2 Basic notions: infinite dimension

FOREWORD

Up to isomorphism, we have for each n only one n-dimensional Gaussian measure (like a Euclidean space). What about only one infinite-dimensional Gaussian measure (like a Hilbert space)?

Probability in infinite-dimensional spaces is usually overloaded with topological technicalities; spaces are required to be Hilbert or Banach or locally convex (or not), separable (or not), complete (or not) etc. A fuss! We need measurability, not continuity; a σ-field, not a topology. This is the way to a single infinite-dimensional Gaussian measure (up to isomorphism).

Naturally, the infinite-dimensional case is more technical than finite-dimensional. The crucial technique is compactness. Back to topology? Not quite so. The art is, to borrow compactness without paying dearly to topology. Look, how to do it...

2a Random elements of measurable spaces: formulations

Before turning to Gaussian measures on infinite-dimensional spaces we need some more general theory. First, recall that

* a *measurable space* (in other words, Borel space) is a set S endowed with a σ-field $\mathcal{B}$;
* $\mathcal{B} \subset 2^S$ is a *σ-field* (in other words, σ-algebra, or tribe) if it is closed under finite and countable operations (union, intersection and complement);
* a collection closed under finite operations is called a (Boolean) *algebra* (of sets);

- the σ-field $\sigma(C)$ *generated* by a given set $C \subset 2^S$ is the least σ-field that contains C; the algebra $\alpha(C)$ *generated* by C is defined similarly;
- a *measurable map* from $(S_1, \mathcal{B}_1)$ to $(S_2, \mathcal{B}_2)$ is $f : S_1 \to S_2$ such that $f^{-1}(B) \in \mathcal{B}_1$ for all $B \in \mathcal{B}_2$.

Each element of $\alpha(C)$ is a finite Boolean combination of elements of C; alas, the structure of $\sigma(C)$ is much more complicated. If C is finite then $\sigma(C) = \alpha(C)$ is also finite; if C is countable then $\alpha(C)$ is also countable, while $\sigma(C)$ can be of cardinality continuum (but not higher[1]). See also [2], Sect. 1.1.

2a1 Exercise. Let S, T be sets, $f : S \to T$ a map, $B_1, B_2, \dots \subset T$ and $A_k = f^{-1}(B_k)$. Then

$$\alpha(A_1, A_2, \dots) = \{f^{-1}(B) : B \in \alpha(B_1, B_2, \dots)\},$$
$$\sigma(A_1, A_2, \dots) = \{f^{-1}(B) : B \in \sigma(B_1, B_2, \dots)\}.$$

Prove it.

A set $C \subset 2^S$ leads to an equivalence relation,

$$(s_1 \equiv s_2) \iff \forall A \in C \, (s_1 \in A \iff s_2 \in A) .$$

2a2 Exercise. The sets C, $\alpha(C)$ and $\sigma(C)$ lead to the same equivalence relation.

Prove it.

If every two different points are non-equivalent, one says that C *separates points* of S.

A measurable space $(S, \mathcal{B})$ is

- *countably separated,* if some finite or countable subset (or subalgebra) of $\mathcal{B}$ separates points of S;
- *countably generated,* if some finite or countable subset (or subalgebra) of $\mathcal{B}$ generates $\mathcal{B}$ (as a σ-field).

The interval $(0, 1)$ with the Lebesgue (rather than Borel) σ-field is countably separated, but not countably generated.[2] See also [2], Sect. 8.6 (p. 288).

Let $(S, \mathcal{B})$ be a measurable space.

- A *probability measure* on $(S, \mathcal{B})$ is a countably additive function $\mu : \mathcal{B} \to [0, 1]$ such that $\mu(S) = 1$;

[1] See [2], Exercise 8 to Sect. 8.2, or [3], Problem 8 to Sect. 4.2.

[2] Since its cardinality is higher than continuum, see [3], Prop. 3.4.1.

* the *μ-completion* of $\mathcal{B}$ is the σ-field $\mathcal{B}_\mu = \{A : \exists B_-, B_+ \in \mathcal{B} \ \big(B_- \subset A \subset B_+,\ \mu(B_-) = \mu(B_+) \big) \}$; sets of $\mathcal{B}_\mu$ will be called *μ-measurable;*
* the extension of μ to $\mathcal{B}_\mu$ (it evidently exists and is unique) is called the *completion* of μ (and will be denoted by μ again);
* a probability space $(S, \mathcal{B}, \mu)$ is called *complete* if $\mathcal{B}_\mu = \mathcal{B}$;
* if f is a measurable map from $(S_1, \mathcal{B}_1)$ to $(S_2, \mathcal{B}_2)$ and μ is a probability measure on $(S_1, \mathcal{B}_1)$, then $f(\mu)$ (or $\mu \circ f^{-1}$) denotes the completion of the probability measure $B \mapsto \mu(f^{-1}(B))$ on $(S_2, \mathcal{B}_2)$.

All measures will be completed, unless otherwise stated. In particular, Lebesgue σ-field on $(0,1)$ is the completion of Borel σ-field w.r.t. Lebesgue measure. See also [2], Sect. 1.5 or [3], Sect. 3.3.

Unfortunately, a probability space is usually defined as just a measurable space endowed with a probability measure. Such 'probability spaces' can be quite bizarre.[1] The appropriate notion is known as 'standard probability space' (to be defined soon).[2]

In fact, a single probability space can serve all cases, namely, $(0,1)$ with Lebesgue measure. Accordingly, we define:

* an *event* is a measurable subset of $(0,1)$, or its equivalence class;
* the *probability* of an event is its Lebesgue measure;
* a *random element* of a countably separated measurable space $(S, \mathcal{B})$ is a measurable function $X : (0,1) \to S$, or its equivalence class;
* the *distribution* of X is the (complete!) measure $X(\mathrm{mes})$;
* two random elements of $(S, \mathcal{B})$ are *identically distributed,* if their distributions are equal.

