SECTION 3 EXISTENCE AND EXTENSION

The main theorem to be proved here may be compactly stated this way:

THEOREM 3.1
A probability measure on a field has a unique extension to the generated σ-field.

In more detail the assertion is this: Suppose that P is a probability measure on a field $\mathcal{F}_0$ of subsets of Ω, and put $\mathcal{F} = \sigma(\mathcal{F}_0)$. Then there exists a probability measure Q on $\mathcal{F}$ such that $Q(A) = P(A)$ for $A \in \mathcal{F}_0$. Further, if Q' is another probability measure on $\mathcal{F}$ such that $Q'(A) = P(A)$ for $A \in \mathcal{F}_0$, then $Q'(A) = Q(A)$ for $A \in \mathcal{F}$.

Although the measure extended to $\mathcal{F}$ is usually denoted by the same letter as the original measure on $\mathcal{F}_0$, they are really different set functions, since they have different domains of definition. The class $\mathcal{F}_0$ is only assumed finitely additive in the theorem, but the set function P on it must be assumed countably additive (since this of course follows from the conclusion of the theorem).

As shown in Theorem 2.2, λ (initially defined for intervals as length: $\lambda(I) = |I|$) extends to a probability measure on the field $\mathcal{B}_0$ of finite disjoint unions of subintervals of $(0, 1]$. By Theorem 3.1, λ extends in a unique way from $\mathcal{B}_0$ to $\mathcal{B} = \sigma(\mathcal{B}_0)$, the class of Borel sets in $(0,1]$. The extended λ is *Lebesgue measure* on the unit interval. Theorem 3.1 has many other applications as well.

The uniqueness in Theorem 3.1 will be proved later; see Theorem 3.3. The first project is to prove that an extension does exist.

Construction of the Extension

Let P be a probability measure on a field $\mathcal{F}_0$. The construction following extends P to a class that in general is much larger than $\sigma(\mathcal{F}_0)$ but nonetheless does not in general contain all the subsets of Ω.

For each subset A of Ω, define its *outer measure* by

$$P^*(A) = \inf \sum_n P(A_n), \tag{3.1}$$

where the infimum extends over all finite and infinite sequences $A_1, A_2, \ldots$ of $\mathcal{F}_0$-sets satisfying $A \subset \bigcup_n A_n$. If the A_n form an efficient covering of A, in the sense that they do not overlap one another very much or extend much beyond A, then $\sum_n P(A_n)$ should be a good outer approximation to the measure of A if A is indeed to have a measure assigned it at all. Thus (3.1) represents a first attempt to assign a measure to A.

Because of the rule $P(A^c) = 1 - P(A)$ for complements (see (2.6)), it is natural in approximating A from the inside to approximate the complement A^c

from the outside instead and then subtract from 1:

$$P_*(A) = 1 - P^*(A^c). \tag{3.2}$$

This, the *inner measure* of A, is a second candidate for the measure of A.[†] A plausible procedure is to assign measure to those A for which (3.1) and (3.2) agree, and to take the common value $P^*(A) = P_*(A)$ as the measure. Since (3.1) and (3.2) agree if and only if

$$P^*(A) + P^*(A^c) = 1, \tag{3.3}$$

the procedure would be to consider the class of A satisfying (3.3) and use $P^*(A)$ as the measure.

It turns out to be simpler to impose on A the more stringent requirement that

$$P^*(A \cap E) - P^*(A^c \cap E) = P^*(E) \tag{3.4}$$

hold for every set E; (3.3) is the special case $E = \Omega$, because it will turn out that $P^*(\Omega) = 1$.[‡] A set A is called P^*-*measurable* if (3.4) holds for all E; let $\mathscr{M}$ be the class of such sets. What will be shown is that $\mathscr{M}$ contains $\sigma(\mathscr{F}_0)$ and that the restriction of P^* to $\sigma(\mathscr{F}_0)$ is the required extension of P.

The set function P^* has four properties that will be needed:

(i) $P^*(\emptyset) = 0$;

(ii) P^* is nonnegative: $P^*(A) \geq 0$ for every $A \subset \Omega$;

(iii) P^* is monotone: $A \subset B$ implies $P^*(A) \leq P^*(B)$;

(iv) P^* is countably subadditive: $P^*(\bigcup_n A_n) \leq \sum_n P^*(A_n)$.

The others being obvious, only (iv) needs proof. For a given ϵ, choose $\mathscr{F}_0$-sets B_{nk} such that $A_n \subset \bigcup_k B_{nk}$ and $\sum_k P(B_{nk}) < P^*(A_n) + \epsilon 2^{-n}$, which is possible by the definition (3.1). Now $\bigcup_n A_n \subset \bigcup_{n,k} B_{nk}$, so that $P^*(\bigcup_n A_n) \leq \sum_{n,k} P(B_{nk}) < \sum_n P^*(A_n) + \epsilon$, and (iv) follows.[§] Of course, (iv) implies finite subadditivity.

