2

Probability Spaces

This chapter discusses the basic properties of probability spaces, and in particular, probability measures. It also introduces the important ideas of set induction.

2.1 Basic Definitions and Properties

A *probability space* is a triple $(\Omega, \mathcal{B}, P)$ where

- Ω is the sample space corresponding to outcomes of some (perhaps hypothetical) experiment.
- $\mathcal{B}$ is the σ-algebra of subsets of Ω. These subsets are called events.
- P is a probability measure; that is, P is a function with domain $\mathcal{B}$ and range $[0, 1]$ such that
  - (i) $P(A) \geq 0$ for all $A \in \mathcal{B}$.
  - (ii) P is σ-additive: If $\{A_n, n \geq 1\}$ are events in $\mathcal{B}$ that are disjoint, then
$$P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n).$$
  - (iii) $P(\Omega) = 1$.

Here are some simple **consequences** of the definition of a probability measure P.

S.I. Resnick, *A Probability Path*, Modern Birkhäuser Classics,

DOI 10.1007/978-0-8176-8409-9_2,

1. We have
$$P(A^c) = 1 - P(A)$$
since from (iii)
$$1 = P(\Omega) = P(A \cup A^c) = P(A) + P(A^c),$$
the last step following from (ii).

2. We have
$$P(\emptyset) = 0$$
since $P(\emptyset) = P(\Omega^c) = 1 - P(\Omega) = 1 - 1$.

3. For events A, B we have
$$P(A \cup B) = PA + PB - P(AB). \tag{2.1}$$
To see this note
$$\begin{aligned} P(A) =& P(AB^c) + P(AB) \\ P(B) =& P(BA^c) + P(AB) \end{aligned}$$
and therefore
$$\begin{aligned} P(A \cup B) =& P(AB^c \cup BA^c \cup AB) \\ =& P(AB^c) + P(BA^c) + P(AB) \\ =& P(A) - P(AB) + P(B) - P(AB) + P(AB) \\ =& P(A) + P(B) - P(AB). \end{aligned}$$

4. The *inclusion–exclusion formula*: If $A_1, \dots, A_n$ are events, then
$$\begin{aligned} P(\bigcup_{j=1}^{n} A_j) \quad = \quad & \sum_{j=1}^{n} P(A_j) - \sum_{1 \le i < j \le n} P(A_i A_j) \\ & + \sum_{1 \le i < j < k \le n} P(A_i A_j A_k) - \cdots \\ & (-1)^{n+1} P(A_1 \cdots A_n). \end{aligned} \tag{2.2}$$
We may prove (2.2) by induction using (2.1) for $n = 2$. The terms on the right side of (2.2) alternate in sign and give inequalities called Bonferroni inequalities when we neglect remainders. Here are two examples:
$$\begin{aligned} P\left(\bigcup_{j=1}^{n} A_j\right) &\le \sum_{j=1}^{n} PA_j \\ P\left(\bigcup_{j=1}^{n} A_j\right) &\ge \sum_{j=1}^{n} PA_j - \sum_{1 \le i < j \le n} P(A_i A_j). \end{aligned}$$

5. *The monotonicity property*: The measure P is non-decreasing: For events A, B

$$\text{If } A \subset B \text{ then } P(A) \le P(B),$$

since

$$P(B) = P(A) + P(B \setminus A) \ge P(A).$$

6. *Subadditivity*: The measure P is σ-subadditive: For events $A_n, n \ge 1$,

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) \le \sum_{n=1}^{\infty} P(A_n).$$

To verify this we write

$$\bigcup_{n=1}^{\infty} A_n = A_1 + A_1^c A_2 + A_3 A_1^c A_2^c + \cdots ,$$

and since P is σ-additive,

$$\begin{aligned} P(\bigcup_{n=1}^{\infty} A_n) =& P(A_1) + P(A_1^c A_2) + P(A_3 A_1^c A_2^c) + \cdots \\ \le& P(A_1) + P(A_2) + P(A_3) + \cdots \end{aligned}$$

by the non-decreasing property of P.

7. *Continuity*: The measure P is continuous for monotone sequences in the sense that

   (i) If $A_n \uparrow A$, where $A_n \in \mathcal{B}$, then $P(A_n) \uparrow P(A)$.

   (ii) If $A_n \downarrow A$, where $A_n \in \mathcal{B}$, then $P(A_n) \downarrow P(A)$.

   To **prove** (i), assume

$$A_1 \subset A_2 \subset A_3 \subset \cdots \subset A_n \subset \cdots$$

and define

$$B_1 = A_1, B_2 = A_2 \setminus A_1, \dots , B_n = A_n \setminus A_{n-1}, \dots .$$

Then $\{B_i\}$ is a disjoint sequence of events and

$$\bigcup_{i=1}^{n} B_i = A_n, \quad \bigcup_{i=1}^{\infty} B_i = \bigcup_{i} A_i = A.$$

By σ-additivity

$$\begin{aligned}P(A) =&P(\bigcup_{i=1}^{\infty} B_i) = \sum_{i=1}^{\infty} P(B_i) = \lim_{n\to\infty} \uparrow \sum_{i=1}^{n} P(B_i)\\ &= \lim_{n\to\infty} \uparrow P(\bigcup_{i=1}^{n} B_i) = \lim_{n\to\infty} \uparrow P(A_n).\end{aligned}$$

To prove (ii), note if $A_n \downarrow A$, then $A_n^c \uparrow A^c$ and by part (i)

$$P(A_n^c) = 1 - P(A_n) \uparrow P(A^c) = 1 - P(A)$$

so that $PA_n \downarrow PA$. □

8. *More continuity and Fatou's lemma*: Suppose $A_n \in \mathcal{B}$, for $n \geq 1$.

   (i) Fatou Lemma: We have the following inequalities

$$\begin{aligned}P(\liminf_{n\to\infty} A_n) &\leq \liminf_{n\to\infty} P(A_n)\\ &\leq \limsup_{n\to\infty} P(A_n) \leq P(\limsup_{n\to\infty} A_n).\end{aligned}$$

   (ii) If $A_n \to A$, then $P(A_n) \to P(A)$.

Proof of 8. (ii) follows from (i) since, if $A_n \to A$, then

$$\limsup_{n\to\infty} A_n = \liminf_{n\to\infty} A_n = A.$$

Suppose (i) is true. Then we get

$$\begin{aligned}P(A) &= P(\liminf_{n\to\infty} A_n) \leq \liminf_{n\to\infty} P(A_n)\\ &\leq \limsup_{n\to\infty} P(A_n) \leq P(\limsup_{n\to\infty} A_n) = P(A),\end{aligned}$$

so equality pertains throughout.

Now consider the proof of (i): We have

$$\begin{aligned}P(\liminf_{n\to\infty} A_n) =&P(\lim_{n\to\infty} \uparrow (\bigcap_{k\geq n} A_k))\\ &= \lim_{n\to\infty} \uparrow P(\bigcap_{k\geq n} A_k)\end{aligned}$$

(from the monotone continuity property 7)

$$\leq \liminf_{n\to\infty} P(A_n)$$

since $P(\cap_{k\geq n} A_k) \leq P(A_n)$. Likewise

$$\begin{aligned} P(\limsup_{n\to\infty} A_n) &= P(\lim_{n\to\infty} \downarrow (\bigcup_{k\geq n} A_k)) \\ &= \lim_{n\to\infty} \downarrow P(\bigcup_{k\geq n} A_k) \end{aligned}$$

(from continuity property 7)

$$\geq \limsup_{n\to\infty} P(A_n),$$

completing the proof. □

Example 2.1.1 Let $\Omega = \mathbb{R}$, and suppose P is a probability measure on $\mathbb{R}$. Define $F(x)$ by

$$F(x) = P((-\infty, x]), \quad x \in \mathbb{R}. \tag{2.3}$$

Then

(i) F is right continuous,

(ii) F is monotone non-decreasing,

(iii) F has limits at $\pm\infty$

$$\begin{aligned} F(\infty) &:= \lim_{x\uparrow\infty} F(x) = 1 \\ F(-\infty) &:= \lim_{x\downarrow-\infty} F(x) = 0. \end{aligned}$$

Definition 2.1.1 A function $F : \mathbb{R} \mapsto [0, 1]$ satisfying (i), (ii), (iii) is called a (probability) distribution function. We abbreviate distribution function by df.

Thus, starting from P, we get F from (2.3). In practice we need to go in the other direction: we start with a known df and wish to construct a probability space $(\Omega, \mathcal{B}, P)$ such that (2.3) holds. See Section 2.5.

Proof of (i), (ii), (iii). For (ii), note that if $x < y$, then

$$(-\infty, x] \subset (-\infty, y]$$

so by monotonicity of P

$$F(x) = P((-\infty, x]) \leq P((-\infty, y]) \leq F(y).$$

Now consider (iii). We have

$$\begin{aligned}
F(\infty) &= \lim_{x_n \uparrow \infty} F(x_n) \qquad \text{(for any sequence } x_n \uparrow \infty) \\
&= \lim_{x_n \uparrow \infty} \uparrow P((-\infty, x_n]) \\
&= P(\lim_{x_n \uparrow \infty} \uparrow (-\infty, x_n]) \qquad \text{(from property 7)} \\
&= P(\bigcup_n (-\infty, x_n]) = P((-\infty, \infty)) \\
&= P(\mathbb{R}) = P(\Omega) = 1.
\end{aligned}$$

Likewise,

$$\begin{aligned}
F(-\infty) &= \lim_{x_n \downarrow -\infty} F(x_n) = \lim_{x_n \downarrow -\infty} \downarrow P((-\infty, x_n]) \\
&= P(\lim_{x_n \downarrow -\infty} (-\infty, x_n]) \qquad \text{(from property 7)} \\
&= P(\bigcap_n (-\infty, x_n]) = P(\emptyset) = 0.
\end{aligned}$$

For the proof of (i), we may show F is right continuous as follows: Let $x_n \downarrow x$. We need to prove $F(x_n) \downarrow F(x)$. This is immediate from the continuity property 7 of P and

$$(-\infty, x_n] \downarrow (-\infty, x]. \qquad \square$$

Example 2.1.2 (Coincidences) The inclusion-exclusion formula (2.2) can be used to compute the probability of a coincidence. Suppose the integers $1, 2, \dots, n$ are randomly permuted. What is the probability that there is an integer left unchanged by the permutation?

To formalize the question, we construct a probability space. Let Ω be the set of all permutations of $1, 2, \dots, n$ so that

$$\Omega = \{(x_1, \dots, x_n) : x_i \in \{1, \dots, n\}; i = 1, \dots, n; x_i \neq x_j\}.$$

Thus Ω is the set of outcomes from the experiment of sampling n times without replacement from the population $1, \dots, n$. We let $\mathcal{B} = \mathcal{P}(\Omega)$ be the power set of Ω and define for $(x_1, \dots, x_n) \in \Omega$

$$P((x_1, \dots, x_n)) = \frac{1}{n!},$$

and for $B \in \mathcal{B}$

$$P(B) = \frac{1}{n!}\#\text{elements in } B.$$

For $i = 1, \dots, n$, let A_i be the set of all elements of Ω with i in the ith spot. Thus, for instance,

$$\begin{aligned}
A_1 &= \{(1, x_2, \dots, x_n) : (1, x_2, \dots, x_n) \in \Omega\}, \\
A_2 &= \{(x_1, 2, \dots, x_n) : (x_1, 2, \dots, x_n) \in \Omega\}.
\end{aligned}$$

and so on. We need to compute $P(\cup_{i=1}^n A_i)$. From the inclusion-exclusion formula (2.2) we have

$$P(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i) - \sum_{1 \le i < j \le n} P(A_i A_j) + \sum_{1 \le i < j < k \le n} P(A_i A_j A_k)$$
$$- \dots (-1)^{n+1} P(A_1 A_2 \dots A_n).$$

To compute $P(A_i)$, we fix integer i in the ith spot and count the number of ways to distribute $n-1$ objects in $n-1$ spots, which is $(n-1)!$ and then divide by $n!$. To compute $P(A_i A_j)$ we fix i and j and count the number of ways to distribute $n-2$ integers into $n-2$ spots, and so on. Thus

$$P(\bigcup_{i=1}^n A_i) = n\frac{(n-1)!}{n!} - \binom{n}{2}\frac{(n-2)!}{n!} + \binom{n}{3}\frac{(n-3)!}{n!} - \dots (-1)^n \frac{1}{n!}$$
$$= 1 - \frac{1}{2!} + \frac{1}{3!} - \dots (-1)^n \frac{1}{n!}.$$

Taking into account the expansion of e^x for $x = -1$ we see that for large n, the probability of a coincidence is approximately

$$P(\bigcup_{i=1}^n A_i) \approx 1 - e^{-1} \approx 0.632.$$

□

2.2 More on Closure

A σ-field is a collection of subsets of Ω satisfying certain closure properties, namely closure under complementation and countable union. We will have need of collections of sets satisfying different closure axioms. We define a *structure* $\mathcal{G}$ to be a collection of subsets of Ω satisfying certain specified closure axioms. Here are some other structures. Some have been discussed, some will be discussed and some are listed but will not be discussed or used here.

