example, let

$F_1(x) = F_2(x) = 1$ for $x \geq 0$, and 0 for $x < 0$. Then F is the indicator of the first quadrant.

− − − − − − − − − − − −

4. Note that a function $f : (\Omega, \mathcal{F}) \to \left(\overline{\mathbb{R}}, \mathcal{B}\right)$ is Borel measurable if and only if $\{\omega \,|\, f(\omega) > c\} \in \mathcal{F}$ for all $c \in \overline{\mathbb{R}}$ (or iff $\{\omega \,|\, f(\omega) < c\} \in \mathcal{F}$ for all $c \in \overline{\mathbb{R}}$). Then, if $c \in \overline{\mathbb{R}}$,

$$\left\{\omega \,\middle|\, \sup_n f_n(\omega) > c\right\} = \bigcup_{n=1}^{\infty} \{\omega \,|\, f_n(\omega) > c\} \in \mathcal{F},$$

and

$$\left\{\omega \,\middle|\, \inf_n f_n(\omega) < c\right\} = \bigcup_{n=1}^{\infty} \{\omega \,|\, f_n(\omega) < c\} \in \mathcal{F}.$$

− − − − − − − − − − − −

5. Let $\{x_n\}$ be any sequence of points converging to x_0. For each n, $|f(x_n, y)| \leq g(y)$, and g is integrable by hypothesis. By the dominated convergence theorem,

$$\lim_{n\to\infty} \int_a^b f(x_n, y)dy = \int_a^b \left[\lim_{n\to\infty} f(x_n, y)\right] dy,$$

as desired. (Note that $\lim\limits_{n\to\infty} f(x_n, \cdot)$ is a limit of Borel measurable functions, hence is Borel measurable).

− − − − − − − − − − − −

6. The restriction of f to the irrationals is continuous; in other words, if we consider a sequence of irrationals $x_n \to x$, x irrational, then $f(x_n) \to f(x)$. This is not the same thing as saying that f is continuous at each irrational point, since if $\{x_n\}$ is an arbitrary sequence (which can be of rationals) approaching x, $f(x_n)$ need not approach $f(x)$.

− − − − − − − − − − − −

7. Let $r_1, r_2, \ldots$ be an ordering of the rationals in $[a, b]$, and take $f_n = \mathbb{I}_{A_n}$, where $A_n = \{r_1, \ldots, r_n\}$. Since f_n is continuous except at $r_1, \ldots, r_n$, it is Riemann integrable. But $f = \lim_n f_n$ is the indicator of the set of all rationals in $[a, b]$; hence f is discontinuous everywhere and therefore not Riemann integrable.

— — — — — — — — — — — —

8. $\mu(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k) \leq \mu(\bigcup_{k=n}^{\infty} A_k) \leq \sum_{k=n}^{\infty} \mu(A_k) \to 0$, as $n \to \infty$.

— — — — — — — — — — — —

9. Let μ be counting measure on all subsets of $\mathbb{R}$. Take $g \equiv 1$, so that $\lambda = \mu$. Then λ is not σ-finite since $\mathbb{R}$ is not a countable union of finite sets.

— — — — — — — — — — — —

10. We have $\mu\{|f - g| \geq \varepsilon\} \leq \mu\{|f - f_n| \geq \frac{\varepsilon}{2}\} + \mu\{|f_n - g| \geq \frac{\varepsilon}{2}\} \to 0$. Therefore $\mu\{|f - g| \geq \varepsilon\} = 0$ for each $\varepsilon > 0$; hence $f = g$ a.e. with respect to μ.

— — — — — — — — — — — —

11. Note that $\int_{\mathbb{R}} f_n dx = 1$ for all n. Therefore, $\lim_{n\to\infty} \int_{\mathbb{R}} f_n dx = \liminf_{n\to\infty} \int_{\mathbb{R}} f_n dx = 1$. However, $\liminf_{n\to\infty} f_n(x) = 0$ for all $x \in \mathbb{R}$ and, hence, $\int_{\mathbb{R}} \liminf_{n\to\infty} f_n(x) dx = 0$. Thus

$$0 = \int_{\mathbb{R}} \liminf_{n\to\infty} f_n(x) dx < \liminf \int_{\mathbb{R}} f_n dx = 1,$$

which proves that a strict inequality in Fatou's lemma is possible

— — — — — — — — — — — —

12. It is obvious that $f_n \longrightarrow 0$ pointwise so that $\int_{\mathbb{R}} \lim_{n\to\infty} f_n dx = 0$. Note that $\int_{\mathbb{R}} f_n dx = n \cdot \dfrac{1}{n} = 1$ so that $\lim_{n\to\infty} \int_{\mathbb{R}} f_n dx = 1$. Thus, the integral and the limit do not commute in this example.

Note that the sequence of functions $\{f_n\}_{n=1}^{\infty}$ of this exercise is NOT non-decreasing so that we cannot apply the monotone convergence theorem. Moreover, there is no Lebesgue integrable function g such that $|f_n(x)| \leq g(x)$ for all $x \in \mathbb{R}$ and for all n. Therefore, the dominated convergence theorem does not apply either.

— — — — — — — — — — — —

13. Note that $\mu = \mu_1 \times \mu_2$ is the counting measure on $\Omega = \Omega_1 \times \Omega_2$.

$\int\limits_{\Omega} f d\mu = \int\limits_{\Omega} f^+ d\mu - \int\limits_{\Omega} f^- d\mu = (1+2+3+....) - (1+2+3+....) = \infty - \infty,$

so that $\int\limits_{\Omega} f d\mu$ does not exist.

$$\int\limits_{\Omega_1}\int\limits_{\Omega_2} f d\mu_2 d\mu_1 = \sum_i \left(\sum_j f(i,j) \right) = \sum_i (i - i) = \sum_i 0 = 0,$$

$$\int\limits_{\Omega_2}\int\limits_{\Omega_1} f d\mu_1 d\mu_2 = \sum_j \left(\sum_i f(i,j) \right) = \sum_j (-(j-1) + j) = \sum_j 1 = \infty.$$

— — — — — — — — — — — —

14. Note that

$$\int\limits_{\Omega_1} \mu_2(A(\omega_1)) d\mu_1(\omega_1) = \int\limits_{\Omega_1} \mu_2\{\omega_1\} d\mu_1(\omega_1) = \int\limits_{\Omega_1} 1 d\mu_1 = \mu_1(\Omega_1) = \mu_1(\mathbb{R}) = \infty,$$

$$\int\limits_{\Omega_2} \mu_1(A(\omega_2)) d\mu_2(\omega_2) = \int\limits_{\Omega_2} \mu_1\{\omega_2\} d\mu_2(\omega_2) = \int\limits_{\Omega_2} 0 d\mu_2 = 0.$$

— — — — — — — — — — — —