ECE 534 — Random Processes

Lecture Notes: The Cantor Function, Absolute Continuity, Radon–Nikodym Derivatives, Densities, and the Meaning of Integration

Abstract

These notes develop a single, carefully chosen counterexample—the *Cantor distribution*—and use it to motivate the measure-theoretic machinery underlying continuous random variables. We construct the Cantor set, prove it is uncountable yet Lebesgue-null, build the Cantor function, and prove that it is continuous and is a bona fide cumulative distribution function (CDF) which nonetheless admits *no* probability density function (pdf). This forces us to define *absolute continuity* and the *Radon–Nikodym derivative*, which is exactly the object we call a density. We then record the defining properties of pdfs, list the standard continuous laws (Gaussian, uniform, exponential), and close by asking what an integral really *is*, giving the rigorous definition of the Riemann integral and indicating why probability theory ultimately needs the Lebesgue integral.

Contents

1 The Cantor Set

Construction 1.1 (Middle-thirds Cantor set)**.** Begin with $\mathcal{C}_0 = [0,1]$ and build the set by *sequentially removing open middle-third intervals*, always retaining the two endpoints that each removal creates.

Stage 1*.* Remove the open middle third $\left(\frac{1}{3}, \frac{2}{3}\right)$ from $[0,1]$. Its endpoints $\frac{1}{3}$ and $\frac{2}{3}$ are *not* removed, leaving the two closed intervals

$$\mathcal{C}_1 = \left[0, \tfrac{1}{3}\right] \cup \left[\tfrac{2}{3}, 1\right].$$

Stage 2*.* Remove the open middle third of each remaining interval, namely $\left(\frac{1}{9}, \frac{2}{9}\right)$ and $\left(\frac{7}{9}, \frac{8}{9}\right)$; once again their endpoints survive, leaving

$$\mathcal{C}_2 = \left[0, \tfrac{1}{9}\right] \cup \left[\tfrac{2}{9}, \tfrac{1}{3}\right] \cup \left[\tfrac{2}{3}, \tfrac{7}{9}\right] \cup \left[\tfrac{8}{9}, 1\right].$$

Stage $n \to n+1$*.* In general, if $\mathcal{C}_n$ is a disjoint union of 2^n closed intervals each of length 3^{-n}, delete the open middle third of every one of them; the surviving endpoints assemble into $\mathcal{C}_{n+1}$, a union of 2^{n+1} closed intervals each of length $3^{-(n+1)}$. Writing G_n for the (finite) union of the open intervals removed at stage n and $U = \bigcup_{n\geq 1} G_n$ for the total removed set, the *Cantor set* is what remains after all stages:

$$\mathcal{C} \;=\; \bigcap_{n=0}^{\infty} \mathcal{C}_n \;=\; [0,1] \setminus U.$$

Remark 1.2 (The surviving endpoints)*.* Every endpoint of every removed open interval belongs to $\mathcal{C}$: once created, such a point is an endpoint of some closed interval at every later stage and is never interior to a deleted middle third, so it is never removed. These endpoints—$0, 1, \frac{1}{3}, \frac{2}{3}, \frac{1}{9}, \frac{2}{9}, \frac{7}{9}, \frac{8}{9}, \ldots$—are triadic rationals $k/3^n$ and hence form only a *countable* subset of $\mathcal{C}$. They are far from all of it: by Theorem 2.1 the Cantor set is uncountable, so it also contains uncountably many non-endpoints, for instance $\frac{1}{4} = 0.\overline{02}_3 \in \mathcal{C}$, which is not an endpoint of any removed interval (see Figure 1).

Remark 1.3 (A point of $\mathcal{C}$ is an address)*.* It helps to read a point of $\mathcal{C}$ as an infinite itinerary of left/right choices: at each stage the current interval loses its middle third, and the point drops into either the *left* third (L, ternary digit 0) or the *right* third (R, ternary digit 2). Endpoints are precisely the itineraries that eventually *freeze*—all L's or all R's from some point on (e.g. $\frac{1}{3} = 0.0\overline{2}$, an L then R forever). Any itinerary that never freezes names a *non*-endpoint. The simplest is the alternating address $LRLR\cdots = 0.\overline{02}_3 = \frac{1}{4}$: as Figure 1 shows, $\frac{1}{4}$ is pinched between the shrinking interval endpoints, lying strictly inside at every stage—never at an edge, never in a removed gap—so it survives into $\mathcal{C}$ without ever being an endpoint.