- field
- σ-field
- semialgebra
- semiring
- ring
- σ-ring
- monotone class (closed under monotone limits)

- π-system ($\mathcal{P}$ is a π-system, if it is closed under finite intersections: $A, B \in \mathcal{P}$ implies $A \cap B \in \mathcal{P}$).
- λ-system (synonyms: σ-additive class, Dynkin class); this will be used extensively as the basis of our most widely used induction technique.

Fix a structure in mind. Call it $\mathcal{S}$. As with σ-algebras, we can make the following definition.

Definition 2.2.1 *The minimal structure $\mathcal{S}$ generated by a class $\mathcal{C}$ is a non-empty structure satisfying*

(i) $\mathcal{S} \supset \mathcal{C}$,

(ii) If $\mathcal{S}'$ is some other structure containing $\mathcal{C}$, then $\mathcal{S}' \supset \mathcal{S}$.

Denote the minimal structure by $\mathcal{S}(\mathcal{C})$.

Proposition 2.2.1 *The minimal structure $\mathcal{S}$ exists and is unique.*

As we did with generating a minimal σ-field, let

$$\aleph = \{\mathcal{G} : \mathcal{G} \text{ is a structure }, \mathcal{G} \supset \mathcal{C}\}$$

and

$$\mathcal{S}(\mathcal{C}) = \cap_{\mathcal{G} \in \aleph} \mathcal{G}.$$

2.2.1 *Dynkin's theorem*

Dynkin's theorem is a remarkably flexible device for performing set inductions which is ideally suited to probability theory.

A class of subsets $\mathcal{L}$ of Ω is a called a λ-system if it satisfies either the *new* postulates $\lambda_1, \lambda_2, \lambda_3$ or the old postulates $\lambda_1', \lambda_2', \lambda_3'$ given in the following table.

λ-system postulates			
old		new	
λ_1'	$\Omega \in \mathcal{L}$	λ_1	$\Omega \in \mathcal{L}$
λ_2'	$A, B \in \mathcal{L},\ A \subset B \Rightarrow B \setminus A \in \mathcal{L}$	λ_2	$A \in \mathcal{L} \Rightarrow A^c \in \mathcal{L}$
λ_3'	$A_n \uparrow, A_n \in \mathcal{L} \Rightarrow \cup_n A_n \in \mathcal{L}$	λ_3	$n \neq m, A_n A_m = \emptyset,$ $A_n \in \mathcal{L} \Rightarrow \cup_n A_n \in \mathcal{L}.$

The *old* postulates are equivalent to the *new* ones. Here we only check that *old* implies *new*. Suppose $\lambda_1', \lambda_2', \lambda_3'$ are true. Then λ_1 is true. Since $\Omega \in \mathcal{L}$, if $A \in \mathcal{L}$, then $A \subset \Omega$ and by λ_2', $\Omega \setminus A = A^c \in \mathcal{L}$, which shows that λ_2 is true. If $A, B \in \mathcal{L}$ are disjoint, we show that $A \cup B \in \mathcal{L}$. Now $\Omega \setminus A \in \mathcal{L}$ and $B \subset \Omega \setminus A$ (since $\omega \in B$ implies $\omega \notin A$ which means $\omega \in A^c = \Omega \setminus A$) so by λ_2' we have $(\Omega \setminus A) \setminus B = A^c B^c \in \mathcal{L}$ and by λ_2 we have $(A^c B^c)^c = A \cup B \in \mathcal{L}$ which is λ_3 for finitely many sets. Now if $A_j \in \mathcal{L}$ are mutually disjoint for $j = 1, 2, \ldots,$

define $B_n = \cup_{j=1}^n A_j$. Then $B_n \in \mathcal{L}$ by the prior argument for 2 sets and by λ_3' we have $\cup_n B_n = \lim_{n\to\infty} \uparrow B_n \in \mathcal{L}$. Since $\cup_n B_n = \cup_n A_n$ we have $\cup_n A_n \in \mathcal{L}$ which is λ_3. □

Remark. It is clear that a σ-field is always a λ-system since the *new* postulates obviously hold.

Recall that a π-system is a class of sets closed under finite intersections; that is, $\mathcal{P}$ is a π-system if whenever $A, B \in \mathcal{P}$ we have $AB \in \mathcal{P}$.

We are now in a position to state Dynkin's theorem.

Theorem 2.2.2 (Dynkin's theorem) *(a) If $\mathcal{P}$ is a π-system and $\mathcal{L}$ is a λ-system such that $\mathcal{P} \subset \mathcal{L}$, then $\sigma(\mathcal{P}) \subset \mathcal{L}$.*

(b) If $\mathcal{P}$ is a π-system

$$\sigma(\mathcal{P}) = \mathcal{L}(\mathcal{P}),$$

that is, the minimal σ-field over $\mathcal{P}$ equals the minimal λ-system over $\mathcal{P}$.

Note (b) follows from (a). To see this assume (a) is true. Since $\mathcal{P} \subset \mathcal{L}(\mathcal{P})$, we have from (a) that $\sigma(\mathcal{P}) \subset \mathcal{L}(\mathcal{P})$. On the other hand, $\sigma(\mathcal{P})$, being a σ-field, is a λ-system containing $\mathcal{P}$ and hence contains the minimal λ-system over $\mathcal{P}$, so that $\sigma(\mathcal{P}) \supset \mathcal{L}(\mathcal{P})$.

Before the proof of (a), here is a significant application of Dynkin's theorem.

Proposition 2.2.3 *Let P_1, P_2 be two probability measures on $(\Omega, \mathcal{B})$. The class*

$$\mathcal{L} := \{A \in \mathcal{B} : P_1(A) = P_2(A)\}$$

is a λ-system.

Proof of Proposition 2.2.3. We show the *new* postulates hold:

(λ_1) $\Omega \in \mathcal{L}$ since $P_1(\Omega) = P_2(\Omega) = 1$.

(λ_2) $A \in \mathcal{L}$ implies $A^c \in \mathcal{L}$, since $A \in \mathcal{L}$ means $P_1(A) = P_2(A)$, from which

$$P_1(A^c) = 1 - P_1(A) = 1 - P_2(A) = P_2(A^c).$$

(λ_3) If $\{A_j\}$ is a mutually disjoint sequence of events in $\mathcal{L}$, then $P_1(A_j) = P_2(A_j)$ for all j, and hence

$$P_1(\bigcup_j A_j) = \sum_j P_1(A_j) = \sum_j P_2(A_j) = P_2(\bigcup_j A_j)$$

so that

$$\bigcup_j A_j \in \mathcal{L}.$$

□

Corollary 2.2.1 *If P_1, P_2 are two probability measures on $(\Omega, \mathcal{B})$ and if $\mathcal{P}$ is a π-system such that*

$$\forall A \in \mathcal{P}: \quad P_1(A) = P_2(A),$$

then

$$\forall B \in \sigma(\mathcal{P}): \quad P_1(B) = P_2(B).$$

Proof of Corollary 2.2.1. We have

$$\mathcal{L} = \{A \in \mathcal{B} : P_1(A) = P_2(A)\}$$

is a λ-system. But $\mathcal{L} \supset \mathcal{P}$ and hence by Dynkin's theorem $\mathcal{L} \supset \sigma(\mathcal{P})$. □

Corollary 2.2.2 *Let $\Omega = \mathbb{R}$. Let P_1, P_2 be two probability measures on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that their distribution functions are equal:*

$$\forall x \in \mathbb{R}: \quad F_1(x) = P_1((-\infty, x]) = F_2(x) = P_2((-\infty, x]).$$

Then

$$P_1 \equiv P_2$$

on $\mathcal{B}(\mathbb{R})$.

So a probability measure on $\mathbb{R}$ is uniquely determined by its distribution function.

Proof of Corollary 2.2.2. Let

$$\mathcal{P} = \{(-\infty, x] : x \in \mathbb{R}\}.$$

Then $\mathcal{P}$ is a π-system since

$$(-\infty, x] \cap (-\infty, y] = (-\infty, x \wedge y] \in \mathcal{P}.$$

Furthermore $\sigma(\mathcal{P}) = \mathcal{B}(\mathbb{R})$ since the Borel sets can be generated by the semi-infinite intervals (see Section 1.7). So $F_1(x) = F_2(x)$ for all $x \in \mathbb{R}$, means $P_1 = P_2$ on $\mathcal{P}$ and hence $P_1 = P_2$ on $\sigma(\mathcal{P}) = \mathcal{B}(\mathbb{R})$. □

2.2.2 *Proof of Dynkin's theorem*

Recall that we only need to prove: If $\mathcal{P}$ is a π-system and $\mathcal{L}$ is a λ-system then $\mathcal{P} \subset \mathcal{L}$ implies $\sigma(\mathcal{P}) \subset \mathcal{L}$.

We begin by proving the following proposition.

Proposition 2.2.4 *If a class $\mathcal{C}$ is both a π-system and a λ-system, then it is a σ-field.*

Proof of Proposition 2.2.4. First we show $\mathcal{C}$ is a field: We check the field postulates.

(i) $\Omega \in \mathcal{C}$ since $\mathcal{C}$ is a λ-system.

(ii) $A \in \mathcal{C}$ implies $A^c \in \mathcal{C}$ since $\mathcal{C}$ is a λ-system.

(iii) If $A_j \in \mathcal{C}$, for $j = 1, \dots, n$, then $\cap_{j=1}^n A_j \in \mathcal{C}$ since $\mathcal{C}$ is a π-system.

Knowing that $\mathcal{C}$ is a field, in order to show that it is a σ-field we need to show that if $A_j \in \mathcal{C}$, for $j \geq 1$, then $\cup_{j=1}^\infty A_j \in \mathcal{C}$. Since

$$\bigcup_{j=1}^{\infty} A_j = \lim_{n\to\infty} \uparrow \bigcup_{j=1}^{n} A_j$$

and $\cup_{j=1}^n A_j \in \mathcal{C}$ (since $\mathcal{C}$ is a field) it suffices to show $\mathcal{C}$ is closed under monotone non-decreasing limits. This follows from the *old* postulate λ_3'. □

We can now prove Dynkin's theorem.

Proof of Dynkin's Theorem 2.2.2. It suffices to show $\mathcal{L}(\mathcal{P})$ is a π-system since $\mathcal{L}(\mathcal{P})$ is both a π-system and a λ-system, and thus by Proposition 2.2.4 also a σ-field. This means that

$$\mathcal{L} \supset \mathcal{L}(\mathcal{P}) \supset \mathcal{P}.$$

Since $\mathcal{L}(\mathcal{P})$ is a σ-field containing $\mathcal{P}$,

$$\mathcal{L}(\mathcal{P}) \supset \sigma(\mathcal{P})$$

from which

$$\mathcal{L} \supset \mathcal{L}(\mathcal{P}) \supset \sigma(\mathcal{P}),$$

and therefore we get the desired conclusion that

$$\mathcal{L} \supset \sigma(\mathcal{P}).$$

We now concentrate on showing that $\mathcal{L}(\mathcal{P})$ is a π-system. Fix a set $A \in \sigma(\mathcal{P})$ and relative to this A, define

$$\mathcal{G}_A = \{B \in \sigma(\mathcal{P}) : AB \in \mathcal{L}(\mathcal{P})\}.$$

We proceed in a series of steps.

[A] If $A \in \mathcal{L}(\mathcal{P})$, we claim that $\mathcal{G}_A$ is a λ-system.

To prove [A] we check the *new* λ-system postulates.

(i) We have

$$\Omega \in \mathcal{G}_A$$

since $A\Omega = A \in \mathcal{L}(\mathcal{P})$ by assumption.

(ii) Suppose $B \in \mathcal{G}_A$. We have that $B^c A = A \setminus AB$. But $B \in \mathcal{G}_A$ means $AB \in \mathcal{L}(\mathcal{P})$ and since by assumption $A \in \mathcal{L}(\mathcal{P})$, we have $A \setminus AB = B^c A \in \mathcal{L}(\mathcal{P})$ since λ-systems are closed under proper differences. Since $B^c A \in \mathcal{L}(\mathcal{P})$, it follows that $B^c \in \mathcal{G}_A$ by definition.

