Why can't we define PDFs in infinite dimensions?

Alen Alexanderian*

Abstract

Probability density functions are fundamental objects in probability and statistics. Yet, we cannot define them in the usual sense in infinite-dimensional separable Banach spaces. This is due to the fact that an analogue of the Lebesgue measure cannot be defined in such spaces. In this brief note, we provide a concise coverage of these well-known facts.

1 Introduction

A probability density function (PDF) is a familiar notion in probability and statistics. We usually take PDFs for granted. For example, when we talk about a normally distributed random variable, we usually think of the normal PDF and its graph—the bell curve. Another tool we take for granted is the Lebesgue measure. This is the canonical reference measure in finite dimensions. For example, for a real-valued random variable X with PDF f_{X}, we define its expectation by

$$E\{X\} = \int_{\mathbb{R}} x f_{\mathrm{X}}(x)\, dx.$$

The "dx" in this equation indicates integrating with respect to the Lebesgue measure. Random variables taking values in $\mathbb{R}^n$, with $n \in \mathbb{N}$, can be treated similarly. Namely, for a random n-vector $\boldsymbol{X}$ with PDF f_{X}, we have $E\{\boldsymbol{X}\} = \int_{\mathbb{R}^n} \boldsymbol{x} f_{\mathrm{X}}(\boldsymbol{x})\, d\boldsymbol{x}$. In this case, "$d\boldsymbol{x}$" indicates integrating with respect to the n-dimensional Lebesgue measure.

In some applications, one works with random variables taking values in an infinite-dimensional vector space. For example, this is the case when we study Bayesian formulations of inverse problems with infinite-dimensional parameters [3]. The setting of interest is the case where the random variable takes values in an infinite-dimensional separable real Banach space. It is known that we cannot define an analogue of the Lebesgue measure in such spaces. This also entails that we cannot define a PDF in the usual sense. In this brief note, we provide a brief discussion of these basic facts. We begin by recalling some basics about random variables in Section 2. Then, we discuss why an analogue of the Lebesgue measure cannot be defined in the infinite-dimensional setting in Section 3.

2 Random variables

Here we recall some basics from probability; see, e.g., [4] for further reading. A starting point in probabilistic formulations is to consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. In this triple, Ω is the sample space, $\mathcal{F}$ is a suitable σ-algebra on Ω, and $\mathbb{P}$ is a probability measure. Consider a random variable $\boldsymbol{X} : \Omega \to \mathbb{R}^n$. Equipping $\mathbb{R}^n$ with the Borel σ-algebra, $\mathcal{B}(\mathbb{R}^n)$, it is common to consider the "image probability space" $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mu_{\mathrm{X}})$, where μ_{X} is the probability measure

$$\mu_{\mathrm{X}}(E) = \mathbb{P}(\boldsymbol{X} \in E), \quad E \in \mathcal{B}(\mathbb{R}^n).$$

*North Carolina State University. E-mail: alexanderian@ncsu.edu
Last revised: January 1, 2023

The measure μ_X is known as the law of the random variable $\boldsymbol{X}$. If μ_X is absolutely continuous with respect to the Lebesgue measure, by the Radon–Nikodym theorem, there exists a unique (in an almost sure sense) non-negative integrable function f_X such that

$$\mu_X(E) = \int_E f_X(\boldsymbol{x})\, d\boldsymbol{x}, \quad E \in \mathcal{B}(\mathbb{R}^n). \tag{2.1}$$

This f_X is the so called Radon–Nikodym derivative of μ_X with respect to the Lebesgue measure. In probability and statistics, the function f_X in (2.1) is referred to as the *probability density function* of the random variable $\boldsymbol{X}$. Working in the image probability space and using PDFs are important concepts from a practical point of view. For example, consider the expected value of $\boldsymbol{X}$, which is given by

$$E\{\boldsymbol{X}\} = \int_\Omega \boldsymbol{X}(\omega)\, \mathbb{P}(d\omega) = \int_{\mathbb{R}^n} \boldsymbol{x} f_X(\boldsymbol{x})\, d\boldsymbol{x}.$$

While the first integral is over an abstract (unobservable) sample space Ω, the second one is an integral over $\mathbb{R}^n$.

