

Chapter 1

Measure theory and probability

Aim and contents

This chapter contains a number of exercises, aimed at familiarizing the reader with some important measure theoretic concepts, such as: Monotone Class Theorem (Williams [98], II.3, II.4, II.13), uniform integrability (which is often needed when working with a family of probabilities, see Dellacherie and Meyer [23]), L^p convergence (Jacod and Protter [40], Chapter 23), conditioning (this will be developed in a more probabilistic manner in the following chapters), absolute continuity (Fristedt and Gray [33], p. 118).

We would like to emphasize the importance for every probabilist to stand on a “reasonably” solid measure theoretic (back)ground for which we recommend, e.g., Revuz [74].

Exercise **1.12** plays a unifying role, and highlights the fact that the operation of taking a conditional expectation is a contraction (in L^2 , but also in every L^p) in a strong sense.

* 1.1 Some traps concerning the union of σ -fields

1. Show that the union of two σ -fields is never a σ -field unless one is included into the other.
2. Give an example of a filtration $(\mathcal{F}_n)_{n \geq 0}$ which is strictly increasing, i.e. $\mathcal{F}_n \neq \mathcal{F}_{n+1}$, for all n and such that $\cup_{n \geq 0} \mathcal{F}_n$ is not a σ -field.

Comments:

- (a) It is clear from the proof that the assertion of question 1 is also valid if we consider only Boole algebras instead of σ -fields.
- (b) Question 2 raises the following more general problem: are there examples of strictly increasing filtrations $(\mathcal{F}_n)_{n \geq 0}$, such that the union $\cup_{n \geq 0} \mathcal{F}_n$ is a σ -field? The answer to this question is NO. A proof of this fact can be found in: A. Broughton and B. W. Huff: A comment on unions of sigma-fields. *The American Mathematical Monthly*, **84**, no. 7 (Aug.–Sep., 1977), pp. 553–554. We are very grateful to Rodolphe Garbit and Rongli Liu, who have indicated to us some references about this question.

** 1.2 Sets which do not belong in a strong sense, to a σ -field

Let (Ω, \mathcal{F}, P) be a complete probability space. We consider two (\mathcal{F}, P) complete sub- σ -fields of \mathcal{F} , \mathcal{A} and \mathcal{B} , and a set $A \in \mathcal{A}$.

The aim of this exercise is to study the property:

$$0 < P(A|\mathcal{B}) < 1, \quad P \text{ a.s.} \quad (1.2.1)$$

1. Show that (1.2.1) holds if and only (iff) there exists a probability Q , which is equivalent to P on \mathcal{F} , and such that

$$(a) \quad 0 < Q(A) < 1, \quad \text{and} \quad (b) \quad \mathcal{B} \text{ and } A \text{ are independent.}$$

Hint. If (1.2.1) holds, we may consider, for $0 < \alpha < 1$, the probability:

$$Q_\alpha = \left\{ \alpha \frac{1_A}{P(A|\mathcal{B})} + (1 - \alpha) \frac{1_{A^c}}{P(A^c|\mathcal{B})} \right\} \cdot P.$$

2. Assume that (1.2.1) holds. Define $\mathcal{B}^A = \mathcal{B} \vee \sigma(A)$. Let $0 < \alpha < 1$, and Q be a probability which satisfies (a) and (b) together with:

$$(c) \quad Q(A) = \alpha, \quad \text{and} \quad (d) \quad \frac{dQ}{dP} \Big|_{\mathcal{F}} \text{ is } \mathcal{B}^A\text{-measurable.}$$

Show then the existence of a \mathcal{B} -measurable r.v. Z , which is > 0 , P a.s., and such that:

$$E_P(Z) = 1, \quad \text{and} \quad Q = Z \left\{ \alpha \frac{1_A}{P(A|\mathcal{B})} + (1 - \alpha) \frac{1_{A^c}}{P(A^c|\mathcal{B})} \right\} \cdot P.$$