(iii) Suppose $\{B_j\}$ is a mutually disjoint sequence and $B_j \in \mathcal{G}_A$. Then

$$A \cap (\bigcup_{j=1}^{\infty} B_j) = \bigcup_{j=1}^{\infty} AB_j$$

is a disjoint union of sets in $\mathcal{L}(\mathcal{P})$, and hence in $\mathcal{L}(\mathcal{P})$.

[B] Next, we claim that if $A \in \mathcal{P}$, then $\mathcal{L}(\mathcal{P}) \subset \mathcal{G}_A$.

To prove this claim, observe that since $A \in \mathcal{P} \subset \mathcal{L}(\mathcal{P})$, we have from [A] that $\mathcal{G}_A$ is a λ-system.

For $B \in \mathcal{P}$, we have $AB \in \mathcal{P}$ since by assumption $A \in \mathcal{P}$ and $\mathcal{P}$ is a π-system. So if $B \in \mathcal{P}$, then $AB \in \mathcal{P} \subset \mathcal{L}(\mathcal{P})$ implies $B \in \mathcal{G}_A$; that is

$$\mathcal{P} \subset \mathcal{G}_A. \tag{2.4}$$

Since $\mathcal{G}_A$ is a λ-system, $\mathcal{G}_A \supset \mathcal{L}(\mathcal{P})$.

[B′] We may rephrase [B] using the definition of $\mathcal{G}_A$ to get the following statement. If $A \in \mathcal{P}$, and $B \in \mathcal{L}(\mathcal{P})$, then $AB \in \mathcal{L}(\mathcal{P})$. (So we are making progress toward our goal of showing $\mathcal{L}(\mathcal{P})$ is a π-system.)

[C] We now claim that if $A \in \mathcal{L}(\mathcal{P})$, then $\mathcal{L}(\mathcal{P}) \subset \mathcal{G}_A$.

To prove [C]: If $B \in \mathcal{P}$ and $A \in \mathcal{L}(\mathcal{P})$, then from [B′] (interchange the roles of the sets A and B) we have $AB \in \mathcal{L}(\mathcal{P})$. So when $A \in \mathcal{L}(\mathcal{P})$,

$$\mathcal{P} \subset \mathcal{G}_A.$$

From [A], $\mathcal{G}_A$ is a λ-system so $\mathcal{L}(\mathcal{P}) \subset \mathcal{G}_A$.

[C′] To finish, we rephrase [C]: If $A \in \mathcal{L}(\mathcal{P})$, then for any $B \in \mathcal{L}(\mathcal{P})$, $B \in \mathcal{G}_A$. This says that

$$AB \in \mathcal{L}(\mathcal{P})$$

as desired. □

2.3 Two Constructions

Here we give two simple examples of how to construct probability spaces. These examples will be familiar from earlier probability studies and from Example 2.1.2, but can now be viewed from a more mature perspective. The task of constructing more general probability models will be considered in the next Section 2.4

(i) Discrete models: Suppose $\Omega = \{\omega_1, \omega_2, \dots\}$ is countable. For each i, associate to ω_i the number p_i where

$$\forall i \geq 1,\ p_i \geq 0 \text{ and } \sum_{i=1}^{\infty} p_i = 1.$$

Define $\mathcal{B} = \mathcal{P}(\Omega)$, and for $A \in \mathcal{B}$, set

$$P(A) = \sum_{\omega_i \in A} p_i.$$

Then we have the following properties of P:

(i) $P(A) \geq 0$ for all $A \in \mathcal{B}$.

(ii) $P(\Omega) = \sum_{i=1}^{\infty} p_i = 1$.

(iii) P is σ-additive: If A_j, $j \geq 1$ are mutually disjoint subsets, then

$$\begin{aligned} P(\bigcup_{j=1}^{\infty} A_j) = \sum_{\omega_i \in \cup_j A_j} p_i &= \sum_j \sum_{\omega_i \in A_j} p_i \\ &= \sum_j P(A_j). \end{aligned}$$

Note this last step is justified because the series, being positive, can be added in any order.

This gives the general construction of probabilities when Ω is countable. Next comes a time honored specific example of countable state space model.

(ii) Coin tossing N times: What is an appropriate probability space for the experiment "toss a weighted coin N times"? Set

$$\Omega = \{0, 1\}^N = \{(\omega_1, \dots, \omega_N) : \omega_i = 0 \text{ or } 1\}.$$

For $p \geq 0, q \geq 0, p + q = 1$, define

$$p_{(\omega_1, \dots, \omega_N)} = p^{\sum_{j=1}^{N} \omega_j} q^{N - \sum_{j=1}^{N} \omega_j} = p^{\#1's} q^{\#0's}.$$

Construct a probability measure P as in (i) above: Let $\mathcal{B} = \mathcal{P}(\Omega)$ and for $A \subset \Omega$ define

$$P(A) = \sum_{\omega \in A} p_\omega.$$

As in (i) above, this gives a probability model provided $\sum_{\omega\in\Omega} p_\omega = 1$. Note the product form

$$p_{(\omega_1,\dots,\omega_N)} = \prod_{i=1}^{N} p^{\omega_i} q^{1-\omega_i}$$

so

$$\begin{aligned}\sum_{\omega_1,\dots,\omega_N} p_{\omega_1,\dots,\omega_N} &= \sum_{\omega_1,\dots,\omega_N} \prod_{i=1}^{n} p^{\omega_i} q^{1-\omega_i} \\ &= \sum_{\omega_1,\dots,\omega_{N-1}} \prod_{i=1}^{N-1} p^{\omega_i} q^{1-\omega_i} \underbrace{(p^1q^0 + p^0q^1)}_{1} = \cdots = 1.\end{aligned}$$

□

2.4 Constructions of Probability Spaces

The previous section described how to construct a probability space when the sample space Ω is countable. A more complex case but very useful in applications is when Ω is uncountable, for example, when $\Omega = \mathbb{R}, \mathbb{R}^k, \mathbb{R}^\infty$, and so on. For these and similar cases, how do we construct a probability space which will have given desirable properties? For instance, consider the following questions.

(i) Given a distribution function $F(x)$, let $\Omega = \mathbb{R}$. How do we construct a probability measure P on $\mathcal{B}(\mathbb{R})$ such that the distribution function corresponding to P is F:
$$P((-\infty, x]) = F(x).$$

(ii) How do you construct a probability space containing an iid sequence of random variables or a sequence of random variables with given finite dimensional distributions.

A simple case of this question: How do we build a model of an infinite sequence of coin tosses so we can answer questions such as:

(a) What is the probability that heads occurs infinitely often in an infinite sequence of coin tosses; that is, how do we compute
$$P[\text{ heads occurs i.o. }]?$$

(b) How do we compute the probability that ultimately the excess of heads over tails is at least 17?

(c) In a gambling game where a coin is tossed repeatedly and a heads results in a gain of one dollar and a tail results in a loss of one dollar, what is the probability that starting with a fortune of x, ruin eventually occurs; that is, eventually my stake is wiped out?

For these and similar questions, we need uncountable spaces. For the coin tossing problems we need the sample space

$$\begin{aligned}\Omega =&\{0,1\}^{\mathbb{N}}\\ =&\{(\omega_1,\omega_2,\dots):\omega_i\in\{0,1\},\ i\geq 1\}.\end{aligned}$$

2.4.1 General Construction of a Probability Model

The general method is to start with a sample space Ω and a restricted, simple class of subsets $\mathcal{S}$ of Ω to which the assignment of probabilities is obvious or natural. Then this assignment of probabilities is extended to $\sigma(\mathcal{S})$. For example, if $\Omega=\mathbb{R}$, the real line, and we are given a distribution function F, we could take $\mathcal{S}$ to be

$$\mathcal{S}=\{(a,b]:-\infty\leq a\leq b\leq\infty\}$$

and then define P on $\mathcal{S}$ to be

$$P((a,b])=F(b)-F(a).$$

The problem is to extend the definition of P from $\mathcal{S}$ to $\mathcal{B}(\mathbb{R})$, the Borel sets.

For what follows, recall the notational convention that $\sum_{i=1}^n A_i$ means a disjoint union; that is, that $A_1,\dots,A_n$ are mutually disjoint and

$$\sum_{i=1}^n A_i=\bigcup_{i=1}^n A_i.$$

The following definitions help clarify language and proceedings. Given two structures $\mathcal{G}_1,\mathcal{G}_2$ of subsets of Ω such that $\mathcal{G}_1\subset\mathcal{G}_2$ and two set functions

$$P_i:\mathcal{G}_i\mapsto[0,1],\quad i=1,2,$$

we say P_2 is an *extension* of P_1 (or P_1 extends to P_2) if P_2 restricted to $\mathcal{G}_1$ equals P_1. This is written

$$P_2|_{\mathcal{G}_1}=P_1$$

and means $P_2(A_1)=P_1(A_1)$ for all $A_1\in\mathcal{G}_1$. A set function P with structure $\mathcal{G}$ as domain and range $[0,1]$,

$$P:\mathcal{G}\mapsto[0,1],$$

is ***additive*** if for any $n\geq 1$ and any disjoint $A_1,\dots,A_n\in\mathcal{G}$ such that $\sum_{i=1}^n A_i\in\mathcal{G}$ we have

$$P(\sum_{i=1}^n A_i)=\sum_{i=1}^n P(A_i).\tag{2.5}$$

Call P ***σ-additive*** if the index n can be replaced by ∞; that is, (2.5) holds for mutually disjoint $\{A_n,n\geq 1\}$ with $A_j\in\mathcal{G}$, $j\geq 1$ and $\sum_{j=1}^\infty A_j\in\mathcal{G}$.

We now define a primitive structure called a *semialgebra*.

Definition 2.4.1 A class $\mathcal{S}$ of subsets of Ω is a *semialgebra* if the following postulates hold:

(i) $\emptyset, \Omega \in \mathcal{S}$.

(ii) $\mathcal{S}$ is a π-system; that is, it is closed under finite intersections.

(iii) If $A \in \mathcal{S}$, then there exist some finite n and disjoint sets $C_1, \dots, C_n$, with each $C_i \in \mathcal{S}$ such that $A^c = \sum_{i=1}^n C_i$.

The plan is to start with a probability measure on the primitive structure $\mathcal{S}$, show there is a unique extension to $\mathcal{A}(\mathcal{S})$, the algebra (field) generated by $\mathcal{S}$ (first extension theorem) and then show there is a unique extension from $\mathcal{A}(\mathcal{S})$ to $\sigma(\mathcal{A}(\mathcal{S})) = \sigma(\mathcal{S})$, the σ-field generated by $\mathcal{S}$ (second extension theorem).

Before proceeding, here are standard examples of semialgebras.

Examples:

(a) Let $\Omega = \mathbb{R}$, and suppose $\mathcal{S}_1$ consists of intervals including $\emptyset$, the empty set:

$$\mathcal{S}_1 = \{(a, b] : -\infty \le a \le b \le \infty\}.$$

If $I_1, I_2 \in \mathcal{S}_1$, then $I_1 I_2$ is an interval and in $\mathcal{S}_1$ and if $I \in \mathcal{S}_1$, then I^c is a union of disjoint intervals.

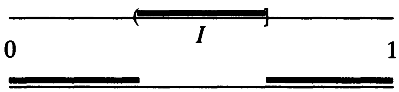

FIGURE 2.1 Intervals

(b) Let

$$\Omega = \mathbb{R}^k = \{(x_1, \dots, x_k) : x_i \in \mathbb{R}, i = 1, \dots, k\}$$
$$\mathcal{S}_k = \text{ all rectangles (including } \emptyset\text{, the empty set).}$$

Note that we call A a rectangle if it is of the form

$$A = I_1 \times \cdots \times I_k$$

where $I_j \in \mathcal{S}_1$ is an interval, $j = 1, \dots, k$ as in item (a) above. Obviously $\emptyset, \Omega$ are rectangles and intersections of rectangles are rectangles. When $k = 2$ and A is a rectangle, the picture of A^c appears in Figure 2.2, showing A^c can be written as a disjoint union of rectangles.