By definition, A lies in the class $\mathscr{M}$ of P^*-measurable sets if it splits each E in 2^Ω in such a way that P^* adds for the pieces—that is, if (3.4) holds. Because of finite subadditivity, this is equivalent to

$$P^*(A \cap E) + P^*(A^c \cap E) \leq P^*(E). \tag{3.5}$$

[†]An idea which seems reasonable at first is to define $P_*(A)$ as the supremum of the sums $\sum_n P(A_n)$ for disjoint sequences of $\mathscr{F}_0$-sets in A. This will not do. For example, in the case where Ω is the unit interval, $\mathscr{F}_0$ is $\mathscr{B}_0$ (Example 2.2), and P is λ as defined by (2.12), the set N of normal numbers would have inner measure 0 because it contains no nonempty elements of $\mathscr{B}_0$; in a satisfactory theory, N will have both inner and outer measure 1.

[‡]It also turns out after the fact, that (3.3) implies that (3.4) holds for all E anyway; see Problem 3.2.

[§]Compare the proof on p. 10 that a countable union of negligible sets is negligible.

Lemma 1. *The class $\mathscr{M}$ is a field.*

Proof. It is clear that $\Omega \in \mathscr{M}$ and that $\mathscr{M}$ is closed under complementation. Suppose that $A, B \in \mathscr{M}$ and $E \subset \Omega$. Then

$$\begin{aligned}
P^*(E) &= P^*(B \cap E) + P^*(B^c \cap E) \\
&= P^*(A \cap B \cap E) + P^*(A^c \cap B \cap E) \\
&\quad + P^*(A \cap B^c \cap E) + P^*(A^c \cap B^c \cap E) \\
&\geq P^*(A \cap B \cap E) \\
&\quad + P^*((A^c \cap B \cap E) \cup (A \cap B^c \cap E) \cup (A^c \cap B^c \cap E)) \\
&= P^*((A \cap B) \cap E) + P^*((A \cap B)^c \cap E),
\end{aligned}$$

the inequality following by subadditivity. Hence† $A \cap B \in \mathscr{M}$, and $\mathscr{M}$ is a field. ■

Lemma 2. *If $A_1, A_2, \ldots$ is a finite or infinite sequence of disjoint $\mathscr{M}$-sets, then for each $E \subset \Omega$,*

$$P^*\left(E \cap \left(\bigcup_k A_k\right)\right) = \sum_k P^*(E \cap A_k). \tag{3.6}$$

Proof. Consider first the case of finitely many A_k, say n of them. For $n = 1$, there is nothing to prove. In the case $n = 2$, if $A_1 \cup A_2 = \Omega$, then (3.6) is just (3.4) with A_1 (or A_2) in the role of A. If $A_1 \cup A_2$ is smaller than Ω, split $E \cap (A_1 \cup A_2)$ by A_1 and A_1^c (or by A_2 and A_2^c) and use (3.4) and disjointness.

Assume (3.6) holds for the case of $n-1$ sets. By the case $n = 2$, together with the induction hypothesis, $P^*(E \cap (\bigcup_{k=1}^n A_k)) = P^*(E \cap (\bigcup_{k=1}^{n-1} A_k)) + P^*(E \cap A_n) = \sum_{k=1}^n P^*(E \cap A_k)$.

Thus (3.6) holds in the finite case. For the infinite case use monotonicity: $P^*(E \cap (\bigcup_{k=1}^\infty A_k)) \geq P^*(E \cap (\bigcup_{k=1}^n A_k)) = \sum_{k=1}^n P^*(E \cap A_k)$. Let $n \to \infty$, and conclude that the left side of (3.6) is greater than or equal to the right. The reverse inequality follows by countable subadditivity. ■

Lemma 3. *The class $\mathscr{M}$ is a σ-field, and P^* restricted to $\mathscr{M}$ is countably additive*

Proof. Suppose that $A_1, A_2, \ldots$ are disjoint $\mathscr{M}$-sets with union A. Since $F_n = \bigcup_{k=1}^n A_k$ lies in the field $\mathscr{M}$, $P^*(E) = P^*(E \cap F_n) + P^*(E \cap F_n^c)$. To the first term on the right apply (3.6), and to the second term apply monotonicity ($F_n^c \supset A^c$): $P^*(E) \geq \sum_{k=1}^n P^*(E \cap A_k) + P^*(E \cap A^c)$. Let $n \to \infty$ and use

---

†This proof does not work if (3.4) is weakened to (3.3).