Proposition 1.4 (Measure zero)**.** *The Cantor set has Lebesgue measure zero,* $\mathfrak{m}(\mathcal{C}) = 0$.

Proof. Since $\mathcal{C} \subseteq \mathcal{C}_n$ and $\mathcal{C}_n$ consists of 2^n intervals of length 3^{-n},

$$\mathfrak{m}(\mathcal{C}) \leq \mathfrak{m}(\mathcal{C}_n) = \left(\tfrac{2}{3}\right)^n \xrightarrow[n\to\infty]{} 0.$$

Equivalently, the total length removed is $\sum_{k\geq 1} 2^{k-1}3^{-k} = \frac{1}{3}\sum_{k\geq 0}\left(\frac{2}{3}\right)^k = 1$, so the complement of $\mathcal{C}$ in $[0,1]$ has full measure. $\square$

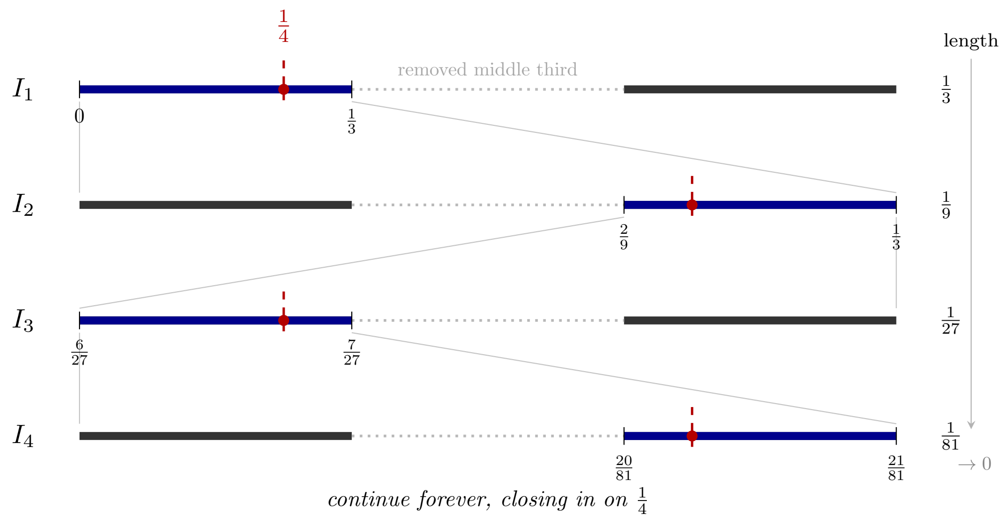

Figure 1: Zooming in on $\frac{1}{4}$. Each black-and-blue bar is the stage-n interval $I_n \ni \frac{1}{4}$, magnified to full width; its middle third (dotted) is deleted, and $\frac{1}{4}$ (red) lands in the surviving *left* or *right* third (blue), alternating $L, R, L, R, \ldots$ $(= 0.\overline{02}_3)$. At every stage $\frac{1}{4}$ lies *strictly between* the two endpoints of its interval—never at an endpoint, never in a gap—while the length 3^{-n} shrinks to 0. The nested intervals pinch down onto $\frac{1}{4}$, so $\frac{1}{4} \in \mathcal{C}$ although it is not an endpoint of any removed interval.

Proposition 1.5 (Ternary characterization)**.** *A point $x \in [0,1]$ lies in $\mathcal{C}$ if and only if it admits a base-3 (ternary) expansion*

$$x = \sum_{k=1}^{\infty} \frac{a_k}{3^k}, \qquad a_k \in \{0, 2\},$$

i.e. an expansion using only the digits 0 and 2.

Sketch. Writing $x = 0.a_1a_2a_3\cdots$ in base 3, the first deletion removes exactly those x forced to have $a_1 = 1$ (with the usual convention that endpoints $\frac{1}{3} = 0.0\overline{2}$ and $\frac{2}{3} = 0.2\overline{0}$ retain a digit-$\{0,2\}$ representation). Inductively, surviving stage n is equivalent to being able to choose $a_1, \ldots, a_n \in \{0, 2\}$. Passing to the intersection gives the claim. $\square$

2 The Cantor Set is Uncountable

Although $\mathcal{C}$ is "small" in the sense of measure (Proposition 1.4), it is "large" in the sense of cardinality.