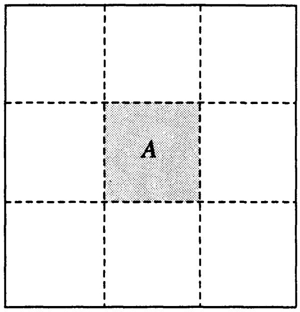

FIGURE 2.2 Rectangles

For general k, let

$$A = I_1 \times \cdots \times I_k, = \bigcap_{i=1}^{k} \{(x_1, \ldots, x_k) : x_i \in I_i\}$$

so that

$$A^c = \left(\bigcap_{i=1}^{k} \{(x_1, \ldots, x_k) : x_i \in I_i\} \right)^c = \bigcup_{i=1}^{k} \{(x_1, \ldots, x_k) : x_i \in I_i^c\}.$$

Since $I_i \in \mathcal{S}_1$, we have $I_i^c = I_i' + I_i''$, where $I_i', I_i'' \in \mathcal{S}_1$ are intervals. Consider

$$\mathcal{D} := \{U_1 \times \cdots \times U_k : U_\alpha = I_\alpha \text{ or } I_\alpha' \text{ or } I_\alpha'',\ \alpha = 1, \ldots, k\}.$$

When $U_\alpha = I_\alpha,\ \alpha = 1, \ldots, k$, then $U_1 \times \ldots \times U_k = A$. So

$$A^c = \sum_{\substack{U_1 \times \cdots \times U_k \in \mathcal{D} \\ \text{Not all } U_\alpha = I_\alpha,\ \alpha = 1, \ldots, k}} U_1 \times \cdots \times U_k.$$

This shows that $\mathcal{S}_k$ is a semialgebra. □

Starting with a semialgebra $\mathcal{S}$, we first discuss the structure of $\mathcal{A}(\mathcal{S})$, the smallest algebra or field containing $\mathcal{S}$.

Lemma 2.4.1 (The field generated by a semialgebra) *Suppose $\mathcal{S}$ is a semialgebra of subsets of Ω. Then*

$$\mathcal{A}(\mathcal{S}) = \{\sum_{i \in I} S_i : I \textit{ finite, } \{S_i, i \in I\} \textit{ disjoint, } S_i \in \mathcal{S}\}, \tag{2.6}$$

is the family of all sums of finite families of mutually disjoint subsets of Ω in $\mathcal{S}$.

Proof. Let Λ be the collection on the right side of (2.6). It is clear that $\Lambda \supset \mathcal{S}$ (take I to be a singleton set) and we claim Λ is a field. We check the field postulates in Definition 1.5.2, Chapter 1 on page 12:

(i) $\Omega \in \Lambda$ since $\Omega \in \mathcal{S}$.

(iii) If $\sum_{i \in I} S_i$ and $\sum_{j \in J} S'_j$ are two members of Λ, then

$$\left(\sum_{i \in I} S_i\right) \bigcap \left(\sum_{j \in J} S'_j\right) = \sum_{(i,j) \in I \times J} S_i S'_j \in \Lambda$$

since $\{S_i S'_j, (i, j) \in I \times J\}$ is a finite, disjoint collection of members of the π-system $\mathcal{S}$.

(ii) To check closure under complementation, let $\sum_{i \in I} S_i \in \Lambda$ and observe

$$\left(\sum_{i \in I} S_i\right)^c = \bigcap_{i \in I} S_i^c.$$

But from the axioms defining a semialgebra, $S_i \in \mathcal{S}$ implies

$$S_i^c = \sum_{j \in J_i} S_{ij}$$

for a finite index set J_i and disjoint sets $\{S_{ij}, j \in J_i\}$ in $\mathcal{S}$. Now observe that $\cap_{i \in I} S_i^c \in \Lambda$ by the previously proven (iii).

So Λ is a field, $\Lambda \supset \mathcal{S}$ and hence $\Lambda \supset \mathcal{A}(\mathcal{S})$. Since also

$$\sum_{i \in I} S_i \in \Lambda \text{ implies } \sum_{i \in I} S_i \in \mathcal{A}(\mathcal{S}),$$

we get $\Lambda \subset \mathcal{A}(\mathcal{S})$ and thus, as desired, $\Lambda = \mathcal{A}(\mathcal{S})$. □

It is now relatively easy to extend a probability measure from $\mathcal{S}$ to $\mathcal{A}(\mathcal{S})$.

Theorem 2.4.1 (First Extension Theorem) *Suppose $\mathcal{S}$ is a semialgebra of subsets of Ω and $P : \mathcal{S} \mapsto [0, 1]$ is σ-additive on $\mathcal{S}$ and satisfies $P(\Omega) = 1$. There is a unique extension P' of P to $\mathcal{A}(\mathcal{S})$, defined by*

$$P'(\sum_{i \in I} S_i) = \sum_{i \in I} P(S_i), \tag{2.7}$$

which is a probability measure on $\mathcal{A}(\mathcal{S})$; that is $P'(\Omega) = 1$ and P' is σ-additive on $\mathcal{A}(\mathcal{S})$.

Proof. We must first check that (2.7) defines P' unambiguously and to do this, suppose $A \in \mathcal{A}(\mathcal{S})$ has two distinct representations

$$A = \sum_{i \in I} S_i = \sum_{j \in J} S'_j.$$

We need to verify that

$$\sum_{i\in I} P(S_i) = \sum_{j\in J} P(S'_j) \tag{2.8}$$

so that P' has a unique value at A. Confirming (2.8) is easy since $S_i \subset A$ and therefore

$$\begin{aligned}\sum_{i\in I} P(S_i) &= \sum_{i\in I} P(S_i A) = \sum_{i\in I} P(S_i \cap \sum_{j\in J} S'_j) \\ &= \sum_{i\in I} P(\sum_{j\in J} S_i S'_j)\end{aligned}$$

and using the fact that $S_i = \sum_{j\in J} S_i S'_j \in \mathcal{S}$ and P is additive on $\mathcal{S}$, we get the above equal to

$$= \sum_{i\in I}\sum_{j\in J} P(S_i S'_j) = \sum_{j\in J}\sum_{i\in I} P(S_i S'_j).$$

Reversing the logic, this equals

$$= \sum_{j\in J} P(S'_j)$$

as required.

Now we check that P' is σ-additive on $\mathcal{A}(\mathcal{S})$. Thus suppose for $i \geq 1$,

$$A_i = \sum_{j\in J_i} S_{ij} \in \mathcal{A}(\mathcal{S}), \quad S_{ij} \in \mathcal{S},$$

and $\{A_i, i \geq 1\}$ are mutually disjoint and

$$A = \sum_{i=1}^{\infty} A_i \in \mathcal{A}(\mathcal{S}).$$

Since $A \in \mathcal{A}(\mathcal{S})$, A also has a representation

$$A = \sum_{k\in K} S_k, \quad S_k \in \mathcal{S},\ k \in K,$$

where K is a finite index set. From the definition of P', we have

$$P'(A) = \sum_{k\in K} P(S_k).$$

Write

$$S_k = S_k A = \sum_{i=1}^{\infty} S_k A_i = \sum_{i=1}^{\infty}\sum_{j\in J_i} S_k S_{ij}.$$

Now $S_k S_{ij} \in \mathcal{S}$ and $\sum_{i=1}^{\infty} \sum_{j \in J_i} S_k S_{ij} = S_k \in \mathcal{S}$, and since P is σ-additive on $\mathcal{S}$, we have

$$\sum_{k \in K} P(S_k) = \sum_{k \in K} \sum_{i=1}^{\infty} \sum_{j \in J_i} P(S_k S_{ij}) = \sum_{i=1}^{\infty} \sum_{j \in J_i} \sum_{k \in K} P(S_k S_{ij}).$$

Again observe

$$\sum_{k \in K} S_k S_{ij} = A S_{ij} = S_{ij} \in \mathcal{S}$$

and by additivity of P on $\mathcal{S}$

$$\sum_{i=1}^{\infty} \sum_{j \in J_i} \sum_{k \in K} P(S_k S_{ij}) = \sum_{i=1}^{\infty} \sum_{j \in J_i} P(S_{ij}),$$

and continuing in the same way, we get this equal to

$$= \sum_{i=1}^{\infty} P(\sum_{j \in J_i} S_{ij}) = \sum_{i=1}^{\infty} P'(A_i)$$

as desired.

Finally, it is clear that P has a unique extension from $\mathcal{S}$ to $\mathcal{A}(\mathcal{S})$, since if P_1' and P_2' are two additive extensions, then for any

$$A = \sum_{i \in I} S_i \in \mathcal{A}(\mathcal{S})$$

we have

$$P_1'(A) = \sum_{i \in I} P(S_i) = P_2'(A).$$

□

Now we know how to extend a probability measure from $\mathcal{S}$ to $\mathcal{A}(\mathcal{S})$. The next step is to extend the probability measure from the algebra to the σ-algebra.

Theorem 2.4.2 (Second Extension Theorem) *A probability measure P defined on a field $\mathcal{A}$ of subsets has a unique extension to a probability measure on $\sigma(\mathcal{A})$, the σ-field generated by $\mathcal{A}$.*

Combining the First and Second Extension Theorems 2.4.1 and 2.4.2 yields the final result.

Theorem 2.4.3 (Combo Extension Theorem) *Suppose $\mathcal{S}$ is a semialgebra of subsets of Ω and that P is a σ-additive set function mapping $\mathcal{S}$ into* $[0, 1]$ *such that* $P(\Omega) = 1$. *There is a unique probability measure on $\sigma(\mathcal{S})$ that extends P.*

The ease with which this result can be applied depends largely on how easily one can check that a set function P defined on $\mathcal{S}$ is σ-additive (as opposed to just being additive). Sometimes some sort of compactness argument is needed.

The proof of the Second Extension Theorem 2.4.2 is somewhat longer than the proof of the First Extension Theorem and is deferred to the next Subsection 2.4.2.

2.4.2 Proof of the Second Extension Theorem

We now prove the Second Extension Theorem. We start with a field $\mathcal{A}$ and a probability measure P on $\mathcal{A}$ so that $P(\Omega) = 1$, and for all $A \in \mathcal{A}$, $P(A) \geq 0$ and for $\{A_i\}$ disjoint, $A_i \in \mathcal{A}$, $\sum_{i=1}^{\infty} A_i \in \mathcal{A}$, we have $P(\sum_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$.

The proof is broken into 3 parts. In Part I, we extend P to a set function Π on a class $\mathcal{G} \supset \mathcal{A}$. In Part II we extend Π to a set function Π^* on a class $\mathcal{D} \supset \sigma(\mathcal{A})$ and in Part III we restrict Π^* to $\sigma(\mathcal{A})$ yielding the desired extension.

PART I. We begin by defining the class $\mathcal{G}$:

$$\begin{aligned}\mathcal{G} :=& \{\bigcup_{j=1}^{\infty} A_j : A_j \in \mathcal{A}\}\\ =& \{\lim_{n\to\infty} \uparrow B_n : B_n \in \mathcal{A}, B_n \subset B_{n+1}, \forall n\}.\end{aligned}$$

So $\mathcal{G}$ is the class of unions of countable collections of sets in $\mathcal{A}$, or equivalently, since $\mathcal{A}$ is a field, $\mathcal{G}$ is the class of non-decreasing limits of elements of $\mathcal{A}$.

We also define a set function $\Pi : \mathcal{G} \mapsto [0, 1]$ via the following definition: If $G = \lim_{n\to\infty} \uparrow B_n \in \mathcal{G}$, where $B_n \in \mathcal{A}$, define

$$\Pi(G) = \lim_{n\to\infty} \uparrow P(B_n). \tag{2.9}$$

Since P is σ-additive on $\mathcal{A}$, P is monotone on $\mathcal{A}$, so the monotone convergence indicated in (2.9) is justified. Call the sequence $\{B_n\}$ the *approximating* sequence to G. To verify that Π is well defined, we need to check that if G has two approximating sequences $\{B_n\}$ and $\{B_n'\}$,

$$G = \lim_{n\to\infty} \uparrow B_n = \lim_{n\to\infty} \uparrow B_n'$$

then

$$\lim_{n\to\infty} \uparrow P(B_n) = \lim_{n\to\infty} \uparrow P(B_n').$$

This is verified in the next lemma whose proof is typical of this sort of uniqueness proof in that some sort of merging of two approximating sequences takes place.

Lemma 2.4.2 *If $\{B_n\}$ and $\{B_n'\}$ are two non-decreasing sequences of sets in $\mathcal{A}$ and*

$$\bigcup_{n=1}^{\infty} B_n \subset \bigcup_{n=1}^{\infty} B_n',$$

then

$$\lim_{n\to\infty} \uparrow P(B_n) \leq \lim_{n\to\infty} \uparrow P(B_n').$$

Proof. For fixed m

$$\lim_{n\to\infty} \uparrow B_m B_n' = B_m. \tag{2.10}$$

Since also

$$B_m B_n' \subset B_n'$$

and P is continuous with respect to monotonely converging sequences as a consequence of being σ-additive (see Item 7 on page 31), we have

$$\lim_{n\to\infty} \uparrow P(B_n') \geq \lim_{n\to\infty} \uparrow P(B_m B_n') = P(B_m),$$

where the last equality results from (2.10) and P being continuous. The inequality holds for all m, so we conclude that

$$\lim_{n\to\infty} \uparrow P(B_n') \geq \lim_{m\to\infty} \uparrow P(B_m)$$

as desired. □

Now we list some properties of Π and $\mathcal{G}$:

Property 1. We have

$$\emptyset \in \mathcal{G}, \quad \Pi(\emptyset) = 0,$$
$$\Omega \in \mathcal{G}, \quad \Pi(\Omega) = 1,$$

and for $G \in \mathcal{G}$

$$0 \leq \Pi(G) \leq 1. \tag{2.11}$$

More generally, we have $\mathcal{A} \subset \mathcal{G}$ and

$$\Pi|_{\mathcal{A}} = P;$$

that is, $\Pi(A) = P(A)$, for $A \in \mathcal{A}$.