Show that there exists a unique probability \hat{Q} which satisfies (a) and (b), together with (c), (d) and (e), where:

$$(e) \quad : \hat{Q} \Big|_{\mathcal{B}} = P \Big|_{\mathcal{B}}.$$

1. Measure theory and probability

3

3. We assume, in this and the two next questions, that $\mathcal{A} = \mathcal{B}^A$, but it is not assumed *a priori* that A satisfies (1.2.1).
 Show then that $A' \in \mathcal{A}$ satisfies (1.2.1) iff the two following conditions are satisfied:
 (f) there exists $B \in \mathcal{B}$ such that: $A' = (B \cap A) \cup (B^c \cap A^c)$, up to a negligible set, and
 (g) A satisfies (1.2.1).
 Consequently, if A does not satisfy (1.2.1), then there exists no set $A' \in \mathcal{A}$ which satisfies (1.2.1).
4. We assume, in this question and in the next one, that $\mathcal{A} = \mathcal{B}^A$, and that A satisfies (1.2.1).
 Show that, if \mathcal{B} is not P -trivial, then there exists a σ -field \mathcal{A}' such that $\mathcal{B} \subsetneq \mathcal{A}' \subsetneq \mathcal{A}$, and that no set in \mathcal{A}' satisfies (1.2.1).
5. (i) We further assume that, under P , A is independent of \mathcal{B} , and that:
 $P(A) = \frac{1}{2}$.
 Show that $A' \in \mathcal{A}$ satisfies (1.2.1) iff A' is P -independent of \mathcal{B} , and
 $P(A') = \frac{1}{2}$.
- (ii) We now assume that, under P , A is independent of \mathcal{B} , and that:
 $P(A) = \alpha$, with: $\alpha \neq \{0, \frac{1}{2}, 1\}$.
 Show that A' (belonging to \mathcal{A} , and assumed to be non-trivial) is independent of \mathcal{B} iff $A' = A$ or $A' = A^c$.
- (iii) Finally, we only assume that A satisfies (1.2.1).
 Show that, if $A' (\in \mathcal{A})$ satisfies (1.2.1), then the equality $\mathcal{A} = \mathcal{B}^{A'}$ holds.

Comments and references. The hypothesis (1.2.1) made at the beginning of the exercise means that A does not belong, in a strong sense, to \mathcal{B} . Such a property plays an important role in:

J. AZÉMA AND M. YOR: Sur les zéros des martingales continues. *Séminaire de Probabilités XXVI*, 248–306, *Lecture Notes in Mathematics*, **1526**, Springer, 1992.

** 1.3 Some criteria for uniform integrability

Consider, on a probability space (Ω, \mathcal{A}, P) , a set H of r.v.s with values in \mathbb{R}_+ , which is bounded in L^1 , i.e.

$$\sup_{X \in H} E(X) < \infty .$$

Recall that H is said to be uniformly integrable if the following property holds:

$$\sup_{X \in H} \int_{(X > a)} X dP \xrightarrow{a \rightarrow \infty} 0 . \quad (1.3.1)$$

To each variable $X \in H$ associate the positive, bounded measure ν_X defined by:

$$\nu_X(A) = \int_A X dP \quad (A \in \mathcal{A}).$$

Show that the property (1.3.1) is equivalent to each of the three following properties:

- (i) the measures $(\nu_X, X \in H)$ are equi-absolutely continuous with respect to P , i.e. they satisfy the criterion:

$$\forall \varepsilon > 0, \exists \delta > 0, \forall A \in \mathcal{A}, P(A) \leq \delta \Rightarrow \sup_{X \in H} \nu_X(A) \leq \varepsilon, \quad (1.3.2)$$

- (ii) for any sequence (A_n) of sets in \mathcal{A} , which decrease to \emptyset , then:

$$\lim_{n \rightarrow \infty} \left(\sup_{X \in H} \nu_X(A_n) \right) = 0, \quad (1.3.3)$$

- (iii) for any sequence (B_n) of disjoint sets of \mathcal{A} ,

$$\lim_{n \rightarrow \infty} \left(\sup_{X \in H} \nu_X(B_n) \right) = 0. \quad (1.3.4)$$

Comments and references:

- (a) The equivalence between properties (1.3.1) and (1.3.2) is quite classical; their equivalence with (1.3.3) and *a fortiori* (1.3.4) may be less known. These equivalences play an important role in the study of weak compactness in:

C. DELLACHERIE, P.A. MEYER AND M. YOR: Sur certaines propriétés des espaces H^1 et BMO, *Séminaire de Probabilités XII*, 98–113, *Lecture Notes in Mathematics*, **649**, Springer, 1978.