(3.6) again: $P^*(E) \geq \sum_{k=1}^{\infty} P^*(E \cap A_k) + P^*(E \cap A^c) = P^*(E \cap A) + P^*(E \cap A^c)$. Hence A satisfies (3.5) and so lies in $\mathscr{M}$, which is therefore closed under the formation of countable disjoint unions.

From the fact that $\mathscr{M}$ is a field closed under the formation of countable disjoint unions it follows that $\mathscr{M}$ is a σ-field (for sets B_k in $\mathscr{M}$, let $A_1 = B_1$ and $A_k = B_k \cap B_1^c \cap \cdots \cap B_{k-1}^c$; then the A_k are disjoint $\mathscr{M}$-sets and $\bigcup_k B_k = \bigcup_k A_k \in \mathscr{M}$). The countable additivity of P^* on $\mathscr{M}$ follows from (3.6): take $E = \Omega$. ■

Lemmas 1, 2, and 3 use only the properties (i) through (iv) of P^* derived above. The next two use the specific assumption that P^* is defined via (3.1) from a probability measure P on the field $\mathscr{F}_0$.

Lemma 4. *If P^* is defined by* (3.1), *then $\mathscr{F}_0 \subset \mathscr{M}$.*

Proof. Suppose that $A \in \mathscr{F}_0$. Given E and ϵ, choose $\mathscr{F}_0$-sets A_n such that $E \subset \bigcup_n A_n$ and $\sum_n P(A_n) \leq P^*(E) + \epsilon$. The sets $B_n = A_n \cap A$ and $C_n = A_n \cap A^c$ lie in $\mathscr{F}_0$ because it is a field. Also, $E \cap A \subset \bigcup_n B_n$ and $E \cap A^c \subset \bigcup_n C_n$; by the definition of P^* and the finite additivity of $P, P^*(E \cap A) + P^*(E \cap A^c) \leq \sum_n P(B_n) + \sum_n P(C_n) = \sum_n P(A_n) \leq P^*(E) + \epsilon$. Hence $A \in \mathscr{F}_0$ implies (3.5), and so $\mathscr{F}_0 \subset \mathscr{M}$. ■

Lemma 5. *If P^* is defined by (3.1), then*

$$P^*(A) = P(A) \quad \text{for } A \in \mathscr{F}_0. \tag{3.7}$$

Proof. It is obvious from the definition (3.1) that $P^*(A) \leq P(A)$ for A in $\mathscr{F}_0$. If $A \subset \bigcup_n A_n$, where A and the A_n are in $\mathscr{F}_0$, then by the countable subadditivity and monotonicity of P on $\mathscr{F}_0$. $P(A) \leq \sum_n P(A \cap A_n) \leq \sum_n P(A_n)$. Hence (3.7). ■

Proof of Extension in Theorem 3.1. Suppose that P^* is defined via (3.1) from a (countably additive) probability measure P on the field $\mathscr{F}_0$. Let $\mathscr{F} = \sigma(\mathscr{F}_0)$. By Lemmas 3 and 4,†

$$\mathscr{F}_0 \subset \mathscr{F} \subset \mathscr{M} \subset 2^{\Omega}.$$

By (3.7), $P^*(\Omega) = P(\Omega) = 1$. By Lemma 3, P^* (which is defined on all of 2^{Ω}) restricted to $\mathscr{M}$ is therefore a probability measure there. And then P^* further restricted to $\mathscr{F}$ is clearly a probability measure on that class as well. This measure on $\mathscr{F}$ is the required extension, because by (3.7) it agrees with P on $\mathscr{F}_0$. ■

†In the case of Lebesgue measure, the relation is $\mathscr{B}_0 \subset \mathscr{B} \subset \mathscr{M} \subset 2^{(0,1]}$, and each of the three inclusions is strict; see Example 2.2 and Problems 3.14 and 3.21.

Uniqueness and the π–λ Theorem

To prove the extension in Theorem 3.1 is unique requires some auxiliary concepts. A class $\mathcal{P}$ of subsets of Ω is a *π-system* if it is closed under the formation of finite intersections:

(π) $A, B \in \mathcal{P}$ implies $A \cap B \in \mathcal{P}$.

A class $\mathcal{L}$ is a *λ-system* if it contains Ω and is closed under the formation of complements and of finite and countable *disjoint* unions:

(λ_1) $\Omega \in \mathcal{L}$;
(λ_2) $A \in \mathcal{L}$ implies $A^c \in \mathcal{L}$;
(λ_3) $A_1, A_2, \ldots, \in \mathcal{L}$ and $A_n \cap A_m = \emptyset$ for $m \neq n$ imply $\bigcup_n A_n \in \mathcal{L}$.