The first statements are clear since, for example, if we set $B_n = \Omega$ for all n, then

$$\mathcal{A} \ni B_n = \Omega \uparrow \Omega,$$

and

$$\Pi(\Omega) = \lim_{n\to\infty} \uparrow P(\Omega) = 1$$

and a similar argument holds for $\emptyset$. The statement (2.11) follows from $0 \leq P(B_n) \leq 1$ for approximating sets $\{B_n\}$ in $\mathcal{A}$. To show $\Pi(A) = P(A)$ for $A_n \in \mathcal{A}$, take the approximating sequence to be identically equal to A.

Property 2. If $G_i \in \mathcal{G}$ for $i = 1, 2$ then

$$G_1 \cup G_2 \in \mathcal{G}, \quad G_1 \cap G_2 \in \mathcal{G},$$

and

$$\Pi(G_1 \cup G_2) + \Pi(G_1 \cap G_2) = \Pi(G_1) + \Pi(G_2). \tag{2.12}$$

This implies Π is additive on $\mathcal{G}$.

To see this, pick approximating sets $B_{n1}, B_{n2} \in \mathcal{A}$ such that $B_{ni} \uparrow G_i$ for $i = 1, 2$ as $n \to \infty$ and then, since $\mathcal{A}$ is a field, it follows that

$$\mathcal{A} \ni B_{n1} \cup B_{n2} \uparrow G_1 \cup G_2,$$
$$\mathcal{A} \ni B_{n1} \cap B_{n2} \uparrow G_1 \cap G_2,$$

showing that $G_1 \cup G_2$ and $G_1 \cap G_2$ are in $\mathcal{G}$. Further

$$P(B_{n1} \cup B_{n2}) + P(B_{n1} \cap B_{n2}) = P(B_{n1}) + P(B_{n2}), \tag{2.13}$$

from (2.1) on page 30. If we let $n \to \infty$ in (2.13), we get (2.12).

Property 3. Π is monotone on $\mathcal{G}$: If $G_i \in \mathcal{G}$, $i = 1, 2$ and $G_1 \subset G_2$, then $\Pi(G_1) \leq \Pi(G_2)$. This follows directly from Lemma 2.4.2.

Property 4. If $G_n \in \mathcal{G}$ and $G_n \uparrow G$, then $G \in \mathcal{G}$ and

$$\Pi(G) = \lim_{n\to\infty} \Pi(G_n).$$

So $\mathcal{G}$ is closed under non-decreasing limits and Π is sequentially monotonely continuous. Combining this with Property 2, we get that if $\{A_i, i \geq 1\}$ is a disjoint sequence of sets in $\mathcal{G}$, $\sum_{i=1}^{\infty} A_i \in \mathcal{G}$ and

$$\begin{aligned}\Pi(\sum_{i=1}^{\infty} A_i) &= \Pi(\lim_{n\to\infty} \uparrow \sum_{i=1}^{n} A_i) = \lim_{n\to\infty} \uparrow \Pi(\sum_{i=1}^{n} A_i)\\ &= \lim_{n\to\infty} \uparrow \sum_{i=1}^{n} \Pi(A_i) = \sum_{i=1}^{\infty} P(A_i).\end{aligned}$$

So Π is σ-additive on $\mathcal{G}$.

For each n, G_n has an approximating sequence $B_{m,n} \in \mathcal{A}$ such that

$$\lim_{m\to\infty} \uparrow B_{m,n} = G_n. \tag{2.14}$$

Define $D_m = \cup_{n=1}^{m} B_{m,n}$. Since $\mathcal{A}$ is closed under finite unions, $D_m \in \mathcal{A}$. We show

$$\lim_{m\to\infty} \uparrow D_m = G, \tag{2.15}$$

and if (2.15) is true, then G has a monotone approximating sequence of sets in $\mathcal{A}$, and hence $G \in \mathcal{G}$.

To show (2.15), we first verify $\{D_m\}$ is monotone:

$$D_m = \bigcup_{n=1}^{m} B_{m,n} \subset \bigcup_{n=1}^{m} B_{m+1,n}$$

(from (2.14))

$$\subset \bigcup_{n=1}^{m+1} B_{m+1,n} = D_{m+1}.$$

Now we show $\{D_m\}$ has the correct limit. If $n \leq m$, we have from the definition of D_m and (2.14)

$$B_{m,n} \subset D_m = \bigcup_{j=1}^{m} B_{m,j} \subset \bigcup_{j=1}^{m} G_j = G_m;$$

that is,

$$B_{m,n} \subset D_m \subset G_m. \tag{2.16}$$

Taking limits on m, we have for any $n \geq 1$,

$$G_n = \lim_{m\to\infty} \uparrow B_{m,n} \subset \lim_{m\to\infty} \uparrow D_m \subset \lim_{m\to\infty} \uparrow G_m = G$$

and now taking limits on n yields

$$G = \lim_{n\to\infty} \uparrow G_n \subset \lim_{m\to\infty} \uparrow D_m \subset \lim_{m\to\infty} \uparrow G_m = G \tag{2.17}$$

which shows $D_m \uparrow G$ and proves $G \in \mathcal{G}$. Furthermore, from the definition of Π, we know $\Pi(G) = \lim_{m\to\infty} \uparrow \Pi(D_m)$.

It remains to show $\Pi(G_n) \uparrow \Pi(G)$. From Property 2, all sets appearing in (2.16) are in $\mathcal{G}$ and from monotonicity property 3, we get

$$\Pi(B_{m,n}) \leq \Pi(D_m) \leq \Pi(G_m).$$

Let $m \to \infty$ and since $G_n = \lim_{m\to\infty} \uparrow B_{m,n}$ we get

$$\Pi(G_n) \leq \lim_{m\to\infty} \uparrow \Pi(D_m) \leq \lim_{m\to\infty} \uparrow \Pi(G_m)$$

which is true for all n. Thus letting $n \to \infty$ gives

$$\lim_{n\to\infty} \uparrow \Pi(G_n) \leq \lim_{m\to\infty} \Pi(D_m) \leq \lim_{m\to\infty} \uparrow \Pi(G_m),$$

and therefore

$$\lim_{n\to\infty} \uparrow \Pi(G_n) = \lim_{m\to\infty} \Pi(D_m).$$

The desired result follows from recalling

$$\lim_{m\to\infty} \Pi(D_m) = \Pi(G).$$

This extends P on $\mathcal{A}$ to a σ-additive set function Π on $\mathcal{G}$. $\square$

PART 2. We next extend Π to a set function Π^* on the power set $\mathcal{P}(\Omega)$ and finally show the restriction of Π^* to a certain subclass $\mathcal{D}$ of $\mathcal{P}(\Omega)$ can yield the desired extension of P.

We define $\Pi^* : \mathcal{P}(\Omega) \mapsto [0, 1]$ by

$$\forall A \in \mathcal{P}(\Omega) : \quad \Pi^*(A) = \inf\{\Pi(G) : A \subset G \in \mathcal{G}\}, \tag{2.18}$$

so $\Pi^*(A)$ is the least upper bound of values of Π on sets $G \in \mathcal{G}$ containing A.

We now consider properties of Π^*:

Property 1. We have on $\mathcal{G}$:

$$\Pi^*|_{\mathcal{G}} = \Pi \tag{2.19}$$

and $0 \leq \Pi^*(A) \leq 1$ for any $A \in \mathcal{P}(\Omega)$.

It is clear that if $A \in \mathcal{G}$, then

$$A \in \{G : A \subset G \in \mathcal{G}\}$$

and hence the infimum in (2.18) is achieved at A.

In particular, from (2.19) we get

$$\Pi^*(\Omega) = \Pi(\Omega) = 1, \quad \Pi^*(\emptyset) = \Pi(\emptyset) = 0.$$

Property 2. We have for $A_1, A_2 \in \mathcal{P}(\Omega)$

$$\Pi^*(A_1 \cup A_2) + \Pi^*(A_1 \cap A_2) \leq \Pi^*(A_1) + \Pi^*(A_2) \tag{2.20}$$

and taking $A_1 = A$, $A_2 = A^c$ in (2.20) we get

$$1 = \Pi^*(\Omega) \leq \Pi^*(A) + \Pi^*(A^c), \tag{2.21}$$

where we used the fact that $\Pi^*(\Omega) = 1$.

To verify (2.20), fix $\epsilon > 0$ and find $G_i \in \mathcal{G}$ such that $G_i \supset A_i$, and for $i = 1, 2$,

$$\Pi^*(A_i) + \frac{\epsilon}{2} \geq \Pi(G_i).$$

Adding over $i = 1, 2$ yields

$$\Pi^*(A_1) + \Pi^*(A_2) + \epsilon \geq \Pi(G_1) + \Pi(G_2).$$

By Property 2 for Π (see (2.12)), the right side equals

$$= \Pi(G_1 \cup G_2) + \Pi(G_1 \cap G_2).$$

Since $G_1 \cup G_2 \supset A_1 \cup A_2$, $G_1 \cap G_2 \supset A_1 \cap A_2$, we get from the definition of Π^* that the above is bounded below by

$$\geq \Pi^*(A_1 \cup A_2) + \Pi^*(A_1 \cap A_2).$$

Property 3. Π^* is monotone on $\mathcal{P}(\Omega)$. This follows from the fact that Π is monotone on $\mathcal{G}$.

Property 4. Π^* is sequentially monotone continuous on $\mathcal{P}(\Omega)$ in the sense that if $A_n \uparrow A$, then $\Pi^*(A_n) \uparrow \Pi^*(A)$.

To prove this, fix $\epsilon > 0$. for each $n \geq 1$, find $G_n \in \mathcal{G}$ such that $G_n \supset A_n$ and

$$\Pi^*(A_n) + \frac{\epsilon}{2^n} \geq \Pi(G_n). \tag{2.22}$$

Define $G'_n = \cup_{m=1}^n G_m$. Since $\mathcal{G}$ is closed under finite unions, $G'_n \in \mathcal{G}$ and $\{G'_n\}$ is obviously non-decreasing. We claim for all $n \geq 1$,

$$\Pi^*(A_n) + \epsilon \sum_{i=1}^{n} 2^{-i} \geq \Pi(G'_n). \tag{2.23}$$

We prove the claim by induction. For $n = 1$, the claim follows from (2.22) and the fact that $G'_1 = G_1$. Make the induction hypothesis that (2.23) holds for n and we verify (2.23) for $n + 1$. We have

$$A_n \subset G_n \subset G'_n \text{ and } A_n \subset A_{n+1} \subset G_{n+1}$$

and therefore $A_n \subset G'_n$ and $A_n \subset G_{n+1}$, so

$$A_n \subset G'_n \cap G_{n+1} \in \mathcal{G}. \tag{2.24}$$

Thus

$$\begin{aligned}\Pi(G'_{n+1}) &= \Pi(G'_n \cup G_{n+1}) \\ &= \Pi(G'_n) + \Pi(G_{n+1}) - \Pi(G'_n \cap G_{n+1})\end{aligned}$$

from (2.12) for Π on $\mathcal{G}$ and using the induction hypothesis, (2.22) and the monotonicity of Π^*, we get the upper bound

$$\begin{aligned}&\leq \left(\Pi^*(A_n) + \epsilon \sum_{i=1}^{n} 2^{-i}\right) + \left(\Pi^*(A_{n+1}) + \frac{\epsilon}{2^n}\right) \\ &\qquad - \Pi^*(A_n) \\ &= \epsilon \sum_{i=1}^{n+1} 2^{-i} + \Pi^*(A_{n+1})\end{aligned}$$

which is (2.23) with n replaced by $n + 1$.