The interval $(0,1)$ is endowed with Lebesgue σ-field and Lebesgue measure, unless otherwise stated.

Let $(S, \mathcal{B}, \mu)$ and $(S', \mathcal{B}', \mu')$ be probability spaces.

* A *measure preserving map* from $(S, \mathcal{B}, \mu)$ to $(S', \mathcal{B}', \mu')$ is a measurable map f from $(S, \mathcal{B}_\mu)$ to $(S', \mathcal{B}'_{\mu'})$ such that $f(\mu) = \mu'$, or the equivalence class of such map;
* a measure preserving map is *one-to-one* mod 0*,* if its restriction to some set of full measure[3] is one-to-one;
* a mod 0 *isomorphism* between $(S, \mathcal{B}, \mu)$ and $(S', \mathcal{B}', \mu')$ is an invertible map $f : S_1 \to S'_1$ for some sets $S_1 \subset S$, $S'_1 \subset S'$ of full measure, such

[1] See [3], Sections 3.5, 10.2 (Problem 6), 12.1 (Problems 1,2). Think also about $\mathbb{R}/\mathbb{Q}$, and $\mathbb{R}^T$.

[2] Most of my papers stipulate standardness of all probability spaces.

[3] A set of full measure in $(S, \mathcal{B}, \mu)$ is a set $A \in \mathcal{B}_\mu$ such that $\mu(A) = 1$.

that both f and $f^{-1} : S_1' \to S_1$ are measurable, measure preserving maps; in other words, for all $A \subset S_1$,

$$A \in \mathcal{B}_\mu \iff f(A) \in \mathcal{B}_{\mu'} ,$$
$$A \in \mathcal{B}_\mu \implies \mu(A) = \mu'(f(A)) ;$$

* a *nonatomic standard probability space* is a probability space isomorphic mod 0 to the probability space $(0,1)$;
* a *standard probability space* is a probability space isomorphic mod 0 to an interval with Lebesgue measure, a finite or countable set of atoms, or a combination of both.

Most of our definitions and arguments related to probability spaces will be invariant under mod 0 isomorphisms.

2a3 Theorem. For every nonatomic[1] random element X of a countably separated measurable space $(S, \mathcal{B})$ there exists a random element Y of $(S, \mathcal{B})$ such that X, Y are identically distributed and the map $Y : (0,1) \to S$ is one-to-one mod 0.

2a4 Theorem. Let Y be a random element of a countably separated measurable space $(S, \mathcal{B})$ such that the map $Y : (0,1) \to S$ is one-to-one mod 0. Then Y is a mod 0 isomorphism.

No matter how bizarre is a countably separated measurable space, a distribution (of a random element) turns it into a standard probability space.[2] Countable separation is essential.

These theorems (and the very idea of standard probability spaces) are due to Rokhlin (1949).[3]

2a5 Exercise. Let $(\Omega, \mathcal{F}, P)$ be a standard probability space, $(S, \mathcal{B})$ a countably separated measurable space, $X : \Omega \to S$ measurable, and μ the distribution of X. If X is one-to-one mod 0, then it is a mod 0 isomorphism between $(\Omega, \mathcal{F}, P)$ and $(S, \mathcal{B}_\mu, \mu)$.

Deduce it from Theorem 2a4.

Hint: first, get rid of atoms (if any).

2a6 Exercise. Let $(S, \mathcal{B})$ be a countably separated measurable space, μ the distribution of some random element of $(S, \mathcal{B})$, and $f : S \to \mathbb{R}^\infty$ a measurable

[1] We say that X is nonatomic if $\operatorname{mes} X^{-1}(\{s\}) = 0$ for all $s \in S$.

[2] If $(S, \mathcal{B})$ is bizarre enough then some probability measures on it cannot be distributions of random elements.

[3] See [7], especially Sect. 2.5. For a modernized presentation, see [8] (especially, Theorems 3-2 and 3-5) or [4] (especially, Prop. 9).

map. If f is one-to-one mod 0 then f is a mod 0 isomorphism between (S, μ) and $(\mathbb{R}^\infty, f(\mu))$.

Deduce it from 2a5. But first define the Borel σ-field on $\mathbb{R}^\infty$. How do you interprete the measurability of f?

2b Random elements of measurable spaces: proofs, and more facts

Littlewood's principles[1]

* Every (measurable) set is nearly a finite sum of intervals;
* every function (of class L^λ) is nearly continuous;
* every convergent sequence of functions is nearly uniformly convergent.

Consider measurable sets $A \subset (0,1)$ such that

$$\forall \varepsilon \; \exists U \; (\operatorname{mes} U < \varepsilon \text{ and } \mathbf{1}_A|_{\complement U} \text{ is continuous}) \, ; \tag{2b1}$$

here U runs over open sets, $\mathbf{1}_A$ is the indicator of A (1 on A and 0 on $\complement A$), and $\mathbf{1}_A|_{\complement U}$ is the restriction of $\mathbf{1}_A$ to the complement of U.

2b2 Exercise. Let $A_1, A_2, \dots$ satisfy (2b1): $\forall n \, \forall \varepsilon \, \exists U \, (\dots)$. Then they satisfy it uniformly: $\forall \varepsilon \, \exists U \, \forall n \, (\dots)$.

Prove it.

Hint: $\frac{\varepsilon}{2} + \frac{\varepsilon}{4} + \dots$

2b3 Exercise. The sets satisfying (2b1) are a σ-field.

Prove it.

2b4 Exercise. All measurable sets satisfy (2b1).

Prove it.