Theorem 2.1. *The Cantor set $\mathcal{C}$ is uncountably infinite; in fact it has the cardinality of the continuum, $|\mathcal{C}| = |[0,1]| = 2^{\aleph_0}$.*

Proof. By Proposition 1.5, the map

$$\Phi : \{0,2\}^{\mathbb{N}} \longrightarrow \mathcal{C}, \qquad \Phi\big((a_k)_{k\geq 1}\big) = \sum_{k=1}^{\infty} \frac{a_k}{3^k},$$

is onto $\mathcal{C}$. It is also injective: two *distinct* ternary expansions can represent the same number only in the terminating/repeating pair $0.d_1 \cdots d_{n-1}d_n\overline{0} = 0.d_1 \cdots d_{n-1}(d_n - 1)\overline{2}$, which requires the digit d_n

to satisfy $d_n \in \{0,2\}$ *and* $d_n - 1 \in \{0,2\}$; no such digit exists. Hence Φ is a bijection from $\{0,2\}^{\mathbb{N}}$ onto $\mathcal{C}$.

Now $\{0,2\}^{\mathbb{N}}$ is in bijection with $\{0,1\}^{\mathbb{N}}$ (replace 2 by 1), and $\{0,1\}^{\mathbb{N}}$ is uncountable by Cantor's diagonal argument: given any sequence $s_1, s_2, \dots$ of elements of $\{0,1\}^{\mathbb{N}}$, the sequence $t = (t_k)$ with $t_k = 1 - s_k(k)$ differs from every s_j in its j-th coordinate, so no enumeration is complete. Therefore $\mathcal{C}$ is uncountable.

Finally, mapping each $(a_k) \in \{0,2\}^{\mathbb{N}}$ to $\sum_k (a_k/2)2^{-k} \in [0,1]$ is onto $[0,1]$ (every binary expansion arises), giving $|\mathcal{C}| \geq |[0,1]|$; the reverse inequality is trivial, so $|\mathcal{C}| = |[0,1]|$. □

Remark 2.2. Theorem 2.1 already signals trouble for the naive dichotomy "discrete versus continuous." A probability measure can be supported on an uncountable set of *measure zero*. That is precisely the Cantor distribution built next.

3 The Cantor Function

Definition 3.1 (Cantor function / Devil's staircase)**.** Define $F : [0,1] \to [0,1]$ as follows. Write $x = 0.a_1a_2a_3\cdots$ in base 3, and let

$$N = \min\{k : \ a_k = 1\} \in \{1,2,\dots\} \cup \{\infty\}.$$

Put $b_k = a_k/2$ for $k < N$ and $b_N = 1$, and set

$$F(x) = \sum_{k=1}^{N} \frac{b_k}{2^k}.$$

If $N = \infty$ (that is, $x \in \mathcal{C}$ so all $a_k \in \{0,2\}$), this reads

$$\boxed{F(x) = \sum_{k=1}^{\infty} \frac{a_k/2}{2^k} \quad \text{for } x \in \mathcal{C}.}$$

Extend to all of $\mathbb{R}$ by $F(x) = 0$ for $x < 0$ and $F(x) = 1$ for $x > 1$.

Equivalently and more geometrically (see Figure 2): F takes the constant value $\frac{1}{2}$ on the first deleted interval $(\frac{1}{3}, \frac{2}{3})$, the values $\frac{1}{4}$ and $\frac{3}{4}$ on the two deleted intervals of the next stage, and so on. On each deleted interval F is *constant*, and the constant equals the common F-value of the two endpoints (which lie in $\mathcal{C}$). One checks, e.g., $F(\frac{1}{3}) = F(0.0\overline{2}) = 0.0\overline{1}_2 = \frac{1}{2} = F(0.2\overline{0}) = F(\frac{2}{3})$, so the constant extension is consistent.

Lemma 3.2 (Monotonicity)**.** *F is non-decreasing on* $[0,1]$.