Because of the disjointness condition in (λ_3), the definition of λ-system is weaker (more inclusive) than that of σ-field. In the presence of (λ_1) and (λ_2), which imply $\emptyset \in \mathcal{L}$, the countably infinite case of (λ_3) implies the finite one.

In the presence of (λ_1) and (λ_3), (λ_2) is equivalent to the condition that $\mathcal{L}$ is closed under the formation of proper differences:

(λ_2') $A, B \in \mathcal{L}$ and $A \subset B$ imply $B - A \in \mathcal{L}$.

Suppose, in fact, that $\mathcal{L}$ satisfies (λ_2) and (λ_3). If $A, B \in \mathcal{L}$ and $A \subset B$, then $\mathcal{L}$ contains B^c, the disjoint union $A \cup B^c$, and its complement $(A \cup B^c)^c = B - A$. Hence (λ_2'). On the other hand, if $\mathcal{L}$ satisfies (λ_1) and (λ_2'), then $A \in \mathcal{L}$ implies $A^c = \Omega - A \in \mathcal{L}$. Hence ($\lambda_2$).

Although a σ-field is a λ-system, the reverse is not true (in a four-point space take $\mathcal{L}$ to consist of $\emptyset, \Omega$, and the six two-point sets). But the connection is close:

Lemma 6. *A class that is both a π-system and a λ-system is a σ-field.*

Proof. The class contains Ω by (λ_1) and is closed under the formation of complements and finite intersections by (λ_2) and (π). It is therefore a field. It is a σ-field because if it contains sets A_n, then it also contains the disjoint sets $B_n = A_n \cap A_1^c \cap \cdots \cap A_{n-1}^c$ and by (λ_3) contains $\bigcup_n A_n = \bigcup_n B_n$. ■

Many uniqueness arguments depend on *Dynkin's π–λ theorem*:

THEOREM 3.2
If $\mathcal{P}$ is a π-system and $\mathcal{L}$ is a λ-system, then $\mathcal{P} \subset \mathcal{L}$ implies $\sigma(\mathcal{P}) \subset \mathcal{L}$.

Proof. Let $\mathscr{L}_0$ be the λ-system generated by $\mathscr{P}$—that is, the intersection of all λ-systems containing $\mathscr{P}$. It is a λ-system, it contains $\mathscr{P}$, and it is contained in every λ-system that contains $\mathscr{P}$ (see the construction of generated σ-fields, p. 21). Thus $\mathscr{P} \subset \mathscr{L}_0 \subset \mathscr{L}$. If it can be shown that $\mathscr{L}_0$ is also a π-system, then it will follow by Lemma 6 that it is a σ-field. From the minimality of $\sigma(\mathscr{P})$ it will then follow that $\sigma(\mathscr{P}) \subset \mathscr{L}_0$, so that $\mathscr{P} \subset \sigma(\mathscr{P}) \subset \mathscr{L}_0 \subset \mathscr{L}$. Therefore, it suffices to show that $\mathscr{L}_0$ is a π-system.

For each A, let $\mathscr{L}_A$ be the class of sets B such that $A \cap B \in \mathscr{L}_0$. If A is assumed to lie in $\mathscr{P}$, or even if A is merely assumed to lie in $\mathscr{L}_0$, then $\mathscr{L}_A$ *is a* λ*-system*: Since $A \cap \Omega = A \in \mathscr{L}_0$ by the assumption, $\mathscr{L}_A$ satisfies (λ_1). If $B_1, B_2 \in \mathscr{L}_A$ and $B_1 \subset B_2$, then the λ-system $\mathscr{L}_0$ contains $A \cap B_1$ and $A \cap B_2$ and hence contains the proper difference $(A \cap B_2) - (A \cap B_1) = A \cap (B_2 - B_1)$, so that $\mathscr{L}_A$ contains $B_2 - B_1$: $\mathscr{L}_A$ satisfies (λ_2'). If B_n are disjoint $\mathscr{L}_A$-sets, then $\mathscr{L}_0$ contains the disjoint sets $A \cap B_n$ and hence contains their union $A \cap (\bigcup_n B_n)$: $\mathscr{L}_A$ satisfies (λ_3).

If $A \in \mathscr{P}$ and $B \in \mathscr{P}$, then ($\mathscr{P}$ is a π-system) $A \cap B \in \mathscr{P} \subset \mathscr{L}_0$, or $B \in \mathscr{L}_A$. Thus $A \in \mathscr{P}$ implies $\mathscr{P} \subset \mathscr{L}_A$, and since $\mathscr{L}_A$ is a λ-system, minimality gives $\mathscr{L}_0 \subset \mathscr{L}_A$.