It follows easily that every measurable set coincides with a finite union of intervals outside some open set of arbitrarily small measure.

A sequence of sets $B_1, B_2, \dots \subset S$ may be treated as a map $f : S \to \{0,1\}^\infty$,

$$f(s) = (\mathbf{1}_{B_1}(s), \mathbf{1}_{B_2}(s), \dots) \, .$$

The map is one-to-one if and only if the sequence of sets separates points of S. Note that $B_k = f^{-1}(C_k)$ where $C_k = \{(x_1, x_2, \dots) : x_k = 1\}$. Elements of the algebra $\alpha(C_1, C_2, \dots)$ are called cylindrical sets in $\{0,1\}^\infty$; elements of the

[1] Sect. 4.1 of: J.E. Littlewood, "Lectures on the theory of functions", Oxford 1944. Quoted also by R. Durrett, "Probability: theory and examples" (second edition), 1996 (Appendix A4).

σ-field $\sigma(C_1, C_2, \dots)$ are called Borel sets in $\{0,1\}^\infty$. By 2a1, $\alpha(B_1, B_2, \dots)$ consists of inverse images of cylindrical sets, and $\sigma(B_1, B_2, \dots)$ consists of inverse images of Borel sets.[1]

Let $(S, \mathcal{B})$ be a measurable space. Clearly, $f : S \to \{0,1\}^\infty$ is measurable if and only if all $B_k = f^{-1}(C_k)$ belong to $\mathcal{B}$.

The set $\{0,1\}^\infty$ is also a compact metrizable topological space.[2] Cylindrical sets are the clopen sets (closed and open simultaneously); open sets are countable unions of cylindrical sets; closed sets are countable intersections of cylindrical sets.

2b5 Lemma. Let $f : (0,1) \to \{0,1\}^\infty$ be measurable, and $Z \subset \{0,1\}^\infty$ be a set[3] such that $f^{-1}(Z)$ is measurable. Then there exists a Borel set $B \subset \{0,1\}^\infty$ such that $B \subset Z \cap f\big((0,1)\big)$ and $\operatorname{mes} f^{-1}(Z \setminus B) = 0$.

Proof (sketch). For any ε, 2b2 and 2b4 give us a compact set $K_\varepsilon \subset f^{-1}(Z)$ such that $\operatorname{mes}\big(f^{-1}(Z) \setminus K_\varepsilon\big) < \varepsilon$ and $f|_{K_\varepsilon}$ is continuous. It follows that $f(K_\varepsilon)$ is compact, therefore, Borel measurable. We take $B = K_{\varepsilon_1} \cup K_{\varepsilon_2} \cup \dots$ for some $\varepsilon_n \to 0$. $\square$

Proof of Theorem 2a3 (sketch). We choose $B_1, B_2, \dots \in \mathcal{B}$ that separate points of S and embed S into $\{0,1\}^\infty$ by the one-to-one map $f(\cdot) = \big(\mathbf{1}_{B_1}(\cdot), \mathbf{1}_{B_2}(\cdot), \dots\big)$. We consider the Borel measure $\mu = f(X(\operatorname{mes}))$ on $\{0,1\}^\infty$.

Claim: the set $f(S)$ is of full measure. *Proof:* Lemma 2b5, applied to the function $f(X(\cdot))$ and the whole $\{0,1\}^\infty$, gives us a Borel set $B \subset f(X((0,1))) \subset f(S)$ such that $\operatorname{mes} X^{-1}(f^{-1}(\complement B)) = 0$, that is, $\mu(\complement B) = 0$.

Thus, the inverse function $f^{-1} : f(S) \to S$ is defined μ-almost everywhere on $\{0,1\}^\infty$.

Claim: f^{-1} is μ-measurable. *Proof:* let $A \in \mathcal{B}$; we have to prove that $f(A)$ is μ-measurable. Lemma 2b5 (again), applied to the function $f(X(\cdot))$ and the set $f(A)$ (note that $X^{-1}(A)$ is measurable) gives us a Borel set $B \subset f(A)$ such that $\operatorname{mes} X^{-1}(f^{-1}(f(A) \setminus B)) = 0$, that is, $\mu(B) = \operatorname{mes} X^{-1}(A)$. The same holds for $\complement A$; namely, $\mu(B') = \operatorname{mes} X^{-1}(\complement A)$ for some Borel set $B' \subset f(\complement A)$. Thus, $B \subset f(A) \subset \complement B'$ and $\mu(B) = \mu(\complement B')$.

So, f^{-1} is a one-to-one mod 0, measure preserving map from $(\{0,1\}^\infty, \mu)$ to $(S, X(\operatorname{mes}))$. It remains to construct a one-to-one mod 0, measure preserving map from $\big((0,1), \operatorname{mes}\big)$ to $(\{0,1\}^\infty, \mu)$. Here we use nonatomicity (of

[1] Thus, elements of $\sigma(B_1, B_2, \dots)$ are described, as far as Borel subsets of $\{0,1\}^\infty$ are evident…

[2] Homeomorphic to the Cantor set, see [2], Exercise 1 to Sect. 8.1 or [3], the example after Theorem 13.1.7.