Proof. On $\mathcal{C}$, the assignment $a = (a_k) \mapsto b = (a_k/2)$ replaces each ternary digit in $\{0,2\}$ by the binary digit $a_k/2 \in \{0,1\}$; this is order preserving because if $x < y$ in $\mathcal{C}$ and m is the first index where their $\{0,2\}$-digits differ, then $a_m(x) < a_m(y)$ forces $b_m(x) < b_m(y)$ and hence $F(x) < F(y)$ (the leading differing digit dominates the tails). Thus F is non-decreasing on $\mathcal{C}$. On each deleted interval F is constant and equals its endpoint values, so monotonicity extends to all of $[0,1]$. □

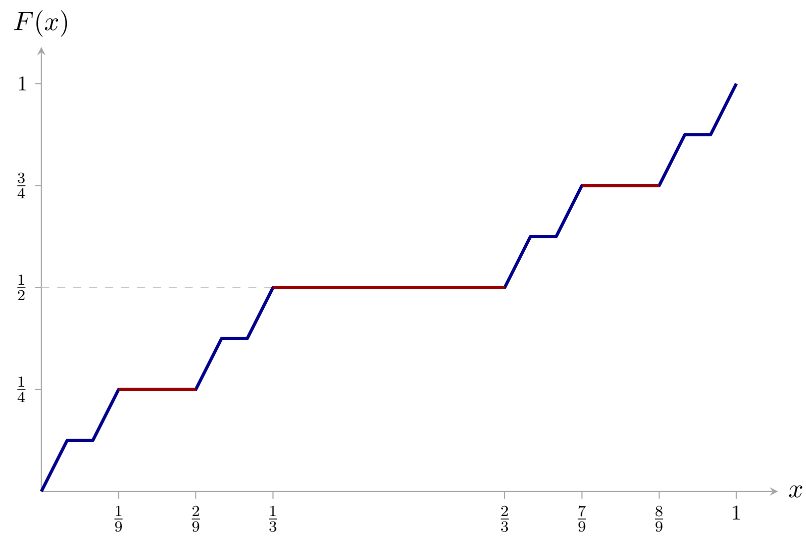

Figure 2: The Cantor function ("devil's staircase"), shown to stage 3. On each removed open interval F is flat (the red plateaus mark the stage-1 and stage-2 gaps $(\frac{1}{3}, \frac{2}{3})$, $(\frac{1}{9}, \frac{2}{9})$, $(\frac{7}{9}, \frac{8}{9})$, at heights $\frac{1}{2}, \frac{1}{4}, \frac{3}{4}$); all of the increase happens on the Cantor set $\mathcal{C}$, which has measure zero. The graph is continuous and non-decreasing, climbing from 0 to 1 while being flat almost everywhere.

4 The Cantor Function is Continuous

The key structural fact is that a monotone function whose *image* has no gaps cannot jump.

Lemma 4.1 (Monotone + surjective ⇒ continuous)**.** *Let $G : [a, b] \to \mathbb{R}$ be non-decreasing and suppose G is onto the interval $[G(a), G(b)]$. Then G is continuous on $[a, b]$.*

Proof. Being monotone, G has one-sided limits everywhere; write $G(x_0^-) = \lim_{x\uparrow x_0} G(x)$ and $G(x_0^+) = \lim_{x\downarrow x_0} G(x)$, with $G(x_0^-) \le G(x_0) \le G(x_0^+)$. Suppose G were discontinuous at an interior point x_0, say $G(x_0^-) < G(x_0)$ (the case $G(x_0) < G(x_0^+)$ is symmetric). Choose v with $G(x_0^-) < v < G(x_0)$. For $x < x_0$ we have $G(x) \le G(x_0^-) < v$, while for $x \ge x_0$ we have $G(x) \ge G(x_0) > v$. Hence v is *not* in the range of G, although $v \in (G(a), G(b))$, contradicting surjectivity. The endpoint cases $x_0 = a, b$ are handled by the corresponding one-sided argument. Therefore G is continuous. □

Theorem 4.2. *The Cantor function F of Definition 3.1 is continuous on $[0, 1]$ (indeed on all of $\mathbb{R}$), and maps $[0, 1]$ onto $[0, 1]$.*

Proof. By Lemma 3.2, F is non-decreasing with $F(0) = 0$ and $F(1) = 1$. We show F is onto $[0, 1]$. Given any $y \in [0, 1]$, take a binary expansion $y = \sum_{k\ge1} b_k 2^{-k}$ with $b_k \in \{0, 1\}$, set $a_k = 2b_k \in \{0, 2\}$, and let $x = \sum_{k\ge1} a_k 3^{-k} \in \mathcal{C}$. By Definition 3.1,

$$F(x) = \sum_{k\ge1} \frac{a_k/2}{2^k} = \sum_{k\ge1} \frac{b_k}{2^k} = y.$$

Thus $F(\mathcal{C}) = [0, 1]$, so in particular $F([0, 1]) = [0, 1]$. A non-decreasing surjection onto an interval is continuous by Lemma 4.1. Continuity on $\mathbb{R}$ follows since F is constant (0, then 1) outside $[0, 1]$ and matches at the endpoints. □

Remark 4.3 (Sharper statement). One can show F is in fact Hölder continuous of exponent $\alpha = \frac{\log 2}{\log 3} \approx 0.6309$: across any interval of one of the deleted intervals' generation the increment of F is controlled by the number of ternary digits resolved, giving $|F(x) - F(y)| \le C\,|x-y|^{\alpha}$. Continuity is all we need here.