- (b) De la Vallée-Poussin's lemma is another very useful criterion for uniform integrability (see Meyer [60]; one may also consult: C. Dellacherie and P.A. Meyer [23]).

The lemma asserts that $(X_i, i \in I)$ is uniformly integrable if and only if there exists a strictly increasing function $\Phi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, such that $\frac{\Phi(x)}{x} \rightarrow \infty$, as $x \rightarrow \infty$ and $\sup_{i \in I} E[\Phi(X_i)] < \infty$. (Prove that the condition is sufficient!) This lemma is often used (in one direction) with $\Phi(x) = x^2$, i.e. a family $(X_i, i \in I)$ which is bounded in L^2 is uniformly integrable. See Exercise 5.9 for an application.

* 1.4 When does weak convergence imply the convergence of expectations?

Consider, on a probability space (Ω, \mathcal{A}, P) , a sequence (X_n) of r.v.s with values in \mathbb{R}_+ , which are uniformly integrable, and such that:

$$X_n(P) \xrightarrow[n \rightarrow \infty]{w} \nu.$$

1. Measure theory and probability

5

1. Show that ν is carried by \mathbb{R}_+ , and that $\int \nu(dx)x < \infty$.
2. Show that $E(X_n)$ converges, as $n \rightarrow \infty$, towards $\int \nu(dx)x$.

Comments:

- (a) Recall that, if $(\nu_n; n \in \mathbb{N})$ is a sequence of probability measures on \mathbb{R}^d (for simplicity), and ν is also a probability on \mathbb{R}^d , then:

$$\nu_n \xrightarrow[n \rightarrow \infty]{w} \nu \text{ if and only if: } \langle \nu_n, f \rangle \xrightarrow[n \rightarrow \infty]{} \langle \nu, f \rangle$$

for every bounded, continuous function f .

- (b) When $\nu_n \xrightarrow[n \rightarrow \infty]{w} \nu$, the question often arises whether $\langle \nu_n, f \rangle \xrightarrow[n \rightarrow \infty]{} \langle \nu, f \rangle$ also for some f 's which may be either unbounded, or discontinuous. Examples of such situations are dealt with in Exercises 5.4 and 5.9.
- (c) Recall Scheffe's lemma: if (X_n) and X are \mathbb{R}_+ -valued r.v.s, with $X_n \xrightarrow[n \rightarrow \infty]{(P)} X$, and $E[X_n] \xrightarrow[n \rightarrow \infty]{} E[X]$, then $X_n \xrightarrow[n \rightarrow \infty]{} X$ in $L^1(P)$, hence the X_n 's are uniformly integrable, thus providing a partial converse to the statement in this exercise.

* 1.5 Conditional expectation and the Monotone Class Theorem

Consider, on a probability space (Ω, \mathcal{F}, P) , a sub- σ -field \mathcal{G} . Assume that there exist two r.v.s, X and Y , with X \mathcal{F} -measurable and Y \mathcal{G} -measurable such that, for every Borel bounded function $g : \mathbb{R} \rightarrow \mathbb{R}_+$, one has:

$$E[g(X) \mid \mathcal{G}] = g(Y) .$$

Prove that: $X = Y$ a.s. *Hint:* Look at the title !

Comments. For a deeper result, see Exercise 1.12.

** 1.6 L^p -convergence of conditional expectations

Let (Ω, \mathcal{F}, P) be a probability space and $X \in L^p(\Omega, \mathcal{F}, P)$, $X \geq 0$, for some $p \geq 1$.

1. Let \mathbb{H} be the set of all sub- σ -fields of \mathcal{F} . Prove that the family of r.v.s $\{(E[X \mid \mathcal{G}]^p) : \mathcal{G} \in \mathbb{H}\}$ is uniformly integrable. (We refer to Exercise 1.2 for the definition of uniform integrability.)
2. Show that if a sequence of r.v.s (Y_n) , with values in \mathbb{R}_+ , is such that (Y_n^p) is uniformly integrable and (Y_n) converges in probability to Y , then (Y_n) converges to Y in L^p .