[3] maybe, bizarre

X, therefore, of μ). We divide $(0,1)$ in two intervals according to the distribution of the first coordinate (that is, $\mu(C_1)$), then subdivide each interval in two according to the distribution of the first two coordinates, and so on. $\square$

Proof of Theorem 2a4 (sketch). Similarly to the proof of Theorem 2a3 we construct $f : S \to \{0,1\}^\infty$ and $\mu = f(Y(\mathrm{mes}))$. Given a measurable $A \subset (0,1)$, we have to prove that $Y(A)$ is measurable w.r.t. $Y(\mathrm{mes})$. We apply Lemma 2b5 to the function $f(Y(\cdot))$ and the set $Z = f(Y(A))$ (note that $Y^{-1}(f^{-1}(Z)) = A$ since Y is one-to-one). We get Borel $B \subset f(Y(A))$ such that $\operatorname{mes} Y^{-1}(f^{-1}(Z \setminus B)) = 0$, that is, $\operatorname{mes} A = \operatorname{mes} Y^{-1}(f^{-1}(B))$. We take $B_- = f^{-1}(B) \in \mathcal{B}$ getting $B_- \subset Y(A)$ and $\operatorname{mes} Y^{-1}(B_-) = \operatorname{mes} A$. The same argument holds for $\complement A$, which provides $B' \in \mathcal{B}$ such that $B' \subset Y(\complement A)$ and $\operatorname{mes} Y^{-1}(B') = \operatorname{mes} \complement A$. Taking $B_+ = \complement B'$ we get $B_- \subset Y(A) \subset B_+$ and $\operatorname{mes} Y^{-1}(B_-) = \operatorname{mes} Y^{-1}(B_+)$. $\square$

Only the first Littlewood's principle was used. The other two are also worthy of our attention.

2b6 Theorem. (Lusin) For every measurable $f : (0,1) \to \mathbb{R}$ and every $\varepsilon > 0$ there exists an open set $U \subset (0,1)$ such that $\operatorname{mes} U < \varepsilon$ and $f|_{\complement U}$ is continuous.

See also [3, 7.5.2] or [2, 7.4.3].

2b7 Exercise. Prove Lusin's theorem.

Hint: apply 2b2 (and 2b4) to $A_k = f^{-1}\big((-\infty, r_k)\big)$ where $(r_1, r_2, \dots)$ is a sequence dense in $\mathbb{R}$.

2b8 Theorem. (Egoroff)

For every sequence of measurable functions $f_1, f_2, \dots : (0,1) \to \mathbb{R}$ such that $f_n \to 0$ almost everywhere, and every $\varepsilon > 0$, there exists an open set $U \subset (0,1)$ such that $\operatorname{mes} U < \varepsilon$ and $f_n|_{\complement U} \to 0$ uniformly.

See also [3, 7.5.1] or [2, 3.1.3].

Proof (sketch). For every δ it is easy to get $\sup_{x \in \complement U} |f_n(x)| < \delta$ for all n large enough. Similarly to 1d4 (by $\frac{\varepsilon}{2} + \frac{\varepsilon}{4} + \dots$) we get it uniform in δ. $\square$

2b9 Remark. Unlike Lusin's theorem, Egoroff's theorem does not use the topology of $(0,1)$ and holds on every probability space (be it standard or not).

2b10 Exercise. (a) For every Borel probability measure on a separable metric space, and every $\varepsilon > 0$, there exists a precompact set of probability $> 1 - \varepsilon$.

(b) For every Borel probability measure on a complete separable metric space, and every $\varepsilon > 0$, there exists a compact set of probability $> 1 - \varepsilon$.

Prove it.

Hint: apply Egoroff's theorem (via 2b9) to $f_n(x) = \operatorname{dist}\big(x, \{x_1, \dots, x_n\}\big) = \min_{k=1,\dots,n} \operatorname{dist}(x, x_k)$, where $(x_1, x_2, \dots)$ is a dense sequence.

See also [3], 7.1.4 (Ulam's theorem).

2b11 Exercise. In 2b1–2b5, replace $\big((0,1), \mathrm{mes}\big)$ with an arbitrary complete separable metric space and an arbitrary Borel probability measure.

2b12 Exercise. A complete separable metric space, endowed with a completed Borel probability measure, is a standard probability space.

Prove it.

Hint. First, get rid of atoms. Then think, what happens to the proof of Theorems 2a3, 2a4 if X is the identical map of the given space to itself, but Y is still defined on $(0,1)$.

This is known as the *isomorphism theorem for measures* (or measure spaces, or probability spaces). See also [7], Sect. 2.4 and 2.7; [4], Example 1 and Prop. 6; [8], Theorems 2-3 and 4-3; [9], Theorem 3.4.23. The corresponding result for measurable spaces is deeper;[1] indeed, the right to neglect some small sets makes our life much easier.[2]

2c Random vectors

A vector is an element of a linear space (called also vector space). However, we have no generally accepted definition of a random vector in infinite dimension. In finite dimension the situation is simple, since every finite-dimensional linear space E carries its Borel σ-field (generated by all linear functionals). Dealing with an infinite-dimensional linear space E we always have a definite σ-field $\mathcal{B}$ on E, not derived from its linear structure.[3] It is not clear, what should be assumed about $\mathcal{B}$ in general.

For example, a separable (infinite-dimensional) Hilbert space (say, l_2) carries several well-known topologies, especially, the norm topology and the

[1] The Borel isomorphism theorem, see [9], Th. 3.3.13.

[2] However, in [9] the isomorphism theorem for measures is deduced from the Borel isomorphism theorem.

[3] And not invariant under the group of *all* linear automorphisms.

weak topology. They generate the same σ-field. This is quite typical; usually we have only one 'reasonable' σ-field, — the natural σ-field.

In most cases the natural σ-field is generated by (some) linear functionals. This is the case for all separable Banach spaces, for example, $L_p(0,1)$ for $p \in [1,\infty)$. However, for $p \in (0,1)$ the space $L_p(0,1)$ has no Borel measurable linear functionals (except for 0). (See also [3], Problem 2(a) to Sect. 6.4 and [5], Th. (9.10).) Of course, it is not a Banach space.[1] Rather, it is a separable F-space. (See also [3, Sect. 6.5].) It means, a linear topological space whose topology is Polish. And a Polish topology is, by definition, a topology that corresponds to some (at least one) complete separable metric (see also [9, Sect. 2.2], [2, Sect. 8.1], [3, Sect. 10.2]).