5 The Cantor Function is a CDF

Definition 5.1 (CDF)**.** A function $F : \mathbb{R} \to [0,1]$ is a *cumulative distribution function* if

(F1) F is non-decreasing;

(F2) F is right-continuous, $F(x) = \lim_{t \downarrow x} F(t)$;

(F3) $\lim\limits_{x\to-\infty} F(x) = 0$ and $\lim\limits_{x\to+\infty} F(x) = 1$.

By the correspondence theorem, every such F is the CDF of a unique Borel probability measure μ on $\mathbb{R}$ determined by $\mu\big((a,b]\big) = F(b) - F(a)$.

Theorem 5.2. *The (extended) Cantor function F is a CDF. The associated law $\mu_{\mathcal{C}}$ — the* Cantor distribution *— is a Borel probability measure that assigns full mass to the Cantor set: $\mu_{\mathcal{C}}(\mathcal{C}) = 1$.*

Proof. Property (F1) is Lemma 3.2. Property (F2) is immediate from Theorem 4.2, since F is in fact continuous, hence certainly right-continuous. Property (F3) holds by construction: $F(x) = 0$ for $x \le 0$ and $F(x) = 1$ for $x \ge 1$. Thus F is a CDF and defines a probability measure $\mu_{\mathcal{C}}$.

For the last claim, the complement $[0,1] \setminus \mathcal{C}$ is the disjoint union of the deleted open intervals (a_i, b_i), on each of which F is constant; hence $\mu_{\mathcal{C}}\big((a_i, b_i]\big) = F(b_i) - F(a_i) = 0$. Summing over the countably many deleted intervals, $\mu_{\mathcal{C}}\big([0,1] \setminus \mathcal{C}\big) = 0$, so $\mu_{\mathcal{C}}(\mathcal{C}) = \mu_{\mathcal{C}}\big([0,1]\big) = 1$. ◻

6 The Cantor Function Has No Density

Definition 6.1 (Having a pdf)**.** A CDF F (equivalently, its law μ) *has a probability density function* if there is a nonnegative integrable $f : \mathbb{R} \to [0,\infty)$ with

$$F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t \qquad \text{for all } x \in \mathbb{R}.$$

Theorem 6.2. *The Cantor function F does* not *have a probability density function. The Cantor distribution is therefore continuous (its CDF has no jumps, so it has no atoms) yet not "continuous" in the pdf sense: it is* singular continuous*.*

Proof. Suppose, for contradiction, that such an $f \ge 0$ exists. Split

$$1 = F(1) - F(0) = \int_0^1 f(t)\,\mathrm{d}t = \int_{\mathcal{C}} f\,\mathrm{d}t \; + \; \int_{[0,1]\setminus\mathcal{C}} f\,\mathrm{d}t.$$

The first integral vanishes because $\mathfrak{m}(\mathcal{C}) = 0$ and the integral of any function over a Lebesgue-null set is 0. For the second, write $[0,1] \setminus \mathcal{C} = \bigsqcup_i (a_i, b_i)$ as the disjoint union of the deleted intervals. On $[a_i, b_i]$ the function F is constant (Theorem 5.2), so

$$\int_{a_i}^{b_i} f(t)\,\mathrm{d}t = F(b_i) - F(a_i) = 0.$$

By countable additivity of the integral of the nonnegative function f,

$$\int_{[0,1]\setminus\mathcal{C}} f\,\mathrm{d}t = \sum_i \int_{a_i}^{b_i} f(t)\,\mathrm{d}t = 0.$$

Adding, we obtain $1 = 0 + 0 = 0$, a contradiction. Hence no density exists. $\square$

Remark 6.3 (Derivative viewpoint). Because F is constant on every deleted interval and these fill up a set of measure 1, we have $F'(x) = 0$ for Lebesgue-a.e. $x \in [0,1]$. If a density f existed, the Lebesgue differentiation theorem would give $f = F' = 0$ a.e., whence $\int_{-\infty}^{x} f = 0$ for all x, contradicting $F(1) = 1$. This is the same obstruction phrased through the derivative: the "mass" of $\mu_{\mathcal{C}}$ hides entirely in the null set $\mathcal{C}$, where it is invisible to Lebesgue integration.

