7 The Radon-Nikodym-Theorem

Given: a measure space $(\Omega, \mathfrak{A}, \mu)$. Put $\overline{\mathfrak{Z}}_+ = \overline{\mathfrak{Z}}_+(\Omega, \mathfrak{A})$.

Definition 1. For f (quasi-)μ-integrable and $A \in \mathfrak{A}$, the *integral of f over A* is

$$\int_A f \, d\mu = \int 1_A \cdot f \, d\mu.$$

(Note: $|1_A \cdot f| \leq |f|$.)

Theorem 1. Let $f \in \overline{\mathfrak{Z}}_+$ and put

$$\nu(A) = \int_A f \, d\mu, \qquad A \in \mathfrak{A}.$$

Then ν is a measure on $\mathfrak{A}$.

Proof. Clearly $\nu(\emptyset) = 0$ and $\nu \geq 0$. For $A_1, A_2, \ldots \in \mathfrak{A}$ pairwise disjoint

$$\begin{aligned} \nu\Big(\bigcup_{i=1}^{\infty} A_i\Big) &= \int \sum_{i=1}^{\infty} 1_{A_i} \cdot f \, d\mu = \int \lim_{n\to\infty} \Big(\sum_{i=1}^{n} 1_{A_i} \cdot f\Big) \, d\mu \\ &= \lim_{n\to\infty} \int \sum_{i=1}^{n} 1_{A_i} \cdot f \, d\mu = \sum_{i=1}^{\infty} \int 1_{A_i} \cdot f \, d\mu \\ &= \sum_{i=1}^{\infty} \nu(A_i) \end{aligned}$$

follows from Theorem 5.1. □

Definition 2. The mapping ν in Theorem 1 is called *measure with μ-density f*, or distribution with density f. Notation: $\nu = f \cdot \mu$ (bad, but common notation: $\mathrm{d}\nu = d \cdot \mathrm{d}\mu$). If $\int f \, d\mu = 1$ then f is called *probability density*.

Example 1. The introductory examples of probability spaces were defined by means of probability densities.

(i) Let $(\Omega, \mathfrak{A}, \mu) = (\mathbb{R}^k, \mathfrak{B}_k, \lambda_k)$. For

$$f(x) = (2\pi)^{-k/2} \cdot \exp\left(-\tfrac{1}{2} \textstyle\sum_{i=1}^{k} x_i^2\right)$$

we get the *k-dimensional standard normal distribution ν*.

For $B \in \mathfrak{B}_k$ such that $0 < \lambda_k(B) < \infty$ and

$$f = \frac{1}{\lambda_k(B)} \cdot 1_B$$

we get the *uniform distribution on B*.

(ii) $\Omega = \mathbb{N}$, $\mathfrak{A} = \mathfrak{P}(\mathbb{N})$, μ the counting measure. A mapping $f : \Omega \to \mathfrak{R}_+$ (i.e., a sequence) is in $\mathfrak{L}^1$ iff it is an absolutely summable sequence (see Übung4.3a)), and for each such f and $A \subseteq \Omega$,

$$\forall\, A \in \mathfrak{A} : \nu(A) = \int_A f\, d\mu = \sum_{n \in A} f(n). \tag{1}$$

Conversely, *any* measure ν on $\mathfrak{A}$ is a measure with density with respect to μ: Put $f(\omega) := \nu(\{\omega\})$, then ((1)) holds.

Theorem 2. Let $\nu = f \cdot \mu$ with $f \in \overline{\mathfrak{Z}}_+$ and $g \in \overline{\mathfrak{Z}}(\Omega, \mathfrak{A})$. Then

$$g \text{ (quasi)-}\nu\text{-integrable} \quad \Leftrightarrow \quad g \cdot f \text{ (quasi)-}\mu\text{-integrable},$$

in which case

$$\int g\, d\nu = \int g \cdot f\, d\mu$$

Proof. First, assume that $g = 1_A$ with $A \in \mathfrak{A}$. Then the statements hold by definition. For $g \in \Sigma_+(\Omega, \mathfrak{A})$ we now use linearity of the integral. For $g \in \overline{\mathfrak{Z}}_+$ we take a sequence $(g_n)_{n \in \mathbb{N}}$ in $\Sigma_+(\Omega, \mathfrak{A})$ such that such that $g_n \uparrow g$. Then $g_n \cdot f \in \overline{\mathfrak{Z}}_+$ and $g_n \cdot f \uparrow g \cdot f$. Hence, by Theorem 5.1 and the previous part of the proof

$$\int g\, d\nu = \lim_{n \to \infty} \int g_n\, d\nu = \lim_{n \to \infty} \int g_n \cdot f\, d\mu = \int g \cdot f\, d\mu.$$