Let us try such a definition.

2c1 Definition. Let E be a linear space endowed with a countably separated σ-field $\mathcal{B}$. An E-valued *random vector* is a random element X of the measurable space $(E,\mathcal{B})$ whose distribution $\mu = X(\mathrm{mes})$ satisfies the following condition.

(2c2) There exist a linear subspace $E_1 \subset E$ of full measure and μ-measurable linear functionals $f_1, f_2, \dots : E_1 \to \mathbb{R}$ that separate points of E_1.

Note that E_1 and f_n need not be $\mathcal{B}$-measurable.

Some $(E,\mathcal{B})$ are such that (2c2) is satisfied always (that is, for every random element of $(E,\mathcal{B})$).

2c3 Exercise. If $\mathcal{B}$ is generated by a sequence of linear functionals $E \to \mathbb{R}$ then (2c2) is always satisfied.

Prove it.

Especially, this is the case for $\mathbb{R}^n$, $\mathbb{R}^\infty$, every separable Banach space, in particular L_p for $p \in [1,\infty)$. The nonseparable Banach space L_∞ also fits, provided that the σ-field is chosen appropriately. What about the F-space $L_p(0,1)$ for $p \in (0,1)$? Here 2c3 is not applicable, but still, (2c2) is always satisfied.[2]

2c4 Question. Is (2c2) always satisfied in every separable F-space? I do not know.[3]

---

[1] Nor a locally convex space.

[2] Hint: almost every point is a Lebesgue point for almost all functions of L_p...

[3] For standard Borel linear spaces the answer is negative, despite an old claim of A. Vershik (1968, "The axiomatics of measure theory in linear spaces", Soviet Math. Dokl. **9**:1, 68–72).

The distribution $\mu = X(\text{mes})$ of a random vector X turns E into a *standard* probability space $(E, \mathcal{B}_\mu, \mu)$.[1] The one-to-one linear operator $f : E_1 \to \mathbb{R}^\infty$,

$$f(x) = \bigl(f_1(x), f_2(x), \dots \bigr) \,,$$

is a mod 0 isomorphism between (E, μ) and $(\mathbb{R}^\infty, f(\mu))$ (recall 2a6). In this sense, $\mathbb{R}^\infty$ is a universal model for all random vectors.

A probability measure μ on $\mathbb{R}^n$ is uniquely determined by its Fourier transform (called also its characteristic function) $\varphi_\mu : \mathbb{R}^n \to \mathbb{C}$,

$$\varphi_\mu(x) = \int \mathrm{e}^{\mathrm{i}\langle x,y \rangle} \, \mu(\mathrm{d}y) \,,$$

see [3, 9.5.1].

2c5 Exercise. A probability measure μ on $\mathbb{R}^\infty$ is uniquely determined by the function $\varphi_\mu : \mathbb{R}_\infty \to \mathbb{C}$ defined by

$$\varphi_\mu(x) = \int \mathrm{e}^{\mathrm{i}\langle x,y \rangle} \, \mu(\mathrm{d}y) \,;$$

here $\mathbb{R}_\infty \subset \mathbb{R}^\infty$ is the subspace of all sequences having only finitely many nonzero coordinates each.

Prove it.

Hint: two probability measures equal on an algebra[2] $\mathcal{A}$ are equal on the σ-field $\sigma(\mathcal{A})$, see [2, 1.6.2] or [3, 3.2.7].

Strangely enough, measurable linear functionals on $(\mathbb{R}^\infty, \mu)$ are not always of the form $(x_1, x_2, \dots) \mapsto c_1 x_1 + c_2 x_2 + \dots$ and moreover, it may happen that a measurable linear functional f cannot be represented as the limit of a sequence of functionals of the form $c_1 x_1 + \dots + c_n x_n$. Like any measurable function, f can be represented as[3]

$$f(x_1, x_2, \dots) = \lim_{n \to \infty} g_n(x_1, \dots, x_n) \quad \mu\text{-almost everywhere,}$$

however, g_n are nonlinear functions, and cannot be chosen among linear functions, even though f is linear. This paradox is due to M. Kanter.[4]

---

[1] Irrespective of (2c2).

[2] Or just a π-system, that is, a collection closed under finite intersections.

[3] See Theorem (5.7) in Chapter 4 of Durrett's textbook (cited on page 14).

[4] See Example 2.3 in: M. Kanter, "Linear sample spaces and stable processes", J. Funct. Anal. **9**:4, 441–459 (1972). See also: W. Smolenski, "An abstract form of a counterexample of Marek Kanter", Lecture Notes in Math. **1080**, 288-291 (1984).

Think, how to translate it into the language of the space $(E, \mathcal{B}_\mu, \mu)$ and linear functionals f_n.

The space $L_0(0,1)$ consists of equivalence classes of *all* measurable functions $(0,1) \to \mathbb{R}$, and is endowed with a metrizable topology that corresponds to the convergence in measure (in probability), see [3, Sect. 9.2]. (It is another F-space.) Similarly, $L_0(\mu)$ is defined for every probability measure μ (on any space).

2c6 Exercise. Let $E, \mathcal{B}, X, \mu$ be as in 2c1. Then equivalence classes of all μ-measurable linear functionals on E are a *closed* linear subspace $L_0^{\mathrm{lin}}(\mu) \subset L_0(\mu)$.

Prove it.

Hint: first, a sequence convergent in probability contains a subsequence convergent a.s. [3, 9.2.1]; second, every lineal functional on a lineal subspace can be extended to the whole space.

(Nothing changes if we consider functionals on linear subspaces of full measure.)