Let $n \to \infty$ in (2.23). Recalling Π^* is monotone on $\mathcal{P}(\Omega)$, Π is monotone on $\mathcal{G}$ and $\mathcal{G}$ is closed under non-decreasing limits, we get

$$\lim_{n\to\infty} \uparrow \Pi^*(A_n) + \epsilon \geq \lim_{n\to\infty} \uparrow \Pi(G'_n) = \Pi(\bigcup_{j=1}^{\infty} G'_j).$$

Since

$$A = \lim_{n\to\infty} \uparrow A_n \subset \bigcup_{j=1}^{\infty} G'_j \in \mathcal{G},$$

we conclude

$$\lim_{n\to\infty} \uparrow \Pi^*(A_n) \geq \Pi^*(A).$$

For a reverse inequality, note that monotonicity gives

$$\Pi^*(A_n) \leq \Pi^*(A)$$

and thus

$$\lim_{n\to\infty} \uparrow \Pi^*(A_n) \leq \Pi^*(A).$$

□

PART 3. We now retract Π^* to a certain subclass $\mathcal{D}$ of $\mathcal{P}(\Omega)$ and show $\Pi^*|_{\mathcal{D}}$ is the desired extension.

We define

$$\mathcal{D} := \{D \in \mathcal{P}(\Omega) : \Pi^*(D) + \Pi^*(D^c) = 1.\}$$

Lemma 2.4.3 *The class $\mathcal{D}$ has the following properties:*

1. *$\mathcal{D}$ is a σ-field.*
2. *$\Pi^*|_{\mathcal{D}}$ is a probability measure on $(\Omega, \mathcal{D})$.*

Proof. We first show $\mathcal{D}$ is a field. Obviously $\Omega \in \mathcal{D}$ since $\Pi^*(\Omega) = 1$ and $\Pi^*(\emptyset) = 0$. To see $\mathcal{D}$ is closed under complementation is easy: If $D \in \mathcal{D}$, then

$$\Pi^*(D) + \Pi^*(D^c) = 1$$

and the same holds for D^c.

Next, we show $\mathcal{D}$ is closed under finite unions and finite intersections. If $D_1, D_2 \in \mathcal{D}$, then from (2.20)

$$\Pi^*(D_1 \cup D_2) + \Pi^*(D_1 \cap D_2) \leq \Pi^*(D_1) + \Pi^*(D_2) \tag{2.25}$$

$$\Pi^*((D_1 \cup D_2)^c) + \Pi^*((D_1 \cap D_2)^c) \leq \Pi^*(D_1^c) + \Pi^*(D_2^c). \tag{2.26}$$

Add the two inequalities (2.25) and (2.26) to get

$$\Pi^*(D_1 \cup D_2) + \Pi^*((D_1 \cup D_2)^c) + \Pi^*(D_1 \cap D_2) + \Pi^*((D_1 \cap D_2)^c) \leq 2 \tag{2.27}$$

where we used $D_i \in \mathcal{D}$, $i = 1, 2$ on the right side. From (2.21), the left side of (2.27) is ≥ 2, so equality prevails in (2.27). Again using (2.21), we see

$$\Pi^*(D_1 \cup D_2) + \Pi^*((D_1 \cup D_2)^c) = 1$$

$$\Pi^*(D_1 \cap D_2) + \Pi^*((D_1 \cap D_2)^c) = 1.$$

Thus $D_1 \cup D_2$, $D_1 \cap D_2 \in \mathcal{D}$ and $\mathcal{D}$ is a field. Also, equality must prevail in (2.25) and (2.26) (else it would fail in (2.27)). This shows that Π^* is finitely additive on $\mathcal{D}$.

Now it remains to show that $\mathcal{D}$ is a σ-field and Π^* is σ-additive on $\mathcal{D}$. Since $\mathcal{D}$ is a field, to show it is a σ-field, it suffices by Exercise 41 of Chapter 1 to show that $\mathcal{D}$ is a monotone class. Since $\mathcal{D}$ is closed under complementation, it is enough to show that $D_n \in \mathcal{D}$, $D_n \uparrow D$ implies $D \in \mathcal{D}$. However, $D_n \uparrow D$ implies, since Π^* is monotone and sequentially monotone continuous, that

$$\lim_{n\to\infty} \uparrow \Pi^*(D_n) = \Pi^*(\bigcup_{n=1}^{\infty} D_n) = \Pi^*(D).$$

Also, for any $m \geq 1$,

$$\Pi^*((\bigcup_{n=1}^{\infty} D_n)^c) = \Pi^*(\bigcap_{n=1}^{\infty} D_n^c) \leq \Pi^*(D_m^c)$$

and therefore, from (2.21)

$$1 \leq \Pi^*(\bigcup_{n=1}^{\infty} D_n) + \Pi^*((\bigcup_{n=1}^{\infty} D_n)^c) \leq \lim_{n\to\infty} \Pi^*(D_n) + \Pi^*(D_m^c) \tag{2.28}$$

and letting $m \to \infty$, we get using $D_n \in \mathcal{D}$

$$\begin{aligned} 1 &\leq \lim_{n\to\infty} \Pi^*(D_n) + \lim_{m\to\infty} \Pi^*(D_m^c) \\ &= \lim_{n\to\infty} \left(\Pi^*(D_n) + \Pi^*(D_n^c)\right) = 1, \end{aligned}$$

and so equality prevails in (2.28). Thus, $D_n \uparrow D$ and $D_n \in \mathcal{D}$ imply $D \in \mathcal{D}$ and $\mathcal{D}$ is both an algebra and a monotone class and hence is a σ-algebra.

Finally, we show $\Pi^*|_{\mathcal{D}}$ is σ-additive. If $\{D_n\}$ is a sequence of disjoint sets in $\mathcal{D}$, then because Π^* is continuous with respect to non-decreasing sequences and $\mathcal{D}$ is a field

$$\begin{aligned} \Pi^*(\sum_{i=1}^{\infty} D_i) &= \Pi^*(\lim_{n\to\infty} \sum_{i=1}^{n} D_i) \\ &= \lim_{n\to\infty} \Pi^*(\sum_{i=1}^{n} D_i) \end{aligned}$$

and because Π^* is finitely additive on $\mathcal{D}$, this is

$$= \lim_{n\to\infty} \sum_{i=1}^{n} \Pi^*(D_i) = \sum_{i=1}^{\infty} \Pi^*(D_i),$$

as desired.

Since $\mathcal{D}$ is a σ-field and $\mathcal{D} \supset \mathcal{A}$, $\mathcal{D} \supset \sigma(\mathcal{A})$. The restriction $\Pi^*|_{\sigma(\mathcal{A})}$ is the desired extension of P on $\mathcal{A}$ to a probability measure on $\sigma(\mathcal{A})$. The extension from $\mathcal{A}$ to $\sigma(\mathcal{A})$ must be unique because of Corollary 2.2.1 to Dynkin's theorem. □

2.5 Measure Constructions

In this section we give two related constructions of probability spaces. The first discussion shows how to construct Lebesgue measure on $(0, 1]$ and the second shows how to construct a probability on $\mathbb{R}$ with given distribution function F.

2.5.1 *Lebesgue Measure on* $(0, 1]$

Suppose

$$\begin{aligned}\Omega &= (0,1],\\ \mathcal{B} &= \mathcal{B}((0,1]),\\ \mathcal{S} &= \{(a,b] : 0 \le a \le b \le 1\}.\end{aligned}$$

Define on $\mathcal{S}$ the function $\lambda : \mathcal{S} \mapsto [0, 1]$ by

$$\lambda(\emptyset) = 0, \quad \lambda(a, b] = b - a.$$

With a view to applying Extension Theorem 2.4.3, note that $\lambda(A) \ge 0$. To show that λ has unique extension we need to show that λ is σ-additive.

We first show that λ is finitely additive on $\mathcal{S}$. Let $(a, b] \in \mathcal{S}$ and suppose

$$(a,b] = \bigcup_{i=1}^{k} (a_i, b_i],$$

where the intervals on the right side are disjoint. Assuming the intervals have been indexed conveniently, we have

$$a_1 = a, b_k = b, b_i = a_{i+1}, \ i = 1, \dots, k-1.$$

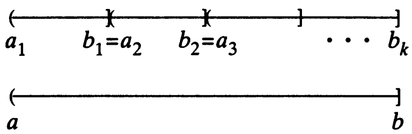

FIGURE 2.3 Abutting Intervals

Then $\lambda(a, b] = b - a$ and

$$\begin{aligned}\sum_{i=1}^{k} \lambda(a_i, b_i] &= \sum_{i=1}^{k} (b_i - a_i)\\ &= b_1 - a_1 + b_2 - a_2 + \cdots + b_k - a_k\\ &= b_k - a_1 = b - a.\end{aligned}$$

This shows λ is finitely additive.

We now show λ is σ-additive. Care must be taken since this involves an infinite number of sets and in fact a compactness argument is employed to cope with the infinities.

Let

$$(a, b] = \bigcup_{i=1}^{\infty}(a_i, b_i]$$

and we first prove that

$$b - a \le \sum_{i=1}^{\infty}(b_i - a_i). \tag{2.29}$$

Pick $\varepsilon < b - a$ and observe

$$[a + \varepsilon, b] \subset \bigcup_{i=1}^{\infty}\left(a_i, b_i + \frac{\varepsilon}{2^i}\right). \tag{2.30}$$

The set on the left side of (2.30) is compact and the right side of (2.30) gives an open cover, so that by compactness, there is a finite subcover. Thus there exists some integer N such that

$$[a + \varepsilon, b] \subset \bigcup_{i=1}^{N}\left(a_i, b_i + \frac{\varepsilon}{2^i}\right). \tag{2.31}$$

It suffices to prove

$$b - a - \varepsilon \le \sum_{1}^{N}\left(b_i - a_i + \frac{\varepsilon}{2^i}\right) \tag{2.32}$$

since then we would have

$$b - a - \varepsilon \le \sum_{1}^{N}\left(b_i - a_i + \frac{\varepsilon}{2^i}\right) \le \sum_{1}^{\infty}(b_i - a_i) + \varepsilon; \tag{2.33}$$

that is,

$$b - a \le \sum_{1}^{\infty}(b_i - a_i) + 2\varepsilon. \tag{2.34}$$

Since ε can be arbitrarily small

$$b - a \le \sum_{1}^{\infty}(b_i - a_i)$$

as desired.

Rephrasing relations (2.31) and (2.32) slightly, we need to prove that

$$[a, b] \subset \bigcup_{1}^{N} (a_i, b_i) \tag{2.35}$$

implies

$$b - a \leq \sum_{1}^{N} (b_i - a_i). \tag{2.36}$$

We prove this by induction. First note that the assertion that (2.35) implies (2.36) is true for $N = 1$. Now we make the induction hypothesis that whenever relation (2.35) holds for $N - 1$, it follows that relation (2.36) holds for $N - 1$. We now must show that (2.35) implies (2.36) for N.

Suppose $a_N = \vee_1^N a_i$, and

$$a_N < b \leq b_N, \tag{2.37}$$

with similar argument if (2.37) fails. Suppose relation (2.35) holds. We consider two cases:

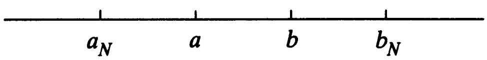

FIGURE 2.4 Case 1

FIGURE 2.5 Case 2

CASE 1: Suppose $a_N \leq a$ Then

$$b - a \leq b_N - a_N \leq \sum_{1}^{N} (b_i - a_i).$$

CASE 2: Suppose $a_N > a$. Then if (2.35) holds

$$[a, a_N] \subset \bigcup_{1}^{N-1} (a_i, b_i)$$

so by the induction hypothesis

$$a_N - a \leq \sum_{i=1}^{N-1} (b_i - a_i)$$

so

$$\begin{aligned}
b-a =& b-a_N+a_N-a \\
\leq& b-a_N+\sum_{i=1}^{N-1}(b_i-a_i) \\
\leq& b_N-a_N+\sum_{i=1}^{N-1}(b_i-a_i) \\
=& \sum_{i=1}^{N}(b_i-a_i)
\end{aligned}$$

which is relation (2.36). This verifies (2.29).