Finally, for $g \in \overline{\mathfrak{Z}}(\Omega, \mathfrak{A})$ we already know that

$$\int g^{\pm} d\nu = \int g^{\pm} \cdot f\, d\mu = \int (g \cdot f)^{\pm}\, d\mu.$$

Use linearity of the integral. □

Remark 1.

$$f, g \in \overline{\mathfrak{Z}}_+ \wedge f = g \;\mu\text{-a.e.} \quad \Rightarrow \quad f \cdot \mu = g \cdot \mu.$$

Theorem 3 (Uniqueness of densities). Let $f, g \in \overline{\mathfrak{Z}}_+$ such that $f \cdot \mu = g \cdot \mu$. Then

(i) f μ-integrable $\Rightarrow f = g$ μ-a.e.,

(ii) μ σ-finite $\Rightarrow f = g$ μ-a.e.

Proof. Ad (i): It suffices to verify the claim: If f, g μ-integrable and

$$\forall\, A \in \mathfrak{A} : \int_A f\, d\mu \le \int_A g\, d\mu \quad \Rightarrow \quad f \le g \;\mu\text{-a.e.}$$

To this end, take $A = \{f > g\}$. By assumption,

$$-\infty < \int_A f\, d\mu \le \int_A g\, d\mu < \infty$$

and therefore $\int_A (f-g)\,d\mu \le 0$. However,

$$1_A \cdot (f-g) \ge 0,$$

hence $\int_A (f-g)\,d\mu \ge 0$. Thus

$$\int 1_A \cdot (f-g)\,d\mu = 0.$$

Theorem 5.3 implies $1_A \cdot (f-g) = 0$ μ-a.e., and by definition of A we get $\mu(A) = 0$. Ad (ii): Assume first that μ is finite. Since for all $k \in \mathbb{N}$,

$$\infty \cdot \mu(\{f=\infty\} \setminus \{g \ge k\}) = \int_{\{f=\infty\}\setminus\{g\ge k\}} f \mathrm{d}\mu = \int_{\{f=\infty\}\setminus\{g\ge k\}} g \mathrm{d}\mu \le k\mu(\Omega)\ ,$$

we have that $\mu(\{f=\infty\} \setminus \{g \ge k\}) = 0$, and by σ–continuity from below, $\mu(\{f = \infty \setminus \{g=\infty\}\}) = 0$. By symmetry, we conclude

$$\mu(\{f=\infty\}\Delta\{g=\infty\}) = 0\ .$$

Set $A_0 = \{f=\infty\} \cup \{g=\infty\}$, $A_1 = A_0^c$; then $\mathbf{1}_{A_0} f = \mathbf{1}_{A_0} g$ μ–a.e., and we claim that

$$\mathbf{1}_{A_1} f = \mathbf{1}_{A_1} g \quad \mu\text{–a.e.}\ . \tag{2}$$

Since

$$A_1 \cap \{f > g\} = \bigcap_{n\in\mathbb{N}} \underbrace{\{n > f > g + 1/n\}}_{=:C_n}\ ,$$

we just have to show $\mu(C_n) = 0$. But

$$\int \mathbf{1}_{C_n} g \mathrm{d}\mu = \int \mathbf{1}_{C_n} f \mathrm{d}\mu \ge \int \mathbf{1}_{C_n} (g + 1/n) = \int \mathbf{1}_{C_n} g \mathrm{d}\mu + \mu(B_n)/n\ .$$