Accordingly, the set $\{f(X(\cdot)) : f \in L_0^{\mathrm{lin}}(\mu)\} \subset L_0(0,1)$ is a closed linear subspace. Kanter's paradox shows that only some special subspaces of $L_0(0,1)$ may appear in this way.[1]

2d Gaussian random vectors

2d1 Definition. The *standard infinite-dimensional Gaussian measure* γ^∞ is the product of countably many copies of γ^1.

Equivalently, γ^∞ is the measure on the space $\mathbb{R}^\infty$ (of all sequences of reals) such that

$$\gamma^\infty\{(x_1, x_2, \dots) : (x_1, \dots, x_n) \in A\} = \gamma^n(A)$$

for all $n = 1, 2, \dots$ and all measurable $A \subset \mathbb{R}^n$. The measure γ^∞ is defined initially on the Borel σ-field of $\mathbb{R}^\infty$ (generated by the coordinate functions $(x_1, x_2, \dots) \mapsto x_n$) and extended to the completed σ-field (of all γ^∞-measurable sets).

2d2 Definition. Let E be a linear space endowed with a countably separated σ-field $\mathcal{B}$, and γ a (completed) probability measure on $(E, \mathcal{B})$.

[1] They must be closed in a special topology, see Def. 6 (p. 954) in: B.S. Tsirelson, "A natural modification of a random process and its application to stochastic functional series and Gaussian measures," Journal of Soviet Math. **16**:2, 940–956 (1981).

(a) γ is a *finite-dimensional centered Gaussian measure,* if for some $n \in \{0, 1, 2, \dots\}$ there exists a one-to-one linear operator $V : \mathbb{R}^n \to E$ such that $V(\gamma^n) = \gamma$.

(b) γ is an *infinite-dimensional centered Gaussian measure,* if there exists a linear subspace $S_1 \subset \mathbb{R}^\infty$ of full measure γ^∞, and a one-to-one linear operator $V : S_1 \to E$ such that $V(\gamma^\infty) = \gamma$, that is, for every $B \in \mathcal{B}$ the set $V^{-1}(B)$ is γ^∞-measurable and $\gamma^\infty(V^{-1}(B)) = \gamma(B)$.

The customary definition is different, see 2b4(b), but equivalent.

As before, we often omit the word 'centered'.

Clearly, every Gaussian measure γ has a dimension $\dim \gamma \in \{0, 1, 2, \dots\} \cup \{\infty\}$. However, the support is well-defined only when $\dim \gamma < \infty$ (as long as E is not topologized).

By a *Gaussian random vector* we mean a random vector whose distribution is a Gaussian measure.

2d3 Exercise. (a) For every $n \in \{0, 1, 2, \dots\}$ there exists a random vector $X : (0, 1) \to \mathbb{R}^n$ distributed γ^n.

(b) There exists a random vector $X : (0, 1) \to \mathbb{R}^\infty$ distributed γ^∞.

Prove it.

Hint: either construct X via binary digits,[1] or appeal to the isomorphism theorem for measures.

2d4 Exercise. Every Gaussian measure is the distribution of some Gaussian random vector, and turns the space into a *standard* probability space (either nonatomic, or a single atom).

Prove it.

Hint: the operator V of 2d2 is a mod 0 isomorphism by 2a5; (2c2) follows.

2d5 Theorem. Let $E, \mathcal{B}, X, \mu$ be as in 2c1, then the following two conditions are equivalent:

(a) μ is a centered Gaussian measure;

(b) each μ-measurable linear functional $f : E \to \mathbb{R}$ is distributed $\mathrm{N}(0, \sigma^2)$ for some σ (depending on f).[2]

The proof is given below, see 2d6–2d13.

The finite-dimensional counterpart of Theorem 2d5 is classical.[3] In the infinite dimension, the implication (a)$\Longrightarrow$(b) is due to Rozanov (1968), see [6], Sect. 9, Prop. 2 (p. 96) and the reference [Roz2] there. See also Th. 2.3 in

[1] See also [3], Sect. 8.2 (p. 200).

[2] That is, $f(\mu) = \mathrm{N}(0, \sigma^2)$.

[3] See [1, p. 4], [6, p. 11].

Kanter's paper (cited on p. 19), and [1, Sect. 2.10]. The implication (b)$\Longrightarrow$(a) is basically the same as [1, Th. 3.4.4].[1]

Let us prove (b)$\Longrightarrow$(a). Assuming (b) we have $L_0^{\text{lin}}(\mu) = L_2^{\text{lin}}(\mu) \subset L_2(\mu)$ and $f(\mu) = \mathrm{N}(0, \|f\|^2)$ for $f \in L_2^{\text{lin}}(\mu)$.

First, the finite-dimensional case: $\dim L_2^{\text{lin}}(\mu) = n < \infty$. We choose an orthonormal basis $(f_1, \dots, f_n)$ of $L_2^{\text{lin}}(\mu)$.

2d6 Exercise. Prove that $f_1, \dots, f_n$ separate points of a linear subspace E_1 of full measure.

Hint: by (2c2), some sequence of elements of $L_2^{\text{lin}}(\mu)$ separates points of some E_1; note however that E_1 of (2c2) need not fit here...

We consider the map $f : E_1 \to \mathbb{R}^n$, $f(x) = \big(f_1(x), \dots, f_n(x)\big)$; it is linear, one-to-one.

2d7 Exercise. Prove that $f(\mu) = \gamma^n$.

Hint: use 2c5, or rather its finite-dimensional counterpart.

Therefore $f(E_1) = \mathbb{R}^n$, and the inverse map $V = f^{-1} : \mathbb{R}^n \to E_1$ sends γ^n to μ, which means that μ is Gaussian.