Thus $A \in \mathscr{P}$ implies $\mathscr{L}_0 \subset \mathscr{L}_A$, or, to put it another way, $A \in \mathscr{P}$ and $B \in \mathscr{L}_0$ together imply that $B \in \mathscr{L}_A$ and hence $A \in \mathscr{L}_B$. (The key to the proof is that $B \in \mathscr{L}_A$ if and only if $A \in \mathscr{L}_B$.) This last implication means that $B \in \mathscr{L}_0$ implies $\mathscr{P} \subset \mathscr{L}_B$. Since $\mathscr{L}_B$ is a λ-system, it follows by minimality once again that $B \in \mathscr{L}_0$ implies $\mathscr{L}_0 \subset \mathscr{L}_B$. Finally, $B \in \mathscr{L}_0$ and $C \in \mathscr{L}_0$ together imply $C \in \mathscr{L}_B$, or $B \cap C \in \mathscr{L}_0$. Therefore, $\mathscr{L}_0$ is indeed a π-system. ■

Since a field is certainly a π-system, the uniqueness asserted in Theorem 3.1 is a consequence of this result:

THEOREM 3.3

Suppose that P_1 and P_2 are probability measures on $\sigma(\mathscr{P})$, where $\mathscr{P}$ is a π-system. If P_1 and P_2 agree on $\mathscr{P}$, then they agree on $\sigma(\mathscr{P})$.

Proof. Let $\mathscr{L}$ be the class of sets A in $\sigma(\mathscr{P})$ such that $P_1(A) = P_2(A)$. Clearly $\Omega \in \mathscr{L}$. If $A \in \mathscr{L}$, then $P_1(A^c) = 1 - P_1(A) = 1 - P_2(A) = P_2(A^c)$, and hence $A^c \in \mathscr{L}$. If A_n are disjoint sets in $\mathscr{L}$, then $P_1(\bigcup_n A_n) = \Sigma_n P_1(A_n) = \Sigma_n P_2(A_n) = P_2(\bigcup_n A_n)$, and hence $\bigcup_n A_n \in \mathscr{L}$. Therefore $\mathscr{L}$ is a λ-system. Since by hypothesis $\mathscr{P} \subset \mathscr{L}$ and $\mathscr{P}$ is a π-system, the $\pi - \lambda$ theorem gives $\sigma(\mathscr{P}) \subset \mathscr{L}$, as required. ■

Note that the π-λ theorem and the concept of λ-system are exactly what are needed to make this proof work: The essential property of probability measures is countable additivity, and this is a condition on countable *disjoint* unions, the only kind involved in the requirement (λ_3) in the definition of λ-system.

In this, as in many applications of the π-λ theorem, $\mathscr{L} \subset \sigma(\mathscr{P})$ and therefore $\sigma(\mathscr{P}) = \mathscr{L}$, even though the relation $\sigma(\mathscr{P}) \subset \mathscr{L}$ itself suffices for the conclusion of the theorem.

Monotone Classes

A class $\mathscr{M}$ of subsets of Ω is *monotone* if it is closed under the formation of monotone unions and intersections:

(i) $A_1, A_2, \ldots \in \mathscr{M}$ and $A_n \uparrow A$ imply $A \in \mathscr{M}$;

(ii) $A_1, A_2, \ldots \in \mathscr{M}$ and $A_n \downarrow A$ imply $A \in \mathscr{M}$.

Halmos's monotone class theorem is a close relative of the π-λ theorem but will be less frequently used in this book.

THEOREM 3.4
If $\mathscr{F}_0$ is a field and $\mathscr{M}$ is a monotone class, then $\mathscr{F}_0 \subset \mathscr{M}$ implies $\sigma(\mathscr{F}_0) \subset \mathscr{M}$.

Proof. Let $m(\mathscr{F}_0)$ be the minimal monotone class over $\mathscr{F}_0$—the intersection of all monotone classes containing $\mathscr{F}_0$. It is enough to prove $\sigma(\mathscr{F}_0) \subset m(\mathscr{F}_0)$; this will follow if $m(\mathscr{F}_0)$ is shown to be a field, because a monotone field is a σ-field.

Consider the class $\mathscr{G} = [A: A^c \in m(\mathscr{F}_0)]$. Since $m(\mathscr{F}_0)$ is monotone, so is $\mathscr{G}$. Since $\mathscr{F}_0$ is a field, $\mathscr{F}_0 \subset \mathscr{G}$, and so $m(\mathscr{F}_0) \subset \mathscr{G}$. Hence $m(\mathscr{F}_0)$ is closed under complementation.