We now obtain a reverse inequality complementary to (2.29). We claim that if $(a,b]=\sum_{i=1}^{\infty}(a_i,b_i]$, then for every n,

$$\lambda((a,b])=b-a\geq\sum_{i=1}^{n}\lambda((a_i,b_i])=\sum_{i=1}^{n}(b_i-a_i). \tag{2.38}$$

This is easily verified since we know λ is finitely additive on $\mathcal{S}$. For any n, $\cup_{i=1}^{n}(a_i,b_i]$ is a finite union of disjoint intervals and so is

$$(a,b]\setminus\bigcup_{i=1}^{n}(a_i,b_i]=:\bigcup_{j=1}^{m}I_j.$$

So by finite additivity

$$\lambda((a,b])=\lambda(\bigcup_{i=1}^{n}(a_i,b_i]\cup\bigcup_{j=1}^{m}I_j),$$

which by finite additivity is

$$\begin{aligned}
=&\sum_{i=1}^{n}\lambda((a_i,b_i])+\sum_{j=1}^{m}\lambda(I_j) \\
\geq&\sum_{i=1}^{n}\lambda((a_i,b_i]).
\end{aligned}$$

Let $n\to\infty$ to achieve

$$\lambda((a,b])\geq\sum_{i=1}^{\infty}\lambda((a_i,b_i]).$$

This plus (2.29) shows λ is σ-additive on $\mathcal{S}$. $\square$

2.5.2 *Construction of a Probability Measure on $\mathbb{R}$ with Given Distribution Function $F(x)$*

Given Lebesgue measure λ constructed in Section 2.5.1 and a distribution function $F(x)$, we construct a probability measure on $\mathbb{R}$, P_F, such that

$$P_F((-\infty, x]) = F(x).$$

Define the left continuous inverse of F as

$$F^{\leftarrow}(y) = \inf\{s : F(s) \geq y\}, \quad 0 < y \leq 1 \tag{2.39}$$

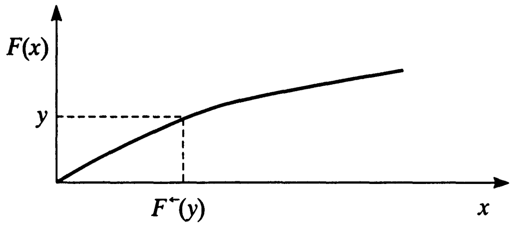

FIGURE 2.6

and define

$$A(y) := \{s : F(s) \geq y\}.$$

Here are the important properties of $A(y)$.

(a) The set $A(y)$ is closed. If $s_n \in A(y)$, and $s_n \downarrow s$, then by right continuity

$$y \leq F(s_n) \downarrow F(s),$$

so $F(s) \geq y$ and $s \in A(y)$. If $s_n \uparrow s$ and $s_n \in A(y)$, then

$$y \leq F(s_n) \uparrow F(s-) \leq F(s)$$

and $y \leq F(s)$ implies $s \in A(y)$.

(b) Since $A(y)$ closed,

$$\inf A(y) \in A(y);$$

that is,

$$F(F^{\leftarrow}(y)) \geq y.$$

(c) Consequently,

$$F^{\leftarrow}(y) > t \text{ iff } y > F(t)$$

or equivalently

$$F^{\leftarrow}(y) \leq t \text{ iff } y \leq F(t).$$

The last property is proved as follows. If $t < F^{\leftarrow}(y) = \inf A(y)$, then $t \notin A(y)$, so that $F(t) < y$. Conversely, if $F^{\leftarrow}(y) \le t$, then $t \in A(y)$ and $F(t) \ge y$.

Now define for $A \subset \mathbb{R}$

$$\xi_F(A) = \{x \in (0,1] : F^{\leftarrow}(x) \in A\}.$$

If A is a Borel subset of $\mathbb{R}$, then $\xi_F(A)$ is a Borel subset of $(0, 1]$.

Lemma 2.5.1 *If $A \in \mathcal{B}(\mathbb{R})$, then $\xi_F(A) \in \mathcal{B}((0,1])$.*

Proof. Define

$$\mathcal{G} = \{A \subset \mathbb{R} : \xi_F(A) \in \mathcal{B}((0,1])\}.$$

$\mathcal{G}$ contains finite intervals of the form $(a, b] \subset \mathbb{R}$ since from Property (c) of $F^{\leftarrow}$

$$\begin{aligned}\xi_F((a,b]) &= \{x \in (0,1] : F^{\leftarrow}(x) \in (a,b]\}\\ &= \{x \in (0,1] : a < F^{\leftarrow}(x) \le b\}\\ &= \{x \in (0,1] : F(a) < x \le F(b)\}\\ &= (F(a), F(b)] \in \mathcal{B}((0,1]).\end{aligned}$$

Also $\mathcal{G}$ is a σ-field since we easily verify the σ-field postulates:

(i) We have

$$\mathbb{R} \in \mathcal{G}$$

since $\xi_F(\mathbb{R}) = (0, 1]$.

(ii) We have that $A \in \mathcal{G}$ implies $A^c \in \mathcal{G}$ since

$$\begin{aligned}\xi_F(A^c) &= \{x \in (0,1] : F^{\leftarrow}(x) \in A^c\}\\ &= \{x \in (0,1] : F^{\leftarrow}(x) \in A\}^c = (\xi_F(A))^c.\end{aligned}$$

(iii) $\mathcal{G}$ is closed under countable unions since if $A_n \in \mathcal{G}$, then

$$\xi_F(\bigcup_n A_n) = \bigcup_n \xi_F(A_n)$$

and therefore

$$\bigcup_n A_n \in \mathcal{G}.$$

So $\mathcal{G}$ contains intervals and $\mathcal{G}$ is a σ-field and therefore

$$\mathcal{G} \supset \mathcal{B}(\text{ intervals }) = \mathcal{B}(\mathbb{R}).$$ □

We now can make our definition of P_F. We define

$$P_F(A) = \lambda(\xi_F(A)),$$

where λ is Lebesgue measure on $(0, 1]$. It is easy to check that P_F is a probability measure. To compute its distribution function and check that it is F, note that

$$\begin{aligned}P_F(-\infty, x] &= \lambda(\xi_F(-\infty, x]) = \lambda\{y \in (0,1] : F^{\leftarrow}(y) \le x\}\\ &= \lambda\{y \in (0,1] : y \le F(x)\}\\ &= \lambda((0, F(x)]) = F(x).\end{aligned}$$ □

2.6 Exercises

1. Let Ω be a non-empty set. Let $\mathcal{F}_0$ be the collection of all subsets such that either A or A^c is finite.

   (a) Show that $\mathcal{F}_0$ is a field.

   Define for $E \in \mathcal{F}_0$ the set function P by

$$P(E) = \begin{cases} 0, & \text{if } E \text{ is finite,} \\ 1, & \text{if } E^c \text{ is finite.} \end{cases}$$

   (b) If Ω is countably infinite, show P is finitely additive but not σ-additive.

   (c) If Ω is uncountable, show P is σ-additive on $\mathcal{F}_0$.

2. Let $\mathcal{A}$ be the smallest field over the π-system $\mathcal{P}$. Use the inclusion-exclusion formula (2.2) to show that probability measures agreeing on $\mathcal{P}$ must agree also on $\mathcal{A}$.

   Hint: Use Exercise 20 of Chapter 1.

3. Let $(\Omega, \mathcal{B}, P)$ be a probability space. Show for events $B_i \subset A_i$ the following generalization of subadditivity:

$$P(\cup_i A_i) - P(\cup_i B_i) \leq \sum_i (P(A_i) - P(B_i)).$$

4. Review Exercise 34 in Chapter 1 to see how to extend a σ-field. Suppose P is a probability measure on a σ-field $\mathcal{B}$ and suppose $A \notin \mathcal{B}$. Let

$$\mathcal{B}_1 = \sigma(\mathcal{B}, A)$$

   and show that P has an extension to a probability measure P_1 on $\mathcal{B}_1$. (Do this without appealing directly to the Combo Extension Theorem 2.4.3.)

5. Let P be a probability measure on $\mathcal{B}(\mathbb{R})$. For any $B \in \mathcal{B}(\mathbb{R})$ and any $\epsilon > 0$, there exists a finite union of intervals A such that

$$P(A \triangle B) < \epsilon.$$

   Hint: Define

$$\mathcal{G} := \{B \in \mathcal{B}(\mathbb{R}) : \forall \epsilon > 0, \text{ there exists a finite union of intervals } A_\epsilon \text{ such that } P(A \triangle B) < \epsilon\}.$$

6. Say events $A_1, A_2, \ldots$ are *almost disjoint* if

$$P(A_i \cap A_j) = 0, \quad i \neq j.$$

   Show for such events

$$P(\bigcup_{j=1}^{\infty} A_j) = \sum_{j=1}^{\infty} P(A_j).$$

7. **Coupon collecting.** Suppose there are N different types of coupons available when buying cereal; each box contains one coupon and the collector is seeking to collect one of each in order to win a prize. After buying n boxes, what is the probability p_n that the collector has at least one of each type? (Consider sampling with replacement from a population of N distinct elements. The sample size is $n > N$. Use inclusion–exclusion formula (2.2).)

8. We know that $P_1 = P_2$ on $\mathcal{B}$ if $P_1 = P_2$ on $\mathcal{C}$, provided that $\mathcal{C}$ generates $\mathcal{B}$ and is a π-system. Show this last property cannot be omitted. For example, consider $\Omega = \{a, b, c, d\}$ with

$$P_1(\{a\}) = P_1(\{d\}) = P_2(\{b\}) = P_2(\{c\}) = \frac{1}{6}$$

and

$$P_1(\{b\}) = P_1(\{c\}) = P_2(\{a\}) = P_2(\{d\}) = \frac{1}{3}.$$

Set

$$\mathcal{C} = \{\{a, b\}, \{d, c\}, \{a, c\}, \{b, d\}\}.$$

9. Background: Call two sets $A_1, A_2 \in \mathcal{B}$ *equivalent* if $P(A_1 \triangle A_2) = 0$. For a set $A \in \mathcal{B}$, define the equivalence class

$$A^{\#} = \{B \in \mathcal{B} : P(B \triangle A) = 0\}.$$

This decomposes $\mathcal{B}$ into equivalences classes. Write

$$P^{\#}(A^{\#}) = P(A), \qquad \forall A \in A^{\#}.$$

In practice we drop #s; that is identify the equivalence classes with the members.

An *atom* in a probability space $(\Omega, \mathcal{B}, P)$ is defined as (the equivalence class of) a set $A \in \mathcal{B}$ such that $P(A) > 0$, and if $B \subset A$ and $B \in \mathcal{B}$, then $P(B) = 0$, or $P(A \setminus B) = 0$. Furthermore the probability space is called *non-atomic* if there are no atoms; that is, $A \in \mathcal{B}$ and $P(A) > 0$ imply that there exists a $B \in \mathcal{B}$ such that $B \subset A$ and $0 < P(B) < P(A)$.

(a) If $\Omega = R$, and P is determined by a distribution function $F(x)$, show that the atoms are $\{x : F(x) - F(x-) > 0\}$.

(b) If $(\Omega, \mathcal{B}, P) = ((0, 1], \mathcal{B}((0, 1]), \lambda)$, where λ is Lebesgue measure, then the probability space is non-atomic.

(c) Show that two distinct atoms have intersection which is the empty set. (The sets A, B are distinct means $P(A \triangle B) > 0$. The exercise then requires showing $P(AB \triangle \emptyset) = 0$.)

(d) A probability space contains at most countably many atoms. (Hint: What is the maximum number of atoms that the space can contain that have probability at least $1/n$?)

(e) If a probability space $(\Omega, \mathcal{B}, P)$ contains no atoms, then for every $a \in (0, 1]$ there exists at least one set $A \in \mathcal{B}$ such that $P(A) = a$. (One way of doing this uses Zorn's lemma.)

(f) For every probability space $(\Omega, \mathcal{B}, P)$ and any $\epsilon > 0$, there exists a finite partition of Ω by $\mathcal{B}$ sets, each of whose elements either has probability $\leq \epsilon$ or is an atom with probability $> \epsilon$.

(g) **Metric space:** On the set of equivalence classes, define

$$d(A_1^{\#}, A_2^{\#}) = P(A_1 \Delta A_2)$$

where $A_i \in A_i^{\#}$ for $i = 1, 2$. Show d is a metric on the set of equivalence classes. Verify

$$|P(A_1) - P(A_2)| \leq P(A_1 \Delta A_2)$$

so that $P^{\#}$ is uniformly continuous on the set of equivalence classes. P is σ-additive is equivalent to

$$\mathcal{B} \ni A_n \downarrow \emptyset \text{ implies } d(A_n^{\#}, \emptyset^{\#}) \to 0.$$

10. Two events A, B on the probability space $(\Omega, \mathcal{B}, P)$ are equivalent (see Exercise 9) if

$$P(A \cap B) = P(A) \vee P(B).$$

11. Suppose $\{B_n, n \geq 1\}$ are events with $P(B_n) = 1$ for all n. Show

$$P(\bigcap_{n=1}^{\infty} B_n) = 1.$$

12. Suppose $\mathcal{C}$ is a class of subsets of Ω and suppose $B \subset \Omega$ satisfies $B \in \sigma(\mathcal{C})$. Show that there exists a countable class $\mathcal{C}_B \subset \mathcal{C}$ such that $B \in \sigma(\mathcal{C}_B)$.

Hint: Define

$$\mathcal{G} := \{B \subset \Omega : \exists \text{ countable } \mathcal{C}_B \subset \mathcal{C} \text{ such that } B \in \sigma(\mathcal{C}_B)\}.$$

Show that $\mathcal{G}$ is a σ-field that contains $\mathcal{C}$.