We turn to the infinite-dimensional case: $\dim L_2^{\text{lin}}(\mu) = \infty$. We choose an orthonormal basis $(f_1, f_2, \dots)$ of $L_2^{\text{lin}}(\mu)$.

2d8 Exercise. Prove that $f_1, f_2, \dots$ separate points of a linear subspace E_1 of full measure.

Hint: similar to 2d6, but there is a difficulty: a series $\sum c_k f_k$ (with $\sum c_k^2 < \infty$) converges in L_2; what about convergence almost everywhere? In fact, it holds,[2] however, you do not need it; instead, choose a subsequence (of the sequence of finite sums) that converges almost everywhere.

We consider the map $f : E_1 \to \mathbb{R}^\infty$, $f(x) = \big(f_1(x), f_2(x), \dots\big)$; it is linear, one-to-one. But do not think that $f(E_1) = \mathbb{R}^\infty$.

2d9 Exercise. Prove that $f(\mu) = \gamma^\infty$.

Hint: similar to 2d7.

2d10 Exercise. Prove that μ is Gaussian.

Hint: by 2a6, f is a mod 0 isomorphism between (E, μ) and $(\mathbb{R}^\infty, \gamma^\infty)$. Consider $V = f^{-1} : f(E_1) \to E_1$ and prove that it is measurable.

---

[1] But the proof given there is much harder.

[2] See [3, 9.7.1].

The implication (b)$\Longrightarrow$(a) of Theorem 2d5 is thus proved.

Before proving (b) for an arbitrary Gaussian measure we will prove it for γ^∞. The natural map $\mathbb{R}_\infty \to L_2^{\rm lin}(\gamma^\infty)$ extends by continuity to a linear isometric map $l_2 \to L_2^{\rm lin}(\gamma^\infty)$. The image of l_2 is a (closed linear) subspace $G \subset L_2^{\rm lin}(\gamma^\infty)$. We want to prove that $G = L_2^{\rm lin}(\gamma^\infty) = L_0^{\rm lin}(\gamma^\infty)$.

2d11 Exercise. For every measurable linear functional f on $(\mathbb{R}^\infty, \gamma^\infty)$ and every n,

$$\biggl| \int \exp\bigl(\mathrm{i} f(x_1,\dots,x_n,0,0,\dots)\bigr)\, \gamma^n(\mathrm{d}x_1 \dots \mathrm{d}x_n) \biggr| \ge \biggl| \int \mathrm{e}^{\mathrm{i}f} \,\mathrm{d}\gamma^\infty \biggr| .$$

Prove it.

Hint: the latter integral decomposes into the product of two integrals.

2d12 Exercise. For every measurable linear functional f on $(\mathbb{R}^\infty, \gamma^\infty)$, its 'coordinates' $c_k = f(0,\dots,0,1,0,0,\dots)$ satisfy $\sum c_k^2 < \infty$.

Prove it.

Hint: first, $| \int \exp(\mathrm{i}\lambda f(x_1,\dots,x_n,0,0,\dots)\, \gamma^n(\mathrm{d}x_1 \dots \mathrm{d}x_n)| = \exp\bigl(-\frac12 \lambda^2 (c_1^2 + \dots + c_n^2)\bigr)$; second, $\int \mathrm{e}^{\mathrm{i}\lambda f}\, \mathrm{d}\gamma^\infty$ is continuous in λ (especially, at $\lambda = 0$).

2d13 Exercise. If a measurable linear functional f on $(\mathbb{R}^\infty, \gamma^\infty)$ vanishes on $\mathbb{R}_\infty$ then it vanishes almost everywhere.

Prove it.

Hint: use Kolmogorov's 0–1 law[1] and the symmetry of γ^∞ w.r.t. $x \mapsto (-x)$.

It follows that $G = L_0^{\rm lin}(\gamma^\infty)$, which gives us the implication (a)$\Longrightarrow$(b) of Theorem 2d5.

The theorem is proved.

2d14 Exercise. Let $E, \mathcal{B}, X, \mu$ be as in 2c1, and there exist μ-measurable linear functionals $f_1, f_2, \dots : E \to \mathbb{R}$ that separate points and are such that every linear combination $c_1 f_1 + \dots + c_n f_n$ (for every n) is distributed $\mathrm{N}(0, \sigma^2)$ for some σ. Then μ is Gaussian.

Prove it.

Hint: similar to the proof of Th. 2d5 (b)$\Longrightarrow$(a), but simpler: 2d8 is not needed.

2d15 Exercise. Let $(E, \mathcal{B})$ and $(E_1, \mathcal{B}_1)$ be linear spaces endowed with countably separated σ-fields, and V a measurable linear operator $(E, \mathcal{B}) \to (E_1, \mathcal{B}_1)$. Let γ be a Gaussian measure on E. Assume that the measure $\gamma_1 = V(\gamma)$ satisfies Condition (2c2). Then γ_1 is Gaussian.

Prove it. Is it enough if V measurable on $(E, \mathcal{B}_\gamma)$ rather than $(E, \mathcal{B})$?

[1] See [3, 8.4.4].

2d16 Question. Can it happen that γ_1 (of 2d15) does not satisfy Condition (2c2), and therefore is not a Gaussian measure? I do not know.

2e Gaussian conditioning

Till now (in 2c, 2d) the linear structure on E and the σ-field $\mathcal{B}$ on E were not related in any way. It does not harm when only a single Gaussian measure on E is considered. However, dealing simultaneously with several Gaussian measures on the same E we will need more.

2e1 Definition. A *measurable linear space* is a linear space E endowed with a σ-field $\mathcal{B}$ such that the maps $(x, y) \mapsto x+y$ and $(c, x) \mapsto cx$ are measurable.