Define $\mathscr{G}_1$ as the class of A such that $A \cup B \in m(\mathscr{F}_0)$ for all $B \in \mathscr{F}_0$. Then $\mathscr{G}_1$ is a monotone class and $\mathscr{F}_0 \subset \mathscr{G}_1$; from the minimality of $m(\mathscr{F}_0)$ follows $m(\mathscr{F}_0) \subset \mathscr{G}_1$. Define $\mathscr{G}_2$ as the class of B such that $A \cup B \in m(\mathscr{F}_0)$ for all $A \in m(\mathscr{F}_0)$. Then $\mathscr{G}_2$ is a monotone class. Now from $m(\mathscr{F}_0) \subset \mathscr{G}_1$ it follows that $A \in m(\mathscr{F}_0)$ and $B \in \mathscr{F}_0$ together imply that $A \cup B \in m(\mathscr{F}_0)$; in other words, $B \in \mathscr{F}_0$ implies that $B \in \mathscr{G}_2$. Thus $\mathscr{F}_0 \subset \mathscr{G}_2$; by minimality, $m(\mathscr{F}_0) \subset \mathscr{G}_2$, and hence $A, B \in m(\mathscr{F}_0)$ implies that $A \cup B \in m(\mathscr{F}_0)$. ■

Lebesgue Measure on the Unit Interval

Consider once again the unit interval (0, 1] together with the field $\mathscr{B}_0$ of finite disjoint unions of subintervals (Example 2.2) and the σ-field $\mathscr{B} = \sigma(\mathscr{B}_0)$ of Borel sets in (0, 1]. According to Theorem 2.2, (2.12) defines a probability measure λ on $\mathscr{B}_0$. By Theorem 3.1, λ extends to $\mathscr{B}$, the extended λ being Lebesgue measure. The probability space $((0, 1], \mathscr{B}, \lambda)$ will be the basis for much of the probability theory in the remaining sections of this chapter. A few geometric properties of λ will be considered here. Since the intervals in (0, 1] from a π-system generating $\mathscr{B}, \lambda$ is the only probability measure on $\mathscr{B}$ that assigns to each interval its length as its measure.

Some Borel sets are difficult to visualize:

EXAMPLE 3.1

Let $\{r_1, r_2, \ldots\}$ be an enumeration of the rationals in (0, 1). Suppose that ϵ is small, and choose an open interval $I_n = (a_n, b_n)$ such that $r_n \in I_n \subset (0, 1)$ and $\lambda(I_n) = b_n - a_n < \epsilon 2^{-n}$. Put $A = \bigcup_{n=1}^{\infty} I_n$. By subadditivity, $0 < \lambda(A) < \epsilon$.

Since A contains all the rationals in (0, 1), it is dense there. Thus A is an open, dense set with measure near 0. If I is an open subinterval of (0, 1), then I must intersect one of the I_n, and therefore $\lambda(A \cap I) > 0$.

If $B = (0, 1) - A$ then $1 - \epsilon < \lambda(B) < 1$. The set B contains no interval and is in fact nowhere dense [A15]. Despite this, B has measure nearly 1.

EXAMPLE 3.2

There is a set defined in probability terms that has geometric properties similar to those in the preceding example. As in Section 1, let $d_n(\omega)$ be the nth digit in the dyadic expansion of ω; see (1.7). Let $A_n = [\omega \in (0, 1]: d_i(\omega) = d_{n+i}(\omega) = d_{2n+i}(\omega), i = 1, \ldots, n]$, and let $A = \bigcup_{n=1}^{\infty} A_n$. Probabilistically, A corresponds to the event that in an infinite sequence of tosses of a coin, some finite initial segment is immediately duplicated twice over. From $\lambda(A_n) = 2^n \cdot 2^{-3n}$ it follows that $0 < \lambda(A) \leq \Sigma_{n=1}^{\infty} 2^{-2n} = \frac{1}{3}$. Again A is dense in the unit interval; its measure, less than $\frac{1}{3}$, could be made less than ϵ by requiring that some initial segment be immediately duplicated k times over with k large.

The outer measure (3.1) corresponding to λ on $\mathscr{B}_0$ is the infimum of the sums $\Sigma_n \lambda(A_n)$ for which $A_n \in \mathscr{B}_0$ and $A \subset \bigcup_n A_n$. Since each A_n is a finite disjoint union of intervals, this outer measure is

$$\lambda^*(A) = \inf \sum_n |I_n|, \tag{3.8}$$

where the infimum extends over coverings of A by intervals I_n. The notion of negligibility in Section 1 can therefore be reformulated: A is negligible if and only if $\lambda^*(A) = 0$. For A in $\mathscr{B}$, this is the same thing as $\lambda(A) = 0$. This covers the set N of normal numbers: Since the complement N^c is negligible and lies in $\mathscr{B}, \lambda(N^c) = 0$. Therefore, the Borel set N itself has probability 1: $\lambda(N) = 1$.

Completeness

This is the natural place to consider completeness, although it enters into probability theory in an essential way only in connection with the study of stochastic processes in continuous time; see Sections 37 and 38.