13. If $\{B_k\}$ are events such that

$$\sum_{k=1}^{n} P(B_k) > n - 1,$$

then

$$P(\bigcap_{k=1}^{n} B_k) > 0.$$

14. If F is a distribution function, then F has at most countably many discontinuities.

15. If $\mathcal{S}_1$ and $\mathcal{S}_2$ are two semialgebras of subsets of Ω, show that the class

$$\mathcal{S}_1\mathcal{S}_2 := \{A_1A_2 : A_1 \in \mathcal{S}_1, A_2 \in \mathcal{S}_2\}$$

is again a semialgebra of subsets of Ω. The field (σ-field) generated by $\mathcal{S}_1\mathcal{S}_2$ is identical with that generated by $\mathcal{S}_1 \cup \mathcal{S}_2$.

16. Suppose $\mathcal{B}$ is a σ-field of subsets of Ω and suppose $Q : \mathcal{B} \mapsto [0, 1]$ is a set function satisfying

    (a) Q is finitely additive on $\mathcal{B}$.

    (b) $0 \le Q(A) \le 1$ for all $A \in \mathcal{B}$ and $Q(\Omega) = 1$.

    (c) If $A_i \in \mathcal{B}$ are disjoint and $\sum_{i=1}^{\infty} A_i = \Omega$, then $\sum_{i=1}^{\infty} Q(A_i) = 1$.

    Show Q is a probability measure; that is, show Q is σ-additive.

17. For a distribution function $F(x)$, define

$$\begin{aligned} F_l^{\leftarrow}(y) &= \inf\{t : F(t) \ge y\} \\ F_r^{\leftarrow}(y) &= \inf\{t : F(t) > y\}. \end{aligned}$$

We know $F_l^{\leftarrow}(y)$ is left-continuous. Show $F_r^{\leftarrow}(y)$ is right continuous and show

$$\lambda\{u \in (0, 1] : F_l^{\leftarrow}(u) \ne F_r^{\leftarrow}(u)\} = 0,$$

where, as usual, λ is Lebesgue measure. Does it matter which inverse we use?

18. Let A, B, C be disjoint events in a probability space with

$$P(A) = .6, \quad P(B) = .3, \quad P(C) = .1.$$

Calculate the probabilities of every event in $\sigma(A, B, C)$.

19. **Completion.** Let $(\Omega, \mathcal{B}, P)$ be a probability space. Call a set N *null* if $N \in \mathcal{B}$ and $P(N) = 0$. Call a set $B \subset \Omega$ *negligible* if there exists a null set N such that $B \subset N$. Notice that for B to be negligible, it is not required that B be measurable. Denote the set of all negligible subsets by $\mathcal{N}$. Call $\mathcal{B}$ *complete* (with respect to P) if every negligible set is null.

    What if $\mathcal{B}$ is not complete? Define

$$\mathcal{B}^* := \{A \cup M : A \in \mathcal{B}, M \in \mathcal{N}\}.$$

    (a) Show $\mathcal{B}^*$ is a σ-field.

(b) If $A_i \in \mathcal{B}$ and $M_i \in \mathcal{N}$ for $i = 1, 2$ and

$$A_1 \cup M_1 = A_2 \cup M_2,$$

then $P(A_1) = P(A_2)$.

(c) Define $P^* : \mathcal{B}^* \mapsto [0, 1]$ by

$$P^*(A \cup M) = P(A), \quad A \in \mathcal{B},\ M \in \mathcal{N}.$$

Show P^* is an extension of P to $\mathcal{B}^*$.

(d) If $B \subset \Omega$ and $A_i \in \mathcal{B}, i = 1, 2$ and $A_1 \subset B \subset A_2$ and $P(A_2 \setminus A_1) = 0$, then show $B \in \mathcal{B}^*$.

(e) Show $\mathcal{B}^*$ is complete. Thus every σ-field has a completion.

(f) Suppose $\Omega = \mathbb{R}$ and $\mathcal{B} = \mathcal{B}(\mathbb{R})$. Let $p_k \geq 0$, $\sum_k p_k = 1$. Let $\{a_k\}$ be any sequence in $\mathbb{R}$. Define P by

$$P(\{a_k\}) = p_k, \quad P(A) = \sum_{a_k \in A} p_k, \quad A \in \mathcal{B}.$$

What is the completion of $\mathcal{B}$?

(g) Say that the probability space $(\Omega, \mathcal{B}, P)$ has a complete extension $(\Omega, \mathcal{B}_1, P_1)$ if $\mathcal{B} \subset \mathcal{B}_1$ and $P_1|_{\mathcal{B}} = P$. The previous problem (c) showed that every probability space has a complete extension. However, this extension may not be unique. Suppose that $(\Omega, \mathcal{B}_2, P_2)$ is a second complete extension of $(\Omega, \mathcal{B}, P)$. Show P_1 and P_2 may not agree on $\mathcal{B}_1 \cap \mathcal{B}_2$. (It should be enough to suppose Ω has a small number of points.)

(h) Is there a *minimal* extension?

20. In $(0, 1]$, let $\mathcal{B}$ be the class of sets that either (a) are of the first category or (b) have complement of the first category. Show that $\mathcal{B}$ is a σ-field. For $A \in \mathcal{B}$, define $P(A)$ to be 0 in case (a) and 1 in case (b). Is P σ-additive?

21. Let $\mathcal{A}$ be a field of subsets of Ω and let μ be a *finitely* additive probability measure on $\mathcal{A}$. (This requires $\mu(\Omega) = 1$.)

    If $\mathcal{A}_n \in \mathcal{A}$ and $\mathcal{A}_n \downarrow \emptyset$, is it the case that $\mu(\mathcal{A}_n) \downarrow 0$? (Hint: Review Problem 2.6.1 with $A_n = \{n, n+1, \dots\}$.)

22. Suppose $F(x)$ is a continuous distribution function on $\mathbb{R}$. Show F is *uniformly* continuous.

23. **Multidimensional distribution functions.** For $\mathbf{a}, \mathbf{b}, \mathbf{x} \in \mathcal{B}(\mathbb{R}^k)$ write

$$\mathbf{a} \le \mathbf{b} \text{ iff } a_i \le b_i,\ i = 1, \dots, k;$$
$$(-\infty, \mathbf{x}] = \{\mathbf{u} \in \mathcal{B}(\mathbb{R}^k) : \mathbf{u} \le \mathbf{x}\}$$
$$(\mathbf{a}, \mathbf{b}] = \{\mathbf{u} \in \mathcal{B}(\mathbb{R}^k) : \mathbf{a} < \mathbf{u} \le \mathbf{b}\}.$$

Let P be a probability measure on $\mathcal{B}(\mathbb{R}^k)$ and define for $\mathbf{x} \in \mathbb{R}^k$

$$F(\mathbf{x}) = P((-\infty, \mathbf{x}]).$$

Let $\mathcal{S}_k$ be the semialgebra of k-dimensional rectangles in $\mathbb{R}^k$.

(a) If $\mathbf{a} \le \mathbf{b}$, show the rectangle $I_k := (\mathbf{a}, \mathbf{b}]$ can be written as

$$I_k = (-\infty, \mathbf{b}] \setminus \Big((-\infty, (a_1, b_2, \dots, b_k)] \cup (-\infty, (b_1, a_2, \dots, b_k)] \cup \cdots \cup (-\infty, (b_1, b_2, \dots, a_k)] \Big) \tag{2.40}$$

where the union is indexed by the vertices of the rectangle other than $\mathbf{b}$.

(b) Show

$$\mathcal{B}(\mathbb{R}^k) = \sigma((-\infty, \mathbf{x}], \mathbf{x} \in \mathbb{R}^k).$$

(c) Check that $\{(-\infty, \mathbf{x}], \mathbf{x} \in \mathbb{R}^k\}$ is a π-system.

(d) Show P is determined by $F(\mathbf{x}), \mathbf{x} \in \mathbb{R}^k$.

(e) Show F satisfies the following properties:
(1) If $x_i \to \infty, i = 1, \dots, k$, then $F(\mathbf{x}) \to 1$.
(2) If for some $i \in \{1, \dots, k\}$ $x_i \to -\infty$, then $F(\mathbf{x}) \to 0$.
(3) For $\mathcal{S}_k \ni I_k = (\mathbf{a}, \mathbf{b}]$, use the inclusion-exclusion formula (2.2) to show

$$P(I_k) = \Delta_{I_k} F.$$

The symbol on the right is explained as follows. Let $\mathcal{V}$ be the vertices of I_k so that

$$\mathcal{V} = \{(x_1, \dots, x_i) : x_i = a_i \text{ or } b_i,\ i = 1, \dots, k\}.$$

Define for $x \in \mathcal{V}$

$$\operatorname{sgn}(\mathbf{x}) = \begin{cases} +1, & \text{if } \operatorname{card}\{i : x_i = a_i\} \text{ is even.} \\ -1, & \text{if } \operatorname{card}\{i : x_i = a_i\} \text{ is odd.} \end{cases}$$

Then

$$\Delta_{I_k} F = \sum_{\mathbf{x} \in \mathcal{V}} \operatorname{sgn}(\mathbf{x}) F(\mathbf{x}).$$

(f) Show F is continuous from above:

$$\lim_{\mathbf{a} \leq \mathbf{x} \downarrow \mathbf{a}} F(\mathbf{x}) = F(\mathbf{a}).$$

(g) Call $F : \mathbb{R}^k \mapsto [0, 1]$ a multivariate distribution function if properties (1), (2) hold as well as F is continuous from above and $\Delta_{I_k} F \geq 0$. Show any multivariate distribution function determines a unique probability measure P on $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$. (Use the extension theorem.)

24. Suppose λ_2 is the uniform distribution on the unit square $[0, 1]^2$ defined by its distribution function

$$\lambda_2([0, \theta_1] \times [0, \theta_2]) = \theta_1 \theta_2, \quad (\theta_1, \theta_2) \in [0, 1]^2.$$

(a) Prove that λ_2 assigns 0 probability to the boundary of $[0, 1]^2$.

(b) Calculate

$$\lambda_2\{(\theta_1, \theta_2) \in [0, 1]^2 : \theta_1 \wedge \theta_2 > \frac{2}{3}\}.$$

(c) Calculate

$$\lambda_2\{(\theta_1, \theta_2) \in [0, 1]^2 : \theta_1 \wedge \theta_2 \leq x, \theta_1 \wedge \theta_2 \leq y\}.$$

25. In the game of bridge 52 distinguishable cards constituting 4 equal suits are distributed at random among 4 players. What is the probability that at least one player has a complete suit?

26. If $A_1, \dots, A_n$ are events, define

$$\begin{aligned} S_1 &= \sum_{i=1}^{n} P(A_i) \\ S_2 &= \sum_{1 \leq i < j \leq n} P(A_i A_j) \\ S_3 &= \sum_{1 \leq i < j < k \leq n} P(A_i A_j A_k) \\ &\vdots \qquad \vdots \end{aligned}$$

and so on.

(a) Show the probability $(1 \leq m \leq n)$

$$p(m) = P[\sum_{i=1}^{n} 1_{A_i} = m]$$

of exactly m of the events occurring is

$$p(m) = S_m - \binom{m+1}{m} S_{m+1} + \binom{m+2}{m} S_{m+2} - + \cdots \pm \binom{n}{m} S_n. \tag{2.41}$$

Verify that the inclusion-exclusion formula (2.2) is a special case of (2.41).

(b) Referring to Example 2.1.2, compute the probability of exactly m coincidences.

27. **Regular measures.** Consider the probability space $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k), P)$. A Borel set A is *regular* if

$$P(A) = \inf\{P(G) : G \supset A, G \text{ open,}\}$$

and

$$P(A) = \sup\{P(F) : F \subset A, F \text{ closed.}\}$$

P is *regular* if all Borel sets are regular. Define $\mathcal{C}$ to be the collection of regular sets.

(a) Show $\mathbb{R}^k \in \mathcal{C}$, $\emptyset \in \mathcal{C}$.

(b) Show $\mathcal{C}$ is closed under complements and countable unions.

(c) Let $\mathcal{F}(\mathbb{R}^k)$ be the closed subsets of $\mathbb{R}^k$. Show

$$\mathcal{F}(\mathbb{R}^k) \subset \mathcal{C}.$$

(d) Show $\mathcal{B}(\mathbb{R}^k) \subset \mathcal{C}$; that is, show regularity.

(e) For any Borel set A

$$P(A) = \sup\{P(K) : K \subset A, K \text{ compact.}\}$$

http://www.springer.com/978-0-8176-8408-2

A Probability Path
Resnick, S.I.
2014, XIV, 453 p. 11 illus., Softcover
ISBN: 978-0-8176-8408-2
A product of Birkhäuser Basel