That is, for every $B \in \mathcal{B}$ the set $\{(x, y) \in E \times E : x + y \in B\}$ belongs to the σ-field on $E \times E$ generated by all $A \times A'$ for $A, A' \in \mathcal{B}$; and the set $\{(c, x) \in \mathbb{R} \times E : cx \in B\}$ belongs to the σ-field on $\mathbb{R} \times E$ generated by all $A \times A'$ for Borel $A \subset \mathbb{R}$ and $A' \in \mathcal{B}$. See also [1, 2.12.1].

2e2 Exercise. Let E be a measurable linear space. Then for every n the map $(c_1, \dots, c_n; x_1, \dots, x_n) \mapsto c_1x_1 + \dots + c_nx_n$ is measurable.

Prove it.

Hint: first, try just $cx + y$.

2e3 Exercise. (a) Is $\mathbb{R}$, endowed with the Borel σ-field, a measurable linear space?

(b) Is $\mathbb{R}$, endowed with the Lebesgue σ-field, a measurable linear space?

(c) What about $\mathbb{R}^n$, $\mathbb{R}^\infty$, Hilbert spaces, Banach spaces, F-spaces (with the Borel σ-field)?

2e4 Theorem. Let $(E, \mathcal{B}), (E_1, \mathcal{B}_1)$ be countably separated measurable linear spaces, γ a Gaussian measure on E, γ_1 a Gaussian measure on E_1, and V a measure preserving linear map $(E, \gamma) \to (E_1, \gamma_1)$. Then there exist a Gaussian measure γ_2 on E and a measurable linear map $\tilde{V} : (E_1, (\mathcal{B}_1)_{\gamma_1}) \to (E, \mathcal{B})$ such that

$$\int f \, \mathrm{d}\gamma = \int_{E_1} \left(\int_E f(\tilde{V}(x) + y) \, \gamma_2(\mathrm{d}y) \right) \gamma_1(\mathrm{d}x)$$

for every bounded γ-measurable function $f : E \to \mathbb{R}$.

Proof (sketch). The map V induces a linear isometric embedding $L_2^{\text{lin}}(\gamma_1) \subset L_2^{\text{lin}}(\gamma)$; we have $L_2^{\text{lin}}(\gamma) = L_2^{\text{lin}}(\gamma_1) \oplus (L_2^{\text{lin}}(\gamma) \ominus L_2^{\text{lin}}(\gamma_1))$. Assume that both summands are infinite-dimensional (other cases are similar but simpler). We choose an orthonormal basis $(f_1, f_2, \dots)$ of $L_2^{\text{lin}}(\gamma)$ such that $f_{2n-1} \in L_2^{\text{lin}}(\gamma_1)$, $f_{2n} \perp L_2^{\text{lin}}(\gamma_1)$. Using mod 0 isomorphisms we reduce the general case to the following special case: $\gamma = \gamma^\infty$, $E \subset \mathbb{R}^\infty$ is a subspace of full measure, and $V(x_1, x_2, \dots) = (x_1, x_3, x_5, \dots)$.

We want to construct γ_2 as the distribution of $(0, x_2, 0, x_4, \dots)$, and $\tilde{V}$ as $\tilde{V}(x_1, x_3, \dots) = (x_1, 0, x_3, 0, \dots)$. It remains to prove that $(0, x_2, 0, x_4, \dots)$ and $(x_1, 0, x_3, 0, \dots)$ belong to E for γ^∞-almost all $(x_1, x_2, \dots)$, and check measurability of $\tilde{V}$.

The measure γ^∞ is invariant under the map $(x_1, x_2, \dots) \mapsto (x_1, -x_2, x_3, -x_4, \dots)$. Thus, γ^∞-almost all $(x_1, x_2, \dots)$ satisfy both $(x_1, x_2, \dots) \in E$ and $(x_1, -x_2, x_3, -x_4, \dots) \in E$; therefore $(x_1, 0, x_3, 0, \dots) \in E$, as well as $(0, x_2, 0, x_4, \dots) \in E$.

Both maps $(\mathbb{R}^\infty, \gamma^\infty) \to E$ considered above are measurable, when $\mathbb{R}^\infty$ is endowed with the σ-field of all γ^∞-measurable sets, and E — with $\mathcal{B}_\gamma$. The more so, when E is endowed with $\mathcal{B}$. Measurability of $\tilde{V}$ follows. $\square$

2e5 Exercise. (a) Is $\tilde{V}$ measurable from $(E_1, (\mathcal{B}_1)_{\gamma_1})$ to $(E, \mathcal{B}_\gamma)$?

(b) Is $(E, \mathcal{B}_\gamma)$ a measurable linear space?

References

[1] V.I. Bogachev, *Gaussian measures,* AMS 1998.

[2] D.L. Cohn, *Measure theory,* Birkhäuser 1980, 1993.

[3] R.M. Dudley, *Real analysis and probability,* Wadsworth 1989.

[4] J. Haezendonck, *Abstract Lebesgue-Rohlin spaces,* Bulletin de la Société Mathématique de Belgique **25**, 243–258 (1973).

[5] A.S. Kechris, *Classical descriptive set theory,* Springer 1995.

[6] M.A. Lifshits, *Gaussian random functions,* Kluwer 1995.

[7] V.A. Rohlin, *On the fundamental ideas of measure theory,* Translations (Amer. Math. Soc.) Series 1, Vol. **10**, 1–54 (1962). (Translated from Russian, 1949.)

[8] T. de la Rue, *Espaces de Lebesgue,* Lect. Notes Math (Séminaire de Probabilités XXVII), Springer, Berlin, **1557**, 15–21 (1993).

[9] S.M. Srivastava, *A course on Borel sets,* Springer 1998.

Index