Note that for a PDF to exist, the law μ_X must be absolutely continuous with respect to the Lebesgue measure. This is not guaranteed to always hold. However, many of the $\mathbb{R}^n$-valued random variables seen in applications do admit PDFs. The situation in the infinite-dimensional setting is different: we cannot define a Lebesgue measure in an infinite-dimensional separable Banach space. Hence, a PDF as we know of cannot be defined in the usual sense in the infinite-dimensional setting.

3 Non-existence of a Lebesgue measure in infinite dimensions

Let $(\mathcal{V}, \|\cdot\|)$ be a normed linear space with $\mathbb{R}$ or $\mathbb{C}$ as the base field. In what follows $B(x,r)$ denotes the open ball centered at x with radius r. We denote by S_1 the unit sphere centered at the origin, $S_1 = \{x \in \mathcal{V} : \|x\| = 1\}$. Moreover, for a subspace $M \subseteq \mathcal{V}$ and a point $p \in \mathcal{V}$, we define

$$\operatorname{dist}(p, M) = \inf_{m\in M} \|p - m\|.$$

The following result, known as Riesz's Lemma (see e.g., [2]), is well-known.

Lemma 3.1 (Riesz). *Let $(\mathcal{V}, \|\cdot\|)$ be a normed linear space and suppose M is a proper closed subspace of $\mathcal{V}$. Then for any $\alpha \in (0,1)$ there exists $x_\alpha \in S_1$ such that $\|x_\alpha - m\| \geq \alpha$ for all $m \in M$.*

Proof. Let $p \in \mathcal{V} \setminus M$ and note that $d := \operatorname{dist}(p, M) > 0$. Take $m_\alpha \in M$ such that $\|p - m_\alpha\| \leq d/\alpha$. Now, set $x_\alpha = \frac{p-m_\alpha}{\|p-m_\alpha\|}$. Thus, $\|x_\alpha\| = 1$ and for every $m \in M$, we have,

$$\begin{aligned} \|x_\alpha - m\| &= \left\| \frac{p - m_\alpha}{\|p - m_\alpha\|} - m \right\| \\ &= \frac{1}{\|p - m_\alpha\|} \left\| p - \big(m_\alpha + \|p - m_\alpha\|\, m\big) \right\| \geq d/(d/\alpha) = \alpha. \quad \square \end{aligned}$$

The following technical result is a consequence of Riesz's Lemma. The result is stated for $B(0,1)$ but can be easily generalized for any open ball.

Lemma 3.2. *Let $\mathcal{V}$ be an infinite dimensional normed linear space. Then there exists a countably infinite collection of disjoint balls $B(x_n, \varepsilon)$, for some $\varepsilon > 0$, inside $B(0,1)$.*

Proof. Let $y_1 \in S_1$ and let $M_1 = \operatorname{span}\{y_1\}$. By Riesz's Lemma, we know there exists $y_2 \in S_1$ such that $\|y_2 - m\| \geq 1/2$ for all $m \in M_1$. We let $M_2 = \operatorname{span}\{y_1, y_2\}$ and proceeding inductively, get $y_3, y_4, y_5, \ldots$, such that $y_n \in S_1$ for all n and for subspaces

$$M_n = \operatorname{span}\{y_1, \ldots, y_n\},$$

we have $\mathrm{dist}(y_{n+1}, M_n) \geq 1/2$. Successive application of Riesz's Lemma is justified, because for all n, M_n is finite-dimensional and is thus a proper closed subspace of $\mathscr{V}$. For the sequence $\{y_n\}_{n=1}^{\infty}$, we have $y_n \in S_1$ and $\|y_{n+1} - y_n\| \geq 1/2$ for all $n \in \mathbb{N}$; the latter also implies, $B(y_n, 1/4) \cap B(y_{n+1}, 1/4) = \emptyset$. Hence, the statement of the lemma holds with the collection of balls given by $\{B(x_n, \varepsilon)\}_{n=1}^{\infty}$, with $x_n = \frac{1}{2} y_n$ and $\varepsilon = 1/8$. □

For any measure on $(\mathscr{V}, \mathcal{B}(\mathscr{V}))$ to behave similar to the Lebesgue measure it must be a translation invariant positive measure that assigns a finite measure to open balls. We can use Lemma 3.2 to prove that such an analogue of the Lebesgue measure cannot be defined in an infinite-dimensional separable Banach space.