A probability measure space $(\Omega, \mathcal{F}, P)$ is *complete* if $A \subset B, B \in \mathcal{F}$, and $P(B) = 0$ together imply that $A \in \mathcal{F}$ (and hence that $P(A) = 0$). If $(\Omega, \mathcal{F}, P)$ is complete, then the conditions $A \in \mathcal{F}, A \Delta A' \subset B \in \mathcal{F}$, and $P(B) = 0$ together imply that $A' \in \mathcal{F}$ and $P(A') = P(A)$.

Suppose that $(\Omega, \mathcal{F}, P)$ is an arbitrary probability space. Define P^* by (3.1) for $\mathcal{F}_0 = \mathcal{F} = \sigma(\mathcal{F}_0)$, and consider the σ-field $\mathcal{M}$ of P^*-measurable sets. The arguments leading to Theorem 3.1 show that P^* restricted to $\mathcal{M}$ is a probability measure. If $P^*(B) = 0$ and $A \subset B$, then $P^*(A \cap E) + P^*(A^c \cap E) \le P^*(B) + P^*(E) = P^*(E)$ by monotonicity, so that A satisfies (3.5) and hence lies in $\mathcal{M}$. Thus $(\Omega, \mathcal{M}, P^*)$ is a complete probability measure space. *In any probability space it is therefore possible to enlarge the σ-field and extend the measure in such a way as to get a complete space.*

Suppose that $((0, 1], \mathcal{B}, \lambda)$ is completed in this way. The sets in the completed σ-field $\mathcal{M}$ are called *Lebesgue* sets, and λ extended to $\mathcal{M}$ is still called Lebesgue measure.

Nonmeasurable Sets

There exist in $(0, 1]$ *sets that lie outside* $\mathcal{B}$. For the construction (due to Vitali) it is convenient to use addition modulo 1 in $(0, 1]$. For $x, y \in (0, 1]$ take $x \oplus y$ to be $x+y$ or $x + y - 1$ according as $x+y$ lies in $(0, 1]$ or not.[†] Put $A \oplus x = [a \oplus x \colon a \in A]$.

Let $\mathcal{L}$ be the class of Borel sets A such that $A \oplus x$ is a Borel set and $\lambda(A \oplus x) = \lambda(A)$. Then $\mathcal{L}$ is a λ-system containing the intervals, and so $\mathcal{B} \subset \mathcal{L}$ by the π-λ theorem. Thus $A \in \mathcal{B}$ implies that $A \oplus x \in \mathcal{B}$ and $\lambda(A \oplus x) = \lambda(A)$. In this sense, λ is translation-invariant.

Define x and y to be equivalent $(x \sim y)$ if $x \oplus r = y$ for some rational r in $(0, 1]$. Let H be a subset of $(0, 1]$ consisting of exactly one representative point from each equivalence class; such a set exists under the assumption of the axiom of choice [A8]. Consider now the countably many sets $H \oplus r$ for rational r.

These sets are disjoint, because no two distinct points of H are equivalent. (If $H \oplus r_1$ and $H \oplus r_2$ share the point $h_1 \oplus r_1 = h_2 \oplus r_2$, then $h_1 \sim h_2$; this is impossible unless $h_1 = h_2$, in which case $r_1 = r_2$.) Each point of $(0, 1]$ lies in one of these sets, because H has a representative from each equivalence class. (If $x \sim h \in H$, then $x = h \oplus r \in H \oplus r$ for some rational r.) Thus $(0, 1] = \bigcup_r (H \oplus r)$, a countable disjoint union.

If H were in $\mathcal{B}$, it would follow that $\lambda(0, 1] = \Sigma_r \lambda(H \oplus r)$. This is impossible: If the value common to the $\lambda(H \oplus r)$ is 0, it leads to $1 = 0$; if the common

[†] This amounts to working in the circle group, where the translation $y \to x \oplus y$ becomes a rotation (1 is the identity). The rationals form a subgroup, and the set H defined below contains one element from each coset.

value is positive, it leads to a convergent infinite series of identical positive terms ($a + a + \cdots < \infty$ and $a > 0$). Thus H lies outside $\mathscr{B}$.

Two Impossibility Theorems†

The argument above, which uses the axiom of choice, in fact proves this: *There exists on* $2^{(0,1]}$ *no probability measure* P *such that* $P(A \oplus x) = P(A)$ *for all* $A \in 2^{(0,1]}$ *and all* $x \in (0, 1]$. In particular it is impossible to extend λ to a translation-invariant probability measure on $2^{(0,1]}$.