The correct language for "μ has a density" is *absolute continuity*, developed next.

7 Absolute Continuity

The Cantor law places all of its probability on the null set $\mathcal{C}$. The property that forbids this—and that a distribution must possess in order to have a density—is *absolute continuity*, which we phrase directly through sets of measure zero.

Definition 7.1 (Absolute continuity of measures)**.** Let μ, ν be measures on a measurable space $(\Omega, \mathcal{F})$. We say ν is *absolutely continuous with respect to* μ, written $\nu \ll \mu$, if

$$\mu(A) = 0 \implies \nu(A) = 0 \qquad \text{for all } A \in \mathcal{F}.$$

When ν is finite, this is equivalent to the ε–δ statement: for every $\varepsilon > 0$ there is $\delta > 0$ with $\mu(A) < \delta \Rightarrow \nu(A) < \varepsilon$. The opposite extreme is *mutual singularity* $\nu \perp \mu$: there is a set S with $\nu(S^c) = 0$ and $\mu(S) = 0$.

In this course the reference measure is almost always Lebesgue measure $\mathfrak{m}$ on $\mathbb{R}$. A law μ then *has a density* precisely when $\mu \ll \mathfrak{m}$—that is, when μ assigns zero probability to every set of Lebesgue measure zero. This is made precise by the Radon–Nikodym theorem in the next section; the density is the resulting derivative $\mathrm{d}\mu/\,\mathrm{d}\mathfrak{m}$.

Example 7.2 (The Cantor law is singular)**.** Let $\mathfrak{m}$ be Lebesgue measure on $[0,1]$ and $\mu_{\mathcal{C}}$ the Cantor distribution. Since $\mu_{\mathcal{C}}(\mathcal{C}) = 1$ while $\mathfrak{m}(\mathcal{C}) = 0$, taking $S = \mathcal{C}$ shows $\mu_{\mathcal{C}} \perp \mathfrak{m}$. In particular $\mu_{\mathcal{C}} \not\ll \mathfrak{m}$, which by the next section is *exactly* why it has no density.

Remark 7.3 (The function-side view: Fundamental Theorem of Lebesgue Calculus). The same dichotomy can be read off the CDF directly. A CDF F with law μ satisfies $\mu \ll \mathfrak{m}$ if and only if F is an indefinite integral of a density,

$$F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t \quad \text{for some integrable } f \geq 0,$$

in which case $F' = f$ almost everywhere; this equivalence is the *Fundamental Theorem of Lebesgue Calculus*. The Cantor function is the canonical failure: $F' = 0$ almost everywhere—it is flat on the removed intervals, whose union has full measure—yet F is not constant, so F cannot equal the integral of F'. Equivalently $\mu_{\mathcal{C}} \not\ll \mathfrak{m}$, in agreement with the previous example. It is the standard example separating "continuous" from "the integral of a density."

8 The Radon–Nikodym Derivative

Theorem 8.1 (Radon–Nikodym)**.** *Let μ, ν be σ-finite measures on $(\Omega, \mathcal{F})$ with $\nu \ll \mu$. Then there exists a measurable $g : \Omega \to [0, \infty)$, unique up to μ-almost-everywhere equality, such that*

$$\nu(A) = \int_A g \, \mathrm{d}\mu \qquad \text{for every } A \in \mathcal{F}.$$

The function g is called the Radon–Nikodym derivative *of ν with respect to μ and is written*

$$g = \frac{\mathrm{d}\nu}{\mathrm{d}\mu}.$$

Corollary 8.2 (A density *is* a Radon–Nikodym derivative)**.** *Let X be a real random variable with law μ (so $\mu(B) = \mathbb{P}(X \in B)$). Then X has a probability density function if and only if $\mu \ll \mathfrak{m}$, and in that case the pdf is exactly*

$$f = \frac{\mathrm{d}\mu}{\mathrm{d}\mathfrak{m}}.$$

Proof. If $\mu \ll \mathfrak{m}$, Theorem 8.1 yields $f = \mathrm{d}\mu / \mathrm{d}\mathfrak{m} \geq 0$ with $\mu(B) = \int_B f \, \mathrm{d}\mathfrak{m}$; taking $B = (-\infty, x]$ recovers Definition 6.1. Conversely, if a density f exists then $\mathfrak{m}(A) = 0 \Rightarrow \mu(A) = \int_A f \, \mathrm{d}\mathfrak{m} = 0$, so $\mu \ll \mathfrak{m}$. □