3. Let (\mathcal{B}_n) be a monotone sequence of sub- σ -fields of \mathcal{F} . We denote by \mathcal{B} the limit of (\mathcal{B}_n) , that is $\mathcal{B} = \vee_n \mathcal{B}_n$ if (\mathcal{B}_n) increases or $\mathcal{B} = \cap_n \mathcal{B}_n$ if (\mathcal{B}_n) decreases. Prove that

$$E(X | \mathcal{B}_n) \xrightarrow{L^p} E(X | \mathcal{B}).$$

Hint. First, prove the result in the case $p = 2$.

Comments and references. These three questions are very classical. We present the end result (of question 3.) as an exercise, although it is an important and classical part of the Martingale Convergence Theorem (see the reference hereafter). We wish to emphasize that here, nonetheless, as for many other questions the L^p convergence results are much easier to obtain than the corresponding almost sure one, which is proved in J. Neveu [62] and D. Williams [98].

* 1.7 Measure preserving transformations

Let (Ω, \mathcal{F}, P) be a probability space, and let $T : (\Omega, \mathcal{F}) \rightarrow (\Omega, \mathcal{F})$ be a transformation which preserves P , i.e. $T(P) = P$.

1. Prove that, if $X : (\Omega, \mathcal{F}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is almost T -invariant, i.e. $X(\omega) = X(T(\omega))$, P a.s., then, for any bounded function $\Phi : (\Omega \times \mathbb{R}, \mathcal{F} \otimes \mathcal{B}(\mathbb{R})) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, one has:

$$E[\Phi(\omega, X(\omega))] = E[\Phi(T(\omega), X(\omega))]. \quad (1.7.1)$$

2. Conversely, prove that, if (1.7.1) is satisfied, then, for every bounded function $g : (\mathbb{R}, \mathcal{B}(\mathbb{R})) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, one has:

$$E[g(X) | T^{-1}(\mathcal{F})] = g(X(T(\omega))), \quad P \text{ a.s.} \quad (1.7.2)$$

3. Prove that (1.7.1) is satisfied if and only if X is almost T -invariant.

Hint: Use Exercise 1.5.

* 1.8 Ergodic transformations

Let (Ω, \mathcal{F}, P) be a probability space, and let $T : (\Omega, \mathcal{F}) \rightarrow (\Omega, \mathcal{F})$ be a transformation which preserves P , i.e. $T(P) = P$.

We denote by \mathcal{J} the invariant σ -field of T , i.e.

$$\mathcal{J} = \{A \in \mathcal{F} : 1_A(T\omega) = 1_A(\omega)\}.$$

T is said to be ergodic if \mathcal{J} is P -trivial.

1. Measure theory and probability

7

1. Prove that T is ergodic if the following property holds:

(a) for every f, g belonging to a vector space \mathcal{H} which is dense in $L^2(\mathcal{F}, P)$,

$$E[f(g \circ T^n)] \xrightarrow{n \rightarrow \infty} E(f)E(g),$$

where T^n is the composition product of T by itself, $(n - 1)$ times: $T^n = T \circ T \circ \dots \circ T$.

2. Prove that, if there exists an increasing sequence $(\mathcal{F}_k)_{k \in \mathbb{N}}$ of sub- σ -fields of \mathcal{F} such that:

(b) $\bigvee_k \mathcal{F}_k = \mathcal{F}$,

(c) for every k , $T^{-1}(\mathcal{F}_k) \subseteq \mathcal{F}_k$,

(d) for every k , $\bigcap_n (T^n)^{-1}(\mathcal{F}_k)$ is P -trivial,

then the property (a) is satisfied.

Consequently, the properties (b)–(c)–(d) imply that T is ergodic.

* 1.9 Invariant σ -fields

Consider, on a probability space (Ω, \mathcal{F}, P) , a measurable transformation T which preserves P , i.e. $T(P) = P$.

Let g be an integrable random variable, i.e. $g \in L^1(\Omega, \mathcal{F}, P)$.

Prove that the two following properties are equivalent:

(i) for every $f \in L^\infty(\Omega, \mathcal{F}, P)$, $E[fg] = E[(f \circ T)g]$,

(ii) g is almost T -invariant, i.e. $g = g \circ T$, P a.s.

