MA 528 Measure Theoretic Probability Notes

Hubert J. Farnsworth

September 3, 2021

Chapter 2 Axioms of Probability

Notes

Let Ω be an abstract space and 2^Ω the power set of Ω. Let $\mathcal{A} \subset 2^\Omega$.

Definition 2.1

$\mathcal{A}$ is an algebra if it satisfies (1), (2), and (3) below. $\mathcal{A}$ is a σ-algebra if it satisfies (1), (2), and (4) below.

1. $\emptyset \in \mathcal{A}$ and $\Omega \in \mathcal{A}$
2. If $A \in \mathcal{A}$, then $A^c \in \mathcal{A}$
3. $\mathcal{A}$ If $A_1, \ldots, A_n \in \mathcal{A}$, then $\bigcup_{i=1}^n A_i \in \mathcal{A}$ and $\bigcap_{i=1}^n A_i \in \mathcal{A}$
4. If the countable sequence $A_1, A_2, \dots \in \mathcal{A}$, then $\bigcup_{i=1}^\infty A_i \in \mathcal{A}$ and $\bigcap_{i=1}^\infty A_i \in \mathcal{A}$

Note If (2) holds then $\emptyset \in \mathcal{A}$ implies $\Omega \in \mathcal{A}$ and $\Omega \in \mathcal{A}$ implies $\emptyset \in \mathcal{A}$. If (1) and (4) are satisfied then (3) is satisfied (every σ-algebra is an algebra).

Definition 2.2

If $\mathcal{C} \subset 2^\Omega$, the σ-algebra generated by $\mathcal{C}$, and written $\sigma(\mathcal{C})$, is the smallest σ-algebra containing $\mathcal{C}$.

Note $\sigma(\mathcal{C})$ always exists. See Exercise 2.2.

Theorem 2.1

The Borel σ-algebra of $\mathbb{R}$, $\mathfrak{B}(\mathbb{R})$, which is the smallest σ-algebra containing the open sets in $\mathbb{R}$ (or equivalently containing the closed sets in $\mathbb{R}$), is generated by intervals of the form $(-\infty, a]$ where $a \in \mathbb{Q}$.

Proof: Let $\mathcal{C}$ denote the set of all open intervals. Since every open set in $\mathbb{R}$ is the countable union of open intervals, $\sigma(\mathcal{C}) = \mathfrak{B}$.

Let $\mathcal{D}$ denote the set of all intervals of the form $(-\infty, a]$, $a \in \mathbb{Q}$. Let $(a, b) \in \mathcal{C}$. Let $(a_n)_{n\geq 1}$ be sequence of rational numbers decreasing strictly to a and $(b_n)_{n\geq 1}$ a sequence of rational numbers increasing strictly to b. Then

$$(a, b) = \cup_{n=1}^\infty (a_n, b_n] = \cup_{n=1}^\infty \left((-\infty, a_n]^c \cap (-\infty, b_n]\right) \implies \mathcal{C} \subset \sigma(\mathcal{D}) \implies \sigma(\mathcal{C}) \subset \sigma(\mathcal{D}).$$

Every $D \in \mathcal{D}$ is a closed set since D^c is open. Since $\mathcal{D}$ is a subset of the set of all closed sets in $\mathbb{R}$, $\sigma(D)$ is contained in the sigma algebra generated by set of closed sets in $\mathbb{R}$. That is, $\sigma(D) \subset \mathfrak{B}$.

$$\mathfrak{B} = \sigma(\mathcal{C}) \subset \sigma(\mathcal{D}) \subset \mathfrak{B} \implies \sigma(\mathcal{D}) = \mathfrak{B}.$$

Definition 2.3

A probability measure defined on a σ-algebra $\mathcal{A}$ is a function $P : \mathcal{A} \to [0, 1]$ that satisfies:

1. $P(\Omega) = 1$
2. For every countable sequence $(A_n)_{n\geq 1}$ of elements of $\mathcal{A}$, pairwise disjoint,

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

Condition (2) is called countable additivity. The number $P(A)$ is called the probability of event A. The more rudimentary property that $A, B \in \mathcal{A}, A \cap B = \emptyset \implies P(A \,\dot\cup\, B) = P(A) + P(B)$ is called additivity. Additivity with respect to two sets (A and B) implies additivity with respect to any finite collection of disjoint sets $(A_1, \ldots, A_m)$.

Theorem 2.2

If P is a probability measure on $(\Omega, \mathcal{A})$, then:

(i) $P(\emptyset) = 0$
(ii) P is additive.

Proof: To prove (i), use (2) of definition 2.3 and the fact that the codomain of P is $[0, 1]$.

$$P(\emptyset) = P\left(\dot{\bigcup_{n=1}^{\infty}} \emptyset\right) = \sum_{n=1}^{\infty} P(\emptyset) = P(\emptyset)\sum_{n=1}^{\infty} 1 \implies P(\emptyset) = 0.$$

To prove (ii), suppose $A, B \in \mathcal{A}, A \cap B = \emptyset$. Let $A_1 = A, A_2 = B$, and let $A_n = \emptyset$ for $n \geq 3$.

$$P(A \cup B) = P\left(\dot{\bigcup_{n=1}^{\infty}} A_n\right) = \sum_{n=1}^{\infty} P(A_n) = P(A) + P(B) + \sum_{n=1}^{\infty} P(\emptyset) = P(A) + P(B).$$

Note It follows from Theorem 2.2 that if $A, C \in \mathcal{A}$ with $A \subset C$ then $P(A) \leq P(C)$. To prove this, let $B = C \backslash A$ so that $A \cap B = \emptyset$ and $P(A) \leq P(A) + P(B) = P(A \,\dot\cup\, B) = P(C)$.

Theorem 2.3

Suppose $P : \mathcal{A} \to [0, 1]$ satisfies (1) of definition 2.3 and P is (finitely) additive. The following are equivalent.

(i) Axiom (2) of definition 2.3.
(ii) $A_n \downarrow \emptyset \implies P(A_n) \downarrow 0$.
(iii) $A_n \downarrow A \implies P(A_n) \downarrow P(A)$.
(iv) $A_n \uparrow \Omega \implies P(A_n) \uparrow 1$.
(v) $A_n \uparrow A \implies P(A_n) \uparrow P(A)$.

Proof:
(iii) $\iff$ (v)
Assume (iii) and suppose $A_n \uparrow A$. Then $A_n^c \downarrow A^c$ so $P(A_n^c) \downarrow P(A^c)$. But then $P(A_n) = (1 - P(A_n^c)) \uparrow (1 - P(A^c)) = P(A)$. Proving the reverse is similar.

(ii) $\iff$ (iv)
Let $A = \Omega$ so that $A^c = \emptyset$ and apply the previous result.

(iv) $\iff$ (v)
Assuming (v) holds, $A_n \uparrow A = \Omega \implies P(A_n) \uparrow P(\Omega) = 1$. Therefore (v) $\implies$ (iv). Now assume (iv) and suppose $A_n \uparrow A$. Define $B_n = A_n \mathbin{\dot\cup} A^c$ so that $B_n \uparrow \Omega$. Since $A_n \cap A^c = \emptyset$ for all n, $P(B_n) = P(A_n) + P(A^c)$ for all n. Since $A_n \subset A_{n+1}$ for each n, $P(A_n) \uparrow P(A)$.

$$1 = \lim P(B_n) = P(A^c) + \lim P(A_n) \implies \lim P(A_n) = 1 - P(A^c) = P(A) \implies P(A_n) \uparrow P(A).$$

(i) $\iff$ (v)
Assume (v) holds and suppose $(A_n)_{n\geq 1}$ are pairwise disjoint. Define $B_n = \dot\bigcup_{p=1}^n A_p$ and $B = \dot\bigcup_{n=1}^\infty A_n$. We have $P(B_n) = \sum_{p=1}^n P(A_p)$ for each n by finite additivity. By (v) $P(B_n) \uparrow P(B)$.

$$P\left(\dot\bigcup_{n=1}^\infty A_n\right) = P(B) = \lim P(B_n) = \lim \sum_{p=1}^n P(A_p) = \sum_{p=1}^\infty P(A_p).$$

Therefore (v) $\implies$ (i). Next assume (i) and suppose $A_n \uparrow A$. Define the sequence of disjoint set $(B_n)_{n\geq 1}$:

$$\begin{aligned} B_1 &= A_1 \\ B_2 &= A_2 \backslash A_1 \\ &\vdots \\ B_n &= A_n \backslash A_{n-1} \end{aligned}$$

Then $\dot\bigcup_{n=1}^\infty B_n = A$ and $P(A) = \sum_{n=1}^\infty B_n$ by (i). Since $A_n \subset A_{n+1}$, $P(A_n) \leq P(A_{n+1})$ for each n.

$$\lim P(A_n) = \lim \sum_{p=1}^n P(B_p) = \sum_{p=1}^\infty P(B_p) = P(A) \implies P(A_n) \uparrow P(A).$$

Theorem 2.4

Let P be a probability measure on $\mathcal{A}$ and A_n a sequence of sets in $\mathcal{A}$ with $A_n \to A$. Then $A \in \mathcal{A}$ and $\lim P(A_n) = P(A)$.

Proof:

$$\begin{aligned} \limsup A_n &:= \cap_{n=1}^\infty \cup_{m\geq n} A_m \\ \liminf A_n &:= \cup_{n=1}^\infty \cap_{m\geq n} A_m \end{aligned}$$

Since $\mathcal{A}$ is a σ-algebra and thus closed under countable union and closed under countable intersection, $\limsup A_n \in \mathcal{A}$ and $\liminf A_n \in \mathcal{A}$.

By hypothesis $A_n \to A$ so $\lim 1_{A_n}(\omega) = 1_A(\omega)$ for each $\omega \in \Omega$. This is equivalent to saying $A = \limsup A_n = \liminf A_n$. Therefore $A \in \mathcal{A}$.

Let $B_n = \cap_{m\geq n} A_n$ and $C_n = \cup_{m\geq n} A_n$. Then $B_n \uparrow A$ and $C_n \downarrow A$ so that $\lim P(B_n) = \lim P(C_n) = P(A)$ by Theorem 2.3. Since $B_n \subset A_n \subset C_n$ for each n, $P(B_n) \leq P(A_n) \leq P(C_n)$ for each n.

$$P(A) = \lim P(B_n) \leq \lim P(A_n) \leq \lim P(C_n) = P(A) \implies \lim P(A_n) = P(A).$$

Exercises

Exercise 2.1

Let Ω be a finite set. Show that the set of all subsets of Ω, 2^Ω, is also finite and that it is a σ-algebra.

Answer:

Claim: If $|\Omega| = n$ for some nonnegative integer n, $|2^\Omega| = 2^{|\Omega|}$.

Proof (Induction): If $\Omega = \emptyset$, $\emptyset$ is the only subset of Ω and $|2^\Omega| = 1 = 2^0 = 2^{|\Omega|}$. Assume the claim holds for a set of cardinality n, $n \geq 0$, and consider the case of $|\Omega| = n + 1$. Select on element $\omega \in \Omega$ and consider all $A \subset \Omega$ such that $\omega \notin A$. By the inductive hypothesis there are 2^n such subsets of Ω. For each of these subsets, we build a new subset of Ω by including ω. In this way we find another 2^n subsets of Ω. Since for any subset A of Ω, either $\omega \in A$ or $\omega \notin A$, conclude that $|2^\Omega| = 2 \cdot 2^n = 2^{n+1} = 2^{|\Omega|}$.

By the claim above, if Ω is a finite set then $|2^\Omega| = 2^{|\Omega|} < +\infty$.

To show that 2^Ω is a σ-algebra, check that 2^Ω satisfies axioms (1), (2), and (4) from Definition 2.1.

1. Since $\emptyset \subseteq \Omega$ and $\Omega \subseteq \Omega$, $\emptyset, \Omega \in 2^\Omega$.
2. Suppose $A \in 2^\Omega$. Then $A \subseteq \Omega$ and $A^c = \{\omega \in \Omega : \omega \notin A\} \subseteq \Omega$. Therefore $A^c \in 2^\Omega$ as well.
4. Suppose $A_1, A_2, \ldots$ is a countable sequence of events in 2^Ω. Since each A_k is a subset of Ω,

$$\bigcup_{k=1}^{\infty} A_k = \{\omega \in \Omega : \omega \in A_k \text{ for some } k\} \subseteq \Omega \implies \bigcup_{k=1}^{\infty} A_k \in 2^\Omega,$$
$$\bigcap_{k=1}^{\infty} A_k = \{\omega \in \Omega : \omega \in A_k \text{ for all } k\} \subseteq \Omega \implies \bigcap_{k=1}^{\infty} A_k \in 2^\Omega.$$

Exercise 2.2

Let $(G_\alpha)_{\alpha\in A}$ be an arbitrary family of σ-algebras defined on an abstract space Ω. Show that $H = \cap_{\alpha\in A} G_\alpha$ is also a σ-algebra.

Answer:

1. Since each G_α is a σ-algebra, $\emptyset, \Omega \in G_\alpha$ for each $\alpha \in A$. Thus $\emptyset, \Omega \in H$.
2. Suppose $A \in H$. Then $A \in G_\alpha$ for each α so that $A^c \in G_\alpha$ for each α. Thus $A^c \in H$.
4. Suppose $A_1, A_2, \ldots$ is a countable sequence of events in H. For each $\alpha \in A$, $A_1, A_2, \ldots$ is a countable sequence of events in G_α. This means

$$\bigcup_{k=1}^{\infty} A_k, \bigcap_{k=1}^{\infty} A_k \in G_\alpha \text{ for each } \alpha \in A \implies \bigcup_{k=1}^{\infty} A_k, \bigcap_{k=1}^{\infty} A_k \in H.$$

Exercise 2.4

Let $\mathcal{A}$ be a σ-algebra and $(A_n)_{n\geq 1}$ a sequence of events in $\mathcal{A}$. Show that

$$\liminf_{n\to\infty} A_n \in \mathcal{A}; \quad \limsup_{n\to\infty} A_n \in \mathcal{A}; \quad \text{and} \quad \liminf_{n\to\infty} A_n \subseteq \limsup_{n\to\infty} A_n.$$

Answer: Recall the definitions

$$\liminf_{n\to\infty} A_n = \cup_{n=1}^{\infty} \cap_{m\geq n} A_m,$$
$$\limsup_{n\to\infty} A_n = \cap_{n=1}^{\infty} \cup_{m\geq n} A_m.$$

For each positive integer n, $(A_m)_{m\geq n}$ is a countable sequence of events in $\mathcal{A}$. By the definition of a σ-algebra, this means both $\cap_{m\geq n} A_m$ and $\cup_{m\geq n} A_m$ belong to $\mathcal{A}$ as $\mathcal{A}$ is closed under countable intersections and unions. But then $(\cap_{m\geq n} A_m)_{n\geq 1}$ and $(\cup_{m\geq n} A_m)_{n\geq 1}$ are each countable sequences of events in $\mathcal{A}$ so that again by the definition of a σ-algebra

$$\liminf_{n\to\infty} A_n = \cup_{n=1}^{\infty} \cap_{m\geq n} A_m \in \mathcal{A} \quad \text{and} \quad \limsup_{n\to\infty} A_n = \cap_{n=1}^{\infty} \cup_{m\geq n} A_m.$$

Suppose $a \in \liminf_{n\to\infty} A_n$. Then there exists a positive integer n such that $a \in \cap_{m\geq n} A_m$. Since $a \in A_m$ for every $m \geq n$, $a \in \cup_{i\geq k}^{\infty} A_i$ for each k (no matter how large we choose k, there is an $m \geq n$ such that $m \geq k$ so that $a \in A_m \subseteq \cup_{i\geq k} A_i$). Thus $a \in \cap_{i=1}^{\infty} \cup_{k\geq i} A_k = \limsup_{n\to\infty} A_n$, which establishes $\liminf_{n\to\infty} A_n \subseteq \limsup_{n\to\infty} A_n$.

Exercise 2.5

Let $(A_n)_{n\geq 1}$ be a sequence of sets. Show that

$$\limsup_{n\to\infty} 1_{A_n} - \liminf_{n\to\infty} 1_{A_n} = 1_{\limsup_n A_n \setminus \liminf_n A_n}.$$

Answer: Assume that whenever we index in what follows, n comes from the set of positive integers.

Lemma: For all $\omega \in \Omega$,

$$\liminf_{n\to\infty} 1_{A_n}(\omega) = 1_{\liminf_n A_n},$$
$$\limsup_{n\to\infty} 1_{A_n}(\omega) = 1_{\limsup_n A_n}.$$

Proof: Note that $1_{\cap_n B_n} = \inf_n 1_{B_n}$ and $1_{\cup_n B_n} = \sup_n 1_{B_n}$. This follows from,

$$\begin{aligned} 1_{\cap_n B_n}(\omega) = 1 &\iff \omega \in \cap_n B_n \\ &\iff \forall n, \omega \in B_n \\ &\iff \forall n, 1_{B_n}(\omega) = 1 \\ &\iff \inf_n 1_{B_n}(\omega) = 1. \end{aligned}$$

$$\begin{aligned} 1_{\cup_n B_n}(\omega) = 1 &\iff \omega \in \cup_n B_n \\ &\iff \exists n, \omega \in B_n \\ &\iff \exists n, 1_{B_n}(\omega) = 1 \\ &\iff \sup_n 1_{B_n}(\omega) = 1. \end{aligned}$$

With only minor changes to the above, we see that $1_{\cap_{m\geq n} B_m} = \inf_{m\geq n} 1_{B_m}$ and $1_{\cup_{m\geq n} B_m} = \sup_{m\geq n} 1_{B_m}$ as well. Therefore,

$$\begin{aligned}
1_{\liminf_n A_n} &= 1_{\cup_n \cap_{m\geq n} A_m} = \sup_n 1_{\cap_{m\geq n} A_m} = \sup_n \inf_{m\geq n} 1_{A_m} = \liminf_{n\to\infty} A_n, \\
1_{\limsup_n A_n} &= 1_{\cap_n \cup_{m\geq n} A_m} = \inf_n 1_{\cup_{m\geq n} A_m} = \inf_n \sup_{m\geq n} 1_{A_m} = \limsup_{n\to\infty} A_n.
\end{aligned}$$

Lemma: For $A, B \subset \Omega$, $1_{A\backslash B} = 1_A - 1_{A\cap B}$.

Proof: For any $\omega \in \Omega$,

$$\begin{aligned}
1_{A\backslash B}(\omega) = 1 &\iff \omega \in A, \omega \notin B \\
&\iff 1_A(\omega) = 1 \text{ and } 1_{A\cap B}(\omega) = 0 \\
&\iff 1_A(\omega) - 1_{A\cap B}(\omega) = 1.
\end{aligned}$$

Using the two lemmas and the result $\liminf_{n\to\infty} A_n \subseteq \limsup_{n\to\infty} A_n$ from Exercise 2.4,

$$\begin{aligned}
\limsup_{n\to\infty} 1_{A_n} - \liminf_{n\to\infty} 1_{A_n} &= 1_{\limsup_n A_n} - 1_{\liminf_n A_n} \\
&= 1_{\limsup_n A_n} - 1_{\liminf_n A_n \cap \limsup_n A_n} \\
&= 1_{\limsup_n A_n \backslash \liminf_n A_n}
\end{aligned}$$

Exercise 2.6

Let $\mathcal{A}$ be a σ-algebra of subsets of Ω and let $B \in \mathcal{A}$. Show that $\mathcal{F} = \{A \cap B : A \in \mathcal{A}\}$ is a σ-algebra of subsets of B. Is it still true when B is a subset of Ω that does not belong to $\mathcal{A}$?

Answer: To prove that $\mathcal{F} \subseteq 2^B$ is a σ-algebra of subsets of B, verify axioms (1), (2), and (4) of Definition 2.1.

1. To prove that $\mathcal{F}$ is a σ-algebra of subsets of B, check that $\emptyset, B \in \mathcal{F}$ (no need to check $\Omega \in \mathcal{F}$). Since $\emptyset, B \in A$, $\emptyset = \emptyset \cap B \in F$ and $B = B \cap B \in \mathcal{F}$.

2. Let $F \in \mathcal{F}$ with $F = A \cap B$ for some $A \in \mathcal{A}$. Since $A, B \in \mathcal{A}$, $F \in \mathcal{A}$ and so $B\backslash F = B \cap F^c \in \mathcal{A}$ as well. Since $F \subseteq B$, the complement of F relative to B is $F^c = B\backslash F = (B\backslash F) \cap B \in \mathcal{F}$.

4. Let $(F_n)_{n\geq 1}$ be a sequence of sets in $\mathcal{F}$ with $F_n = A_n \cap B$ for $A_n \in \mathcal{A}$. Because $\mathcal{A}$ is closed under countable unions and intersections,

$$\bigcup_{k=1}^{\infty} F_k = \bigcup_{k=1}^{\infty} (A_n \cap B) = \left(\bigcup_{k=1}^{\infty} A_n \right) \cap B \in \mathcal{F},$$

$$\bigcap_{k=1}^{\infty} F_k = \bigcap_{k=1}^{\infty} (A_n \cap B) = \left(\bigcap_{k=1}^{\infty} A_n \right) \cap B \in \mathcal{F}.$$

Exercise 2.7

Let f be a function mapping Ω to another space E with a σ-algebra $\mathcal{E}$. Let $\mathcal{A} = \{A \subset \Omega : \exists B \in \mathcal{E}, A = f^{-1}(B)\}$. Show that $\mathcal{A}$ is a σ-algebra on Ω.

1. $\emptyset \in \mathcal{E}$ since $\mathcal{E}$ is a σ-algebra. To see that $f^{-1}(\emptyset) = \emptyset$ suppose instead $f^{-1}(\emptyset) = A \neq \emptyset$. This would mean there is $a \in A \subseteq \Omega$ such that $f(a) \in \emptyset$, contradicting the definition of $\emptyset$. Thus $\emptyset \in \mathcal{A}$. Also $\Omega = \emptyset^c = \Omega \backslash \emptyset \in \mathcal{A}$ by (2), which is proved below.

2. Suppose $A \in \mathcal{A}$ with $A = f^{-1}(B)$. Then $A^c = (f^{-1}(B))^c = f^{-1}(B^c) \in \mathcal{A}$ since $B^c \in \mathcal{E}$ and

$$x \in (f^{-1}(B))^c \iff x \notin f^{-1}(B) \iff f(x) \notin B \iff f(x) \in B^c \iff x \in f^{-1}(B^c).$$

4. Let $(A_n)_{n \geq 1}$ be a sequence of sets in $\mathcal{A}$ with $A_n = f^{-1}(B_n)$.

$$\bigcup_{k=1}^{\infty} A_k = \bigcup_{k=1}^{\infty} f^{-1}(B_k) = f^{-1}\left(\bigcup_{k=1}^{\infty} B_k\right) \in \mathcal{A}$$

as $\bigcup_{k=1}^{\infty} B_k \in \mathcal{E}$ and

$$x \in f^{-1}\left(\bigcup_{k=1}^{\infty} B_k\right) \iff f(x) \in \bigcup_{k=1}^{\infty} B_k \iff \exists k, f(x) \in B_k \iff \exists k, x \in f^{-1}(B_k) \iff x \in \bigcup_{k=1}^{\infty} f^{-1}(B_k).$$

Using this result and the fact that $\mathcal{A}$ is closed under complement by (2), $\bigcap_{k=1}^{\infty} A_k \in \mathcal{A}$ as well.

Exercise 2.8

Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function, and let $\mathcal{A} = \{A \subseteq \mathbb{R} : \exists B \in \mathfrak{B}, A = f^{-1}(B)\}$ where $\mathfrak{B}$ are the Borel subsets of the range space $\mathbb{R}$. Show that $\mathcal{A} \subset \mathfrak{B}$, the Borel subsets of the domain space $\mathbb{R}$.

Answer: Suppose $A \in \mathcal{A}$ so that $A = f^{-1}(B)$ for some $B \in \mathfrak{B}$. Since $B \in \mathfrak{B}$, B is the result of applying a countable number of complements, unions, and/or intersections to a collection of open intervals in $\mathbb{R}$. Since f^{-1} commutes with these set operations, A is the result of applying countably many set operations to the inverse images of open intervals in $\mathbb{R}$. Since f is continuous, the inverse image of an open interval is also an open interval. Applying countably many set operations to a collection of open intervals leaves a Borel set. Thus $A \in \mathfrak{B}$.

Exercise 2.15

Let $\mathcal{A}$ be a σ-algebra on the space Ω and P a probability defined on $(\Omega, \mathcal{A})$. Let $A_i \in \mathcal{A}$ be a sequence of events. Show that

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i) \quad \forall n,$$

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} P(A_i).$$

Answer:

Prove the first inequality (the finite case) by induction. For only one set $A \in \mathcal{A}$ equality holds and thus also inequality. Assume that the inequality holds and consider a sequence $A_1, \dots, A_{n+1}$. Let $A'_{n+1} = A_{n+1} \backslash (A_1 \cup \cdots \cup A_n)$ for some $n \geq 1$. Then $A'_{n+1} \cap (A_1 \cup \dots A_n) = \emptyset$ and $A'_{n+1} \subseteq A_{n+1}$.

$$\begin{aligned}
P\left(\bigcup_{i=1}^{n+1} A_i\right) &= P\left(\left(\bigcup_{i=1}^{n} A_n\right) \cup A'_{n+1}\right) \\
&= \sum_{i=1}^{n} P(A_i) + P(A'_{n+1}) \\
&\leq \sum_{i=1}^{n} P(A_i) + P(A_{n+1}) \\
&= \sum_{i=1}^{n+1} P(A_i)
\end{aligned}$$

To prove countable subadditiviy, let

$$\begin{aligned}
E_1 &:= A_1 \\
E_2 &:= A_2 \backslash E_1 \\
E_3 &:= A_3 \backslash (E_1 \cup E_2) \\
E_4 &:= A_4 \backslash (E_1 \cup E_2 \cup E_3) \\
&\vdots \\
E_n &:= A_n \backslash \left(\bigcup_{i=1}^{n-1} E_i\right) \qquad \vdots
\end{aligned}$$

The E_i are disjoint with $\dot\bigcup_{i=1}^{\infty} E_i = \bigcup_{i=1}^{\infty} A_i$. To see that these unions are equal, first note that $E_i \subseteq A_i$ for each i so $\dot\bigcup_{i=1}^{\infty} E_i \subseteq \bigcup_{i=1}^{\infty} A_i$. If $x \in \bigcup_{i=1}^{\infty} A_i$ we may choose by the Well-Ordering Principle the least index i such that $x \in A_i$. Then $x \in A_i$ and $x \notin A_j$ for $j < i$. Thus $x \in E_i \subseteq \dot\bigcup_{i=1}^{\infty} E_i$. Since $E_i \subset A_i$ for each i, $P(E_i) \leq P(A_i)$ for each i and

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = P(\left(\dot\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i) \leq \sum_{i=1}^{\infty} P(A_i).$$

Exercise 2.17

Suppose that Ω is an infinite set (countable or not), and let $\mathcal{A}$ be the family of all subsets which are either finite or have a finite complement. Show that $\mathcal{A}$ is an algebra, but not a σ- algebra.

1. Both $\emptyset \in \mathcal{A}$ and $\Omega \in \mathcal{A}$ as $\Omega^c = \emptyset$ is finite.

2. Suppose $A \in \mathcal{A}$. If A is finite, then $A^c \in \mathcal{A}$ since $(A^c)^c = A$. If A is infinite, then A^c must be finite so $A^c \in \mathcal{A}$.

3. Suppose $A_1, \dots, A_n \in \mathcal{A}$. If all of the A_i are finite, then the finite union of finite sets $\bigcup_{i=1}^{n} A_i$ is finite. If there is a set A_k, $1 \leq k \leq n$ such that A_k is infinite then $\bigcup_{i=1}^{n} A_i$ is not finite. However, A_k^c must be finite and $(\bigcup_{i=1}^{n} A_i)^c = \bigcap_{i=1}^{n} A_i^c \subset A_k^c$. This shows that $(\bigcup_{i=1}^{n} A_i)^c$ is finite so $\bigcup_{i=1}^{n} A_i \in \mathcal{A}$. Since $\mathcal{A}$ is closed under complement and finite union, $\bigcap_{i=1}^{n} A_i \in \mathcal{A}$ as well.

   Since $\mathcal{A}$ satisfies axioms (1),(2), and (3), $\mathcal{A}$ is an algebra. However $\mathcal{A}$ is not a σ-algebra since it fails axiom (4):

4. $\mathcal{A}$ is not necessarily closed under countable union. Either Ω is countably infinite or uncountable.

   - If Ω is countably infinite, we can list the elements of $\Omega = \{\omega_1, \omega_2, \dots\}$. Let $A_i = x_{2i}$ for each positive integer i. Then both $\bigcup_{i=1}^{\infty} A_i = \{x_2, x_4, \dots\}$ is infinite and $\left(\bigcup_{i=1}^{\infty} A_i\right)^c = \{x_1, x_3, \dots\}$ is infinite so $\bigcup_{i=1}^{\infty} A_i \notin \mathcal{A}$.
   - If Ω is uncountable, let $(A_i)_{n\geq 1}$ be a sequence of pairwise disjoint singleton sets. Then $\bigcup_{i=1}^{\infty} A_i$ has countably infinitely many elements and $\left(\bigcup_{i=1}^{\infty} A_i\right)^c$ must be uncountable (since Ω is uncountable). Since neither of $\bigcup_{i=1}^{\infty} A_i$, $\left(\bigcup_{i=1}^{\infty} A_i\right)^c$ is finite, $\bigcup_{i=1}^{\infty} A_i \notin \mathcal{A}$

Chapter 3 Conditional Probability and Independence

Notes

Definition 3.1

1. Two events A and B are independent if $P(A \cap B) = P(A)P(B)$.
2. A (possibly infinite) collection of events $(A_i)_{i \in I}$ is an independent collection if for every finite subset J of I,

$$P\left(\cap_{i \in J} A_i\right) = \Pi_{i \in J} P(A_i).$$

   The collection $(A_i)_{i \in I}$ is said to be mutually independent.

Theorem 3.1

If A and B are independent, so also are A and B^c, A^c and B, A^c and B^c.
Proof: For A and B^c,

$$\begin{aligned} P(A \cap B^c) &= P(A) - P(A \cap B) \quad \text{(Exercise 2.12)} \\ &= P(A) - P(A)P(B) \quad \text{(Definition 3.1)} \\ &= P(A)(1 - P(B)) \\ &= P(A)P(B^c) \quad \text{(Exercise 2.11).} \end{aligned}$$

To prove that B and A^c are independent, switch A with B and B^c with A^c and repeat the previous argument.
For A^c and B^c,

$$\begin{aligned} P(A^c \cap B^c) &= P((A \cup B)^c) \\ &= 1 - P(A \cup B) \\ &= 1 - (P(A) + P(B) - P(A \cap B)) \quad \text{(Exercise 2.10)} \\ &= 1 - P(A) - P(B) + P(A)P(B) \\ &= (1 - P(B)) - P(A)(1 - P(B)) \\ &= P(B^c) - P(A)P(B^c) \quad \text{(Exercise 2.11)} \\ &= (1 - P(A))P(B^c) \\ &= P(A^c)P(B^c). \end{aligned}$$

Example Let $\Omega = \{1, 2, 3, 4\}$, and $\mathcal{A} = 2^{\Omega}$. Let $P(i) = \frac{1}{4}$, where $i = 1, 2, 3, 4$. Let $A = \{1, 2\}, B = \{1, 3\}, C = \{2, 3\}$. Then A, B, C are pairwise independent but are not independent.

$$\begin{aligned} P(A \cap B) = P(\{1\}) = \frac{1}{4} = \frac{1}{2}\frac{1}{2} &= P(A)P(B), \\ P(A \cap C) = P(\{2\}) = \frac{1}{4} = \frac{1}{2}\frac{1}{2} &= P(A)P(C), \\ P(B \cap C) = P(\{3\}) = \frac{1}{4} = \frac{1}{2}\frac{1}{2} &= P(B)P(C), \\ P(A \cap B \cap C) = P(\emptyset) = 0 \neq \left(\frac{1}{2}\right)^3 &= P(A)P(B)P(C). \end{aligned}$$

Definition 3.2

Let A, B be events, $P(B) > 0$. The conditional probability of A given B is $P(A|B) = P(A \cap B)/P(B)$.

Theorem 3.2

Suppose $P(B) > 0$.

1. A and B are independent if and only if $P(A|B) = P(A)$.
2. The operation $A \to P(A|B)$ from $\mathcal{A} \to [0,1]$ defines a new probability measure on $\mathcal{A}$, called the conditional probability measure given B.

Proof:

1. If A and B are independent, $P(A|B) = P(A \cap B)/P(B) = P(A)P(B)/P(B) = P(A)$. If $P(A|B) = P(A)$, $P(A \cap B)/P(B) = P(A) \implies P(A \cap B) = P(A)P(B)$, so A and B are independent.
2. Let $Q(A) := P(A|B)$. Verify that Q satisfies Definition 2.3.

$$Q(\Omega) = P(\Omega|B) = P(\Omega \cap B)/P(B) = P(B)/P(B) = 1.$$

If $(A_n)_{n\geq 1}$ is a countable sequence of pairwise disjoint elements of $\mathcal{A}$ then $(A_n \cap B)_{n\geq 1}$ is also a sequence of pairwise disjoint elements of $\mathcal{A}$ (If $i \neq j$, $(A_i \cap B) \cap (A_j \cap B) = A_i \cap A_j \cap B = \emptyset \cap B = \emptyset$).

$$\begin{aligned}
Q\left(\left(\bigcup_{k=1}^{\infty} A_k\right) \cap B\right) &= \frac{P\left(\left(\bigcup_{k=1}^{\infty} A_k\right) \cap B\right)}{P(B)} \\
&= \frac{P\left(\bigcup_{k=1}^{\infty} (A_k \cap B)\right)}{P(B)} \\
&= \frac{\sum_{k=1}^{\infty} P(A_k \cap B)}{P(B)} \\
&= \sum_{k=1}^{\infty} \frac{P(A_k \cap B)}{P(B)} \\
&= \sum_{k=1}^{\infty} P(A_k|B) \\
&= \sum_{k=1}^{\infty} Q(A_k).
\end{aligned}$$

Theorem 3.3

If $A_1, \ldots, A_n \in \mathcal{A}$ with $P(A_1 \cap \cdots \cap A_{n-1}) > 0$,

$$P(A_1 \cap \cdots \cap A_n) = P(A_1)P(A_2|A_1)P(A_3|(A_1 \cap A_2)) \ldots P(A_n|(A_1 \cap \cdots \cap A_{n-1})).$$

Proof (Induction): For $n = 2$, the equality holds by Definition 3.2. Suppose the theorem holds for n events, $n \geq 2$. Let $B = A_1 \cap \ldots A_n$.

$$\begin{aligned}
&P(A_1 \cap \dots \cap A_n \cap A_{n+1}) \\
=&P(A_{n+1} \cap B) \\
=&P(A_{n+1}|B)P(B) \\
=&P(A_{n+1}|B)P(A_1)P(A_2|A_1)P(A_3|(A_1 \cap A_2)) \dots P(A_n|(A_1 \cap \dots \cap A_{n-1})) \\
=&P(A_1)P(A_2|A_1)P(A_3|(A_1 \cap A_2)) \dots P(A_n|(A_1 \cap \dots \cap A_{n-1})P(A_{n+1}|(A_1 \cap \dots \cap A_n))
\end{aligned}$$

Theorem 3.4 (Partition Equation)

A collection of events (E_n), $E_n \in \mathcal{A}$, is called a partition of Ω if they are pairwise disjoint, $P(E_n) > 0$ for each n, and $\cup_n E_n = \Omega$. Let $(E_n)_{n \geq 1}$ be a finite or countable partition of Ω. If $A \in \mathcal{A}$,

$$P(A) = \sum_n P(A|E_n)P(E_n).$$

Proof: Since the E_n are pairwise disjoint, the $A \cap E_n$ are also pairwise disjoint.

$$P(A) = P(A \cap \Omega) = P(A \cap (\cup_n E_n)) = P(\cup_n (A \cap E_n)) = \sum_n P(A \cap E_n) = \sum_n P(A|E_n)P(E_n).$$

Theorem 3.5 (Baye's Theorem)

Let $(E_n)_{n \geq 1}$ be a finite or countable partition of Ω and $P(A) > 0$.

$$P(E_n|A) = \frac{P(A|E_n)P(E_n)}{\sum_m P(A|E_m)P(E_m)}.$$

Proof: By Theorem 3.4,

$$\frac{P(A|E_n)P(E_n)}{\sum_m P(A|E_m)P(E_m)} = \frac{P(A|E_n)P(E_n)}{P(A)} = \frac{P(A \cap E_n)}{P(A)} = P(E_n|A).$$

Exercises

Exercise 3.1

Show that if $A \cap B = \emptyset$, then A and B cannot be independent unless $P(A) = 0$ or $P(B) = 0$.

Answer: Unless one or both of $P(A), P(B)$ is zero, $P(A)P(B) \neq 0 = P(A \cap B)/P(B) = P(A|B)$, meaning that A and B are not independent.

Exercise 3.2