Remark 8.3 (Why the Cantor distribution has no density—final word). By Corollary 8.2, a density exists precisely when $\mu \ll \mathfrak{m}$. But $\mu_{\mathcal{C}} \perp \mathfrak{m}$, so $\mu_{\mathcal{C}}$ is about as far from absolutely continuous as possible, and no Radon–Nikodym derivative $\mathrm{d}\mu_{\mathcal{C}} / \mathrm{d}\mathfrak{m}$ exists. This reproves Theorem 6.2 conceptually: "no density" = "not absolutely continuous with respect to Lebesgue measure." The general *Lebesgue decomposition* writes any law as $\mu = \mu_{\mathrm{ac}} + \mu_{\mathrm{d}} + \mu_{\mathrm{sc}}$ (absolutely continuous + discrete + singular continuous); the Cantor law is the purely singular-continuous piece.

9 Probability Density Functions

Definition 9.1 (pdf)**.** A function $f : \mathbb{R} \to \mathbb{R}$ is a *probability density function* if

(D1) $f(x) \geq 0$ for (almost) every x (nonnegativity);

(D2) f is integrable with $\int_{-\infty}^{\infty} f(x) \, \mathrm{d}x = 1$ (normalization).

Given a pdf f, the induced random variable X satisfies, for Borel sets A and points $a \leq b$,

$$F(x) = \int_{-\infty}^{x} f(t) \, \mathrm{d}t, \qquad \mathbb{P}(a < X \leq b) = \int_a^b f(t) \, \mathrm{d}t, \qquad \mathbb{P}(X \in A) = \int_A f \, \mathrm{d}\mathfrak{m},$$

and $F'(x) = f(x)$ at every point of continuity of f (and a.e. in general). Expectations of measurable h are computed by the law-of-the-unconscious-statistician,

$$\mathbb{E}\big[h(X)\big] = \int_{-\infty}^{\infty} h(x) \, f(x) \, \mathrm{d}x.$$

Remark 9.2 (What a pdf is *not*). A pdf need not be continuous or bounded, may take values greater than 1, and is only determined up to modification on a Lebesgue-null set (any two densities of the same law agree a.e.). It is a *density* (a Radon–Nikodym derivative), not a probability: $\mathbb{P}(X = x) = 0$ for every single point in the continuous case.

10 Some Standard Continuous Laws

Example 10.1 (Uniform on $[a, b]$, $a < b$)**.**

$$f(x) = \frac{1}{b-a}\,\mathbf{1}_{[a,b]}(x), \qquad F(x) = \frac{x-a}{b-a} \quad (a \le x \le b).$$

$\mathbb{E}[X] = \dfrac{a+b}{2}, \quad \mathrm{Var}(X) = \dfrac{(b-a)^2}{12}$. It is the maximum-entropy law on $[a, b]$ and the reference for pseudorandom generation.

Example 10.2 (Exponential with rate $\lambda > 0$)**.**

$$f(x) = \lambda e^{-\lambda x}\,\mathbf{1}_{[0,\infty)}(x), \qquad F(x) = \big(1 - e^{-\lambda x}\big)\,\mathbf{1}_{[0,\infty)}(x).$$

$\mathbb{E}[X] = \dfrac{1}{\lambda}, \quad \mathrm{Var}(X) = \dfrac{1}{\lambda^2}$. It is the unique continuous *memoryless* law, $\mathbb{P}(X > s + t \mid X > s) = \mathbb{P}(X > t)$, and models waiting times in Poisson processes.

Example 10.3 (Gaussian / Normal $\mathcal{N}(\mu, \sigma^2)$, $\sigma > 0$)**.**

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\!\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \qquad x \in \mathbb{R}.$$

$\mathbb{E}[X] = \mu, \quad \mathrm{Var}(X) = \sigma^2$. The normalization $\int_{\mathbb{R}} f = 1$ follows from $\int_{\mathbb{R}} e^{-u^2}\,\mathrm{d}u = \sqrt{\pi}$. Its CDF has no elementary closed form; one writes $F(x) = \frac{1}{2}\big(1 + \mathrm{erf}\big(\frac{x-\mu}{\sigma\sqrt{2}}\big)\big)$. The Gaussian is central via the Central Limit Theorem and is the maximum-entropy law for fixed mean and variance.

Remark 10.4. Each example is absolutely continuous with respect to $\mathfrak{m}$; the displayed f is its Radon–Nikodym derivative $\mathrm{d}\mu/\,\mathrm{d}\mathfrak{m}$. Contrast the Cantor distribution, which has no such f.

11 What Does Integration Really Mean?

Every density statement above is an assertion about an *integral*: $F(x) = \int_{-\infty}^{x} f$, $\mathbb{P}(X \in A) = \int_A f$, $\mathbb{E}[h(X)] = \int hf$. The Cantor distribution showed that the meaning of "$\int$" is delicate: the naive expectation "a continuous CDF comes from integrating a density" is false. So we must ask, squarely:

$$\textit{What is the object } \int_a^b f(x)\,\mathrm{d}x, \textit{ exactly?}$$

The historically first rigorous answer is Riemann's.

Definition 11.1 (Tagged partitions and Riemann sums)**.** Let $f : [a, b] \to \mathbb{R}$. A *tagged partition* of $[a, b]$ consists of a partition

$$P: \quad a = x_0 < x_1 < \cdots < x_n = b,$$

together with *tags* $t_i \in [x_{i-1}, x_i]$, one chosen in each subinterval. Write $\Delta x_i = x_i - x_{i-1}$ and call $\|P\| = \max_i \Delta x_i$ the *mesh* of P. The associated *Riemann sum* is

$$S\big(f, P, \{t_i\}\big) = \sum_{i=1}^{n} f(t_i)\,\Delta x_i,$$

the total signed area of the rectangles with width Δx_i and height $f(t_i)$.

Definition 11.2 (The Riemann integral)**.** The function f is *Riemann integrable* on $[a, b]$, with *Riemann integral*

$$\int_a^b f(x)\,\mathrm{d}x \;=\; I,$$

if the Riemann sums converge to the number I as the mesh shrinks, *uniformly over the choice of tags*: for every $\varepsilon > 0$ there is a $\delta > 0$ such that *every* tagged partition $(P, \{t_i\})$ with mesh $\|P\| < \delta$ satisfies

$$\Big| S\big(f, P, \{t_i\}\big) - I \Big| = \Big| \sum_{i=1}^{n} f(t_i)\,\Delta x_i \;-\; I \Big| < \varepsilon.$$

Equivalently, $\displaystyle\int_a^b f(x)\,\mathrm{d}x = \lim_{\|P\|\to 0} \sum_i f(t_i)\,\Delta x_i$, the limit existing and taking the same value I no matter how the tags t_i are selected. When it exists, I is unique.

Theorem 11.3 (Lebesgue's characterization)**.** *A bounded $f : [a, b] \to \mathbb{R}$ is Riemann integrable if and only if its set of discontinuities has Lebesgue measure zero.*

Remark 11.4 (Why Riemann is not enough for us)*.* Theorem 11.3 exposes the limits of the Riemann integral. The Dirichlet function $\mathbf{1}_{\mathbb{Q}\cap[0,1]}$ is discontinuous everywhere, hence *not* Riemann integrable, even though it differs from 0 only on a null set and "should" integrate to 0. More to the point of this course, the Riemann integral integrates over intervals weighted by length; it has no native way to integrate against a singular law such as $\mu_{\mathcal{C}}$, nor to guarantee that pointwise or monotone limits of integrable functions remain integrable. The *Lebesgue integral* $\int f\,\mathrm{d}\mu$ repairs all of this: it integrates measurable f against an arbitrary measure μ, makes $\mathbf{1}_{\mathbb{Q}}$ integrable with integral 0, and supplies the convergence theorems (monotone, dominated) and the Radon–Nikodym theorem (Theorem 8.1) on which the whole theory of densities rests. It agrees with the Riemann integral whenever the latter exists, so nothing is lost—and the Cantor distribution is finally expressible as $\mathbb{P}(X \in A) = \mu_{\mathcal{C}}(A) = \int_A \mathrm{d}\mu_{\mathcal{C}}$, a legitimate integral with no density in sight.