Hint: One may use the following form of the ergodic theorem:

$$\frac{1}{n} \sum_{i=1}^n f \circ T^i \xrightarrow[n \rightarrow \infty]{L^1} E[f | \mathcal{J}] ,$$

where \mathcal{J} is the invariant σ -field of T .

Comments and references on Exercises 1.7, 1.8, 1.9:

(a) These are featured at the very beginning of every book on Ergodic Theory. See, for example, K. Petersen [67] and P. Billingsley [8].

- (b) Of course, many examples of ergodic transformations are provided in books on Ergodic Theory. Let us simply mention that if (B_t) denotes Brownian motion, then the scaling operation, $B \mapsto \left(\frac{1}{\sqrt{c}}B_c\right)$ is ergodic for $c \neq 1$. Can you prove this result? Actually, the same result holds for the whole class of stable processes, as proved in Exercise 5.17.
- (c) Exercise 1.12 yields a proof of (i) \Rightarrow (ii) which does not use the Ergodic Theorem.

** 1.10 Extremal solutions of (general) moments problems

Consider, on a measurable space (Ω, \mathcal{F}) , a family $\Phi = (\varphi_i)_{i \in I}$ of real-valued random variables, and let $c = (c_i)_{i \in I}$ be a family of real numbers.

Define $\mathcal{M}_{\Phi, c}$ to be the family of probabilities P on (Ω, \mathcal{F}) such that:

$$(a) \Phi \subset L^1(\Omega, \mathcal{F}, P) ; \quad (b) \text{ for every } i \in I, E_P(\varphi_i) = c_i .$$

A probability measure P in $\mathcal{M}_{\Phi, c}$ is called extremal if whenever $P = \alpha P_1 + (1 - \alpha)P_2$, with $0 < \alpha < 1$ and $P_1, P_2 \in \mathcal{M}_{\Phi, c}$, then $P = P_1 = P_2$.

1. Prove that, if $P \in \mathcal{M}_{\Phi, c}$, then P is extremal in $\mathcal{M}_{\Phi, c}$, if and only if the vector space generated by 1 and Φ is dense in $L^1(\Omega, \mathcal{F}, P)$.
2. (i) Prove that, if P is extremal in $\mathcal{M}_{\Phi, c}$, and $Q \in \mathcal{M}_{\Phi, c}$, such that $Q \ll P$, and $\frac{dQ}{dP}$ is bounded, then $Q = P$.
 (ii) Prove that, if P is not extremal in $\mathcal{M}_{\Phi, c}$, and $Q \in \mathcal{M}_{\Phi, c}$, such that $Q \simeq P$, with $0 < \varepsilon \leq \frac{dQ}{dP} \leq C < \infty$, for some $\varepsilon, C > 0$, then Q is not extremal in $\mathcal{M}_{\Phi, c}$.
3. Let T be a measurable transformation of (Ω, \mathcal{F}) , and define \mathcal{M}_T to be the family of probabilities P on (Ω, \mathcal{F}) which are preserved by T , i.e. $T(P) = P$. Prove that, if $P \in \mathcal{M}_T$, then P is extremal in \mathcal{M}_T if, and only if, T is ergodic under P .

Comments and references:

- (a) The result of question 1 appears to have been obtained independently by:
 M.A. NAIMARK: Extremal spectral functions of a symmetric operator. *Bull. Acad. Sci. URSS. Sér. Math.*, **11**, 327–344 (1947).
 (see e.g. N.I. AKHIEZER: *The Classical Moment Problem and Some Related Questions in Analysis*. Publishing Co., New York, p. 47, 1965), and
 R. DOUGLAS: On extremal measures and subspace density II. *Michigan Math. J.*, **11**, 243–246 (1964). *Proc. Amer. Math. Soc.*, **17**, 1363–1365 (1966).

1. Measure theory and probability

9

It is often used in the study of indeterminate moment problems, see e.g.

CH. BERG: Recent results about moment problems. *Probability Measures on Groups and Related Structures, XI* (Oberwolfach, 1994), 1–13, World Sci. Publishing, River Edge, NJ, 1995.

