Example 2.2 (Cantor-distributed random variable)
Let $C = \left\{x \in [0,1] \,\middle|\, x = \sum_{n=1}^{\infty} \frac{d_n}{3^n},\ d_n = 0 \text{ or } 2,\ n = 1, 2, \dots\right\}$ **(Cantor set)**

- C is a continuum (C has the same cardinal # as $[0,1]$)
- $\lambda(C) = 0$

Let $Y_1, Y_2, \dots$ be i.i.d. random variables s.t. $\mathsf{P}(Y_n = 0) = \mathsf{P}(Y_n = 2) = \frac{1}{2}$ and let $Y = \sum_{n=1}^{\infty} Y_n/3^n$
$\Rightarrow Y$ takes values in C (Y is continuous), but has no density function since $\lambda(C) = 0$ **(singular continuous)**.

2.1 Definition of Lebesgue Integrals

- Expectation = Lebesgue integral w.r.t. probability measure

$$\mathsf{E}(X) = \int_{\Omega} X(\omega)\, \mathsf{P}(\mathrm{d}\omega)$$

- Define the Lebesgue integral in three steps.

$(\Omega, \mathcal{F}, \mu)$**:** Measure space

h**:** $\Omega \to \overline{\mathbb{R}} = [-\infty, +\infty]$: $\mathcal{F}/\mathcal{B}(\overline{\mathbb{R}})$-measurable function

Def. 2.2 (Integral of simple functions)

h is **simple** $\Leftrightarrow$ h takes only a finite # of distinct values $x_1, x_2, \dots, x_n$ s.t.

$$h(\omega) = \sum_{i=1}^{n} x_i \mathbf{1}_{A_i}(\omega), \quad \omega \in \Omega,$$

where $A_1, A_2, \dots, A_n \in \mathcal{F}$ s.t. $\bigcup_{i=1}^{n} A_i = \Omega$ & $A_i \cap A_j = \emptyset$ $(i \neq j)$

h is simple $\Rightarrow$ $\displaystyle\int_{\Omega} h(\omega)\, \mu(\mathrm{d}\omega) = \sum_{i=1}^{n} x_i\, \mu(A_i)$

Example 2.3 $A \in \mathcal{F}$, $\mathsf{E}(\mathbf{1}_A) = 1 \times \mathsf{P}(A) + 0 \times \mathsf{P}(A^c) = \mathsf{P}(A)$

Def. 2.3 (Integral of nonnegative functions)

$$h \geq 0 \text{ a.e.-}\mu \Rightarrow \quad \int_\Omega h(\omega)\,\mu(\mathrm{d}\omega) = \sup_{g \in \mathcal{S}_h} \int_\Omega g(\omega)\,\mu(\mathrm{d}\omega)$$

where $\mathcal{S}_h = \{\text{measurable and simple } g \text{ s.t. } g \leq h \text{ a.e.-}\mu\}$

- For $A \in \mathcal{F}$, "A a.e.-μ" $\Leftrightarrow$ "$\mu(B^c) = 0$" (almost everywhere w.r.t. measure μ)
- a.e.-P = a.s. (almost surely)

Remark 2.1 It is possible that $\int h\,\mathrm{d}\mu = +\infty$

Def. 2.4 (Integral of general functions)

$h^+ = h\,\mathbf{1}_{\{h \geq 0\}} = \max(h, 0)$ & $h^- = -h\,\mathbf{1}_{\{h<0\}} = -\min(h, 0)$
($h^+ \geq 0$ & $h^- \geq 0$ a.e.-μ) $\Rightarrow$

$$\int_\Omega h(\omega)\,\mu(\mathrm{d}\omega) = \int_\Omega h^+(\omega)\,\mu(\mathrm{d}\omega) - \int_\Omega h^-(\omega)\,\mu(\mathrm{d}\omega)$$

- $\int h^+\,\mathrm{d}\mu = \int h^-\,\mathrm{d}\mu = +\infty \Leftrightarrow \int h\,\mathrm{d}\mu$ does not exist
- h is **μ-integrable** $\Leftrightarrow -\infty < \int h\,\mathrm{d}\mu < \infty$ ($\int h^+\,\mathrm{d}\mu < \infty$ & $\int h^-\,\mathrm{d}\mu < \infty$)

Properties of Integrals

i) $\int h\,\mathrm{d}\mu$ exists $\Rightarrow {}^\forall c \in \mathbb{R},\ \int c\,h\,\mathrm{d}\mu$ exists and $= c\int h\,\mathrm{d}\mu$

ii) $h \leq g$ a.e.-$\mu \Rightarrow \int h\,\mathrm{d}\mu \leq \int g\,\mathrm{d}\mu$

iii) $\int h\,\mathrm{d}\mu$ exists $\Rightarrow |\int h\,\mathrm{d}\mu| \leq \int |h|\,\mathrm{d}\mu$

Remark 2.2 X is a r.v. either discrete or absolutely continuous with probability density function $f \Rightarrow$

$$\mathsf{E}(X) = \int_\Omega X(\omega)\,\mathsf{P}(\mathrm{d}\omega) = \begin{cases} \displaystyle\sum_{i=1}^{\infty} x_i\,\mathsf{P}(X = x_i) & X \text{ is discrete,} \\ \displaystyle\int x\,f(x)\,\mathrm{d}x & X \text{ is continuous with density } f \end{cases}$$

Example 2.4 (X has density function f and $0 \le X \le 1$)

$$\begin{aligned}
\int_0^1 x\, f(x)\, \mathrm{d}x &= \lim_{n\to\infty} \sum_{i=0}^{n-1} \int_{i/n}^{(i+1)/n} x\, f(x)\, \mathrm{d}x \\
&\le \lim_{n\to\infty} \sum_{i=0}^{n-1} \frac{i+1}{n}\, \mathsf{P}\Big(\frac{i}{n} \le X < \frac{i+1}{n}\Big) \\
&= \lim_{n\to\infty} \Big(\sum_{i=0}^{n-1} \frac{i}{n}\, \mathsf{P}\Big(\frac{i}{n} \le X < \frac{i+1}{n}\Big) + \frac{1}{n}\Big) \\
&= \lim_{n\to\infty} \mathsf{E}\Big(\frac{\lfloor n\, X \rfloor}{n}\Big) \le \mathsf{E}(X)
\end{aligned}$$

$\int x\, f(x)\, \mathrm{d}x \ge \mathsf{E}(X)$ is similar.

To show the general case, we use the monotone convergence theorem.