There is a stronger result: *There exists on* $2^{(0,1]}$ *no probability measure* P *such that* $P\{x\} = 0$ *for each* x. Since $\lambda\{x\} = 0$, this implies that it is impossible to extend λ to $2^{(0,1]}$ at all.‡

The proof of this second impossibility theorem requires the well-ordering principle (equivalent to the axiom of choice) and also the continuum hypothesis. Let S be the set of sequences $(s(1), s(2), \ldots)$ of positive integers. Then S has the power of the continuum. (Let the nth partial sum of a sequence in S be the position of the nth 1 in the nonterminating dyadic representation of a point in (0, 1]; this gives a one-to-one correspondence.) By the continuum hypothesis, the elements of S can be put in a one-to-one correspondence with the set of ordinals preceding the first uncountable ordinal. Carrying the well ordering of these ordinals over to S by means of the correspondence gives to S a well-ordering relation $\le_w$ with the property that each element has only countably many predecessors.

For s, t in S write $s \le t$ if $s(i) \le t(i)$ for all $i \ge 1$. Say that t *rejects* s if $t <_w s$ and $s \le t$; this is a transitive relation. Let T be the set of unrejected elements of S. Let V_s be the set of elements that reject s, and assume it is nonempty. If t is the first element (with respect to $\le_w$) of V_s, then $t \in T$ (if t' rejects t, then it also rejects s, and since $t' <_w t$, there is a contradiction). Therefore, if s is rejected at all, it is rejected by an element of T.

Suppose T is countable and let $t_1, t_2, \ldots$ be an enumeration of its elements. If $t^*(k) = t_k(k) + 1$, then t^* is not rejected by any t_k and hence lies in T, which is impossible because it is distinct from each t_k. Thus T is uncountable and must by the continuum hypothesis have the power of (0, 1].

Let x be a one-to-one map of T onto (0, 1]; write the image of t as x_t. Let $A_k^i = [x_t : t(i) = k]$ be the image under x of the set of t in T for which $t(i) = k$. Since $t(i)$ must have some value k, $\bigcup_{k=1}^{\infty} A_k^i = (0, 1]$. Assume that P is countably additive and choose u in S in such a way that $P(\bigcup_{k=1}^{u(i)} A_k^i) \ge 1 - 1/2^{i+1}$ for

†This topic may be omitted. It uses more set theory than is assumed in the rest of the book.

‡This refers to a countably additive extension, of course. If one is content with finite additivity, there *is* an extension to $2^{(0,1]}$; see Problem 3.8.

$i \geq 1$. If

$$A = \bigcap_{i=1}^{\infty} \bigcup_{k=1}^{u(i)} A_k^i = \bigcap_{i=1}^{\infty} [x_t \colon t(i) \leq u(i)] = [x_t \colon t \leq u],$$

then $P(A) > 0$. It A is shown to be countable, this will contradict the hypothesis that each singleton has probability 0.

Now, there is some t_0 in T such that $u \leq t_0$ (if $u \in T$, take $t_0 = u$; otherwise, u is rejected by some t_0 in T). If $t \leq u$ for a t in T, then $t \leq t_0$ and hence $t \leq_w t_0$ (since otherwise t_0 rejects t). This means that $[t \colon t \leq u]$ is contained in the countable set $[t \colon t \leq_w t_0]$, and A is indeed countable.

PROBLEMS

3.1. **(a)** In the proof of Theorem 3.1 the assumed finite additivity of P is used twice and the assumed countable additivity of P is used once. Where?

(b) Show by example that a finitely additive probability measure on a field may not be countably subadditive. Show in fact that if a finitely additive probability measure is countably subadditive, then it is necessarily countably additive as well.

(c) Suppose Theorem 2.1 were weakened by strengthening its hypothesis to the assumption that $\mathscr{F}$ is a σ-field. Why would this weakened result not suffice for the proof of Theorem 3.1?

3.2. Let P be a probability measure on a field $\mathscr{F}_0$ and for every subset A of Ω define $P^*(A)$ by (3.1). Denote also by P the extension (Theorem 3.1) of P to $\mathscr{F} = \sigma(\mathscr{F}_0)$.

(a) Show that

$$P^*(A) = \inf[P(B) \colon A \subset B, B \in \mathscr{F}] \tag{3.9}$$

and (see (3.2))

$$P_*(A) = \sup[P(C) \colon C \subset A, C \in \mathscr{F}], \tag{3.10}$$

and show that the infimum and supremum are always achieved.

(b) Show that A is P^*-measurable if and only if $P_*(A) = P^*(A)$.

(c) The outer and inner measures associated with a probability measure P on a σ-field $\mathscr{F}$ are usually *defined* by (3.9) and (3.10). Show that (3.9) and (3.10) are the same as (3.1) and (3.2) with $\mathscr{F}$ in the role of $\mathscr{F}_0$.