Since further

$$\int \mathbf{1}_{C_n} g \mathrm{d}\mu = \int \mathbf{1}_{C_n} f \mathrm{d}\mu \le n \cdot \mu(\Omega) < \infty\ ,$$

this entails $\mu(C_n) = 0$, and hence $\mu(A_1 \cap \{f > g\}) = 0$; by symmetry, also $\mu(A_1 \cap \{g > f\}) = 0$, i.e., (2) follows.
Let now μ be just σ–finite, and let $B_n \in \mathfrak{A}$ be disjoint such that $\mu(B_n) < \infty$, $\bigcup_n B_n = \Omega$. Set $\mu_n(A) := \mu(A \cap B_n)$. Then μ_n are measures, and for all $A \in \mathfrak{A}$,

$$\mu(A) = \sum_n \mu_n(A)\ .$$

Moreover, $f \cdot \mu_n = g \cdot \mu_n$, and by the first part we know that

$$f = g \qquad \mu_n - -\text{a.e.}, \quad \forall n \in \mathbb{N}\ .$$

But then

$$\mu(\{f\ not = g\}) = \sum_n \mu_n(\{f \ne g\}) = 0\ .$$

□

Remark 2. Let $(\Omega, \mathfrak{A}, \mu) = (\mathbb{R}^k, \mathfrak{B}_k, \lambda_k)$ and $x \in \mathbb{R}^k$. There is no density $f \in \overline{\mathfrak{Z}}_+$ w.r.t. λ_k such that $\delta_x = f \cdot \lambda_k$ (recall δ_x the Dirac point measure). This follows from $\varepsilon_x(\{x\}) = 1$ and

$$(f \cdot \lambda_k)(\{x\}) = \int_{\{x\}} f \, d\lambda_k = 0.$$

Definition 3. A measure ν on $\mathfrak{A}$ is *absolutely continuous w.r.t.* μ if

$$\forall A \in \mathfrak{A} : \mu(A) = 0 \Rightarrow \nu(A) = 0.$$

Notation: $\nu \ll \mu$.

Remark 3.

(i) $\nu = f \cdot \mu \Rightarrow \nu \ll \mu$.

(ii) In Remark 2 neither $\varepsilon_x \ll \lambda_1$ nor $\lambda_1 \ll \varepsilon_x$.

(iii) Let μ denote the counting measure on $\mathfrak{A}$. Then $\nu \ll \mu$ for every measure ν on $\mathfrak{A}$.

(iv) Let μ denote the counting measure on $\mathfrak{B}_1$. Then there is no density $f \in \overline{\mathfrak{Z}}_+$ such that $\lambda_1 = f \cdot \mu$.

Lemma 1. Let $f_n \xrightarrow{\mathfrak{L}^p} f$ and $A \in \mathfrak{A}$. If $p = 1$ or $\mu(A) < \infty$ then

$$\int_A f_n \, d\mu \to \int_A f \, d\mu.$$

Proof. For $p = 1$, this follows from

$$\Big| \int_A f_n \, d\mu - \int_A f \, d\mu \Big| \le \int_A |f_n - f| \, \mathrm{d}\mu \to 0 \; ;$$

if $\mu(A) < \infty$ and $p > 1$ set $1/q = 1 - 1/p$; then by Theorem 6.1,

$$\int \mathbf{1}_A \cdot |f_n - f| \, \mathrm{d}\mu \le \underbrace{\Big(\int \mathbf{1}_A^q \Big)^{1/q}}_{= \mu(A)^{1/q} < \infty} \cdot \underbrace{\Big(\int |f - f_n|^p \Big)^{1/p}}_{\to 0} .$$

□

Theorem 4 (Radon, Nikodym). For every σ-finite measure μ and every measure ν on $\mathfrak{A}$ we have

$$\nu \ll \mu \quad \Rightarrow \quad \exists \, f \in \overline{\mathfrak{Z}}_+ : \nu = f \cdot \mu.$$

Proof. We will prove this only for finite measures (since we need it only for finite measures).

Step 1: We assume the stronger condition

$$\forall A \in \mathfrak{A} : \nu(A) \le \mu(A) \wedge \mu(\Omega) < \infty.$$

A class $\mathfrak{U} = \{A_1, \ldots, A_n\}$ is called a (finite measurable) partition of Ω iff $A_1, \ldots, A_n \in \mathfrak{A}$ are pairwise disjoint and $\bigcup_{i=1}^n A_i = \Omega$. The set of all partitions is partially ordered by

$$\mathfrak{U} \sqsubset \mathfrak{V} \quad \text{iff} \quad \forall\, A \in \mathfrak{U}\, \exists\, B \in \mathfrak{V} : A \subset B.$$

The infimum of two partitions is given by

$$\mathfrak{U} \wedge \mathfrak{V} = \{A \cap B : A \in \mathfrak{U}, B \in \mathfrak{V}\}\,.$$