Proposition 3.3. *Let $\mathscr{V}$ be an infinite dimensional separable Banach space. Then there exists no non-trivial translation invariant positive measure on $(\mathscr{V}, \mathcal{B}(\mathscr{V}))$ that is finite on open balls.*

Proof. Suppose μ is a translation invariant positive measure on $(\mathscr{V}, \mathcal{B}(\mathscr{V}))$ that assigns finite measures to open balls. By Lemma 3.2 we know that $B(0, 1)$ contains a countably infinite collection of disjoint balls $\{B(x_n, \varepsilon)\}_{n=1}^{\infty}$. By translation invariance, $\mu(B(x_n, \varepsilon))$ is the same for every $n \in \mathbb{N}$. That is, $\mu(B(x_n, \varepsilon)) = \nu$ for a constant $\nu \in [0, \infty)$, for every $n \in \mathbb{N}$. If $\nu > 0$, then we have $\mu(B(0, 1)) \geq \mu(\cup_n B(x_n, \varepsilon)) = \sum_n \mu(B(x_n, \varepsilon)) = \sum_n \nu = \infty$, which is a contradiction. Note that we also used the fact that $\{B(x_n, \varepsilon)\}_{n=1}^{\infty}$ are disjoint. On the other hand, if $\nu = 0$, then by separability we can cover the whole space $\mathscr{V}$ with countably many open balls of radius ε and get that $\mu(\mathscr{V}) = 0$; i.e., μ is the trivial (zero) measure. □

4 Concluding remarks

We cannot define a PDF for function-valued random variables in the sense we are used to in finite dimensions. However, we can still consider their distribution law. The latter can be characterized, for example, using its Fourier transform; see, e.g., [1]. When applicable, we can also use the Radon–Nikodym Theorem to define the density of a probability measure with respect to another (reference) measure. For example, given two equivalent Gaussian measures μ_1 and μ_2 on a Hilbert space, we can define the Radon–Nikodym derivative $\frac{d\mu_1}{d\mu_2}$; see e.g., [1, Chapter 2]. A more explicit example comes from Bayesian inversion in infinite-dimensions. Consider a (properly formulated) Bayesian inverse problem on a real separable Hilbert space $\mathscr{H}$. In that context, the reference measure is defined by a prior probability measure μ_{pr} on $(\mathscr{H}, \mathcal{B}(\mathscr{H}))$. This prior measure defines a distribution law that encodes ones prior knowledge about an inversion parameter. The solution of the Bayesian inverse problem is given by a posterior probability measure μ_{post} on $(\mathscr{H}, \mathcal{B}(\mathscr{H}))$. This posterior measure is a distribution law for the inversion parameter that is conditioned on the observed data and is consistent with the prior. In this case, for an event $E \in \mathcal{B}(\mathscr{H})$,

$$\mu_{\mathrm{post}}(E) = \int_E \rho(z)\, \mu_{\mathrm{pr}}(dz),$$

where $\rho(z)$ is given by a suitably normalized data-likelihood. This follows from the infinite-dimensional version of the Bayes Theorem; see [3], for further details.

References

[1] Giuseppe Da Prato. *An introduction to infinite-dimensional analysis*. Springer Science & Business Media, 2006.

[2] Erwin Kreyszig. *Introductory functional analysis with applications*, volume 81. wiley New York, 1989.

[3] Andrew M Stuart. Inverse problems: a bayesian perspective. *Acta numerica*, 19:451–559, 2010.

[4] David Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.