CH. BERG: Indeterminate moment problems and the theory of entire functions. *Proceedings of the International Conference on Orthogonality, Moment Problems and Continued Fractions* (Delft, 1994). *J. Comput. Appl. Math.*, **65**, no. 1–3, 27–55 (1995).

(b) Some variants are presented in:

E.B. DYNKIN: Sufficient statistics and extreme points. *Ann. Probab.*, **6**, no. 5, 705–730 (1978).

For some applications to martingale representations as stochastic integrals, see:

M. YOR: Sous-espaces denses dans L^1 et H^1 et représentations des martingales, *Séminaire de Probabilités XII*, 264–309, *Lecture Notes in Mathematics*, **649**, Springer, 1978.

(c) The next exercise gives the most classical example of a non-moments determinate probability on \mathbb{R} . It is those particular moments problems which motivated the general statement of Naimark–Douglas.

* 1.11 The log normal distribution is moments indeterminate

Let N_{σ^2} be a centered Gaussian variable with variance σ^2 . Associate to N_{σ^2} the log normal variable:

$$X_{\sigma^2} = \exp(N_{\sigma^2}) .$$

1. Compute the density of X_{σ^2} ; its expression gives an explanation for the term “log normal”.
2. Prove that for every $n \in \mathbb{Z}$, and $p \in \mathbb{Z}$,

$$E \left[X_{\sigma^2}^n \sin \left(\frac{p\pi}{\sigma^2} N_{\sigma^2} \right) \right] = 0 . \quad (1.11.1)$$

3. Show that there exist infinitely many probability laws μ on \mathbb{R}_+ such that:

(i) for every $n \in \mathbb{Z}$,

$$\int \mu(dx) x^n = \exp \left(\frac{n^2 \sigma^2}{2} \right) .$$

(ii) μ has a bounded density with respect to the law of $\exp(N_{\sigma^2})$.

Comments and references:

- (a) This exercise and its proof go back to T. Stieltjes' fundamental memoir:

T.J. STIELTJES: Recherches sur les fractions continues. Reprint of *Ann. Fac. Sci. Toulouse* **9**, (1895), A5–A47. Reprinted in *Ann. Fac. Sci. Toulouse Math.*, **6**, no. 4, A5–A47 (1995).

There are many other examples of elements of \mathcal{M}_{σ^2} , including some with countable support; see, e.g., Stoyanov ([86], p. 104).

- (b) In his memoir, Stieltjes also provides other similar elementary proofs for different moment problems. For instance, for any $a > 0$, if Z_a denotes a gamma variable, then for $c > 2$, the law of $(Z_a)^c$ is not moments determinate. See, e.g., J.M. Stoyanov [86], § 11.4.

- (c) There are some sufficient criteria which bear upon the sequence of moments $m_n = E[X^n]$ of an r.v. X and ensure that the law of X is determined uniquely from the (m_n) sequence. (In particular, the classical *sufficient* Carleman criterion asserts that if $\sum_n (m_{2n})^{-1/2n} = \infty$, then the law of X is moments determinate.) But, these are unsatisfactory in a number of cases, and the search continues. See, for example, the following.

J. STOYANOV: Krein condition in probabilistic moment problems. *Bernoulli*, **6**, no. 5, 939–949 (2000).

A. GUT: On the moment problem. *Bernoulli*, **8**, no. 3, 407–421 (2002).

* 1.12 Conditional expectations and equality in law

Let $X \in L^1(\Omega, \mathcal{F}, P)$, and \mathcal{G} be a sub- σ -field of \mathcal{F} . The objective of this exercise is to prove that if X and $Y \stackrel{(\text{def})}{=} E[X | \mathcal{G}]$ have the same distribution, then X is \mathcal{G} measurable (hence $X = Y$).

1. Prove the result if X belongs to L^2 .
2. Prove that for every $a, b \in \mathbb{R}_+$,

$$E[(X \wedge a) \vee (-b) | \mathcal{G}] = (Y \wedge a) \vee (-b), \quad (1.12.1)$$

and conclude.

3. Prove the result of Exercise 1.5 using the previous question.
4. In Exercise 1.9, prove, without using the Ergodic Theorem that (i) implies (ii).