For any partition $\mathfrak{U}$ we define

$$f_{\mathfrak{U}} = \sum_{A \in \mathfrak{U}} \alpha_A \cdot 1_A$$

with

$$\alpha_A = \begin{cases} \nu(A)/\mu(A) & \text{if } \mu(A) > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Clearly $f_{\mathfrak{U}} \in \Sigma_+(\Omega, \sigma(\mathfrak{U})) \subset \Sigma_+(\Omega, \mathfrak{A})$, $\sigma(\mathfrak{U}) = \mathfrak{U}^+ \cup \{\emptyset\}$, and

$$\forall\, A \in \sigma(\mathfrak{U}) : \nu(A) = \int_A f_{\mathfrak{U}}\, d\mu.$$

(Thus we have $\nu|_{\sigma(\mathfrak{U})} = f_{\mathfrak{U}} \cdot \mu|_{\sigma(\mathfrak{U})}$.) Let $\mathfrak{U} \sqsubset \mathfrak{V}$ and $A \in \mathfrak{V}$. Then

$$\nu(A) = \int_A f_{\mathfrak{V}}\, d\mu = \int_A f_{\mathfrak{U}}\, d\mu,$$

since $A \in \sigma(\mathfrak{U})$. Hence

$$\int_A f_{\mathfrak{V}}^2\, d\mu = \int_A f_{\mathfrak{V}} \cdot f_{\mathfrak{U}}\, d\mu,$$

since $f_{\mathfrak{V}}|_A$ is constant, and therefore

$$0 \le \int (f_{\mathfrak{U}} - f_{\mathfrak{V}})^2\, d\mu = \int f_{\mathfrak{U}}^2\, d\mu - \int f_{\mathfrak{V}}^2\, d\mu. \tag{3}$$

Put

$$\beta = \sup \left\{ \int f_{\mathfrak{U}}^2\, d\mu : \mathfrak{U} \text{ partition} \right\},$$

and note that $0 \le \beta \le \mu(\Omega) < \infty$, since $f_{\mathfrak{U}} \le 1$. Consider a sequence of functions $f_n = f_{\mathfrak{U}_n}$ such that

$$\lim_{n\to\infty} \int f_n^2\, d\mu = \beta.$$

Due to (3) we may assume that $\mathfrak{U}_{n+1} \sqsubset \mathfrak{U}_n$. Then, by (3), $(f_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathfrak{L}^2$, so that there exists $f \in \mathfrak{L}^2$ with

$$\lim_{n\to\infty} \|f_n - f\|_2 = 0 \quad \wedge \quad 0 \le f \le 1 \ \mu\text{-a.e.},$$

see Theorem 6.3.

We claim that $\nu = f \cdot \mu$. Let $A \in \mathfrak{A}$. Put

$$\widetilde{\mathfrak{U}}_n = \mathfrak{U}_n \wedge \{A, A^c\}$$

and

$$\widetilde{f}_n = f_{\widetilde{\mathfrak{U}}_n}.$$

Then

$$\nu(A) = \int_A \widetilde{f}_n \, d\mu = \int_A f_n \, d\mu + \int_A (\widetilde{f}_n - f_n) \, d\mu,$$

and (3) yields $\lim_{n\to\infty} \|\widetilde{f}_n - f_n\|_2 = 0$. It remains to apply Lemma 1.

Step 2: We assume only that μ, ν are finite, and $\nu \ll \nu$. Then $\mu, \nu \leq \mu + \nu =: \tau$; by Step 1, we have densities $g, h : \Omega \to [0,1]$ with $\mu = g \cdot \tau$, $\nu = h \cdot \tau$. Since

$$\mu(\{g = 0\}) = \int_{\{g=0\}} \mathrm{d}\mu = \int_{\{g=0\}} g \; \mathrm{d}\tau = 0$$

and $\nu \ll \mu$, $\nu(\{g = 0\}) = 0$. The function

$$f(x) := \begin{cases} h(x)/g(x), & g(x) \neq 0, \\ 0, & g(x) = 0, \end{cases}$$

is now a density for ν:

$$\nu(A) = \int_{A\cap\{g\neq 0\}} \underbrace{h}_{=fg} \; \mathrm{d}\tau = \int_{A\cap\{g\neq 0\}} f \; \mathrm{d}\mu = \int_A f \; \mathrm{d}\mu \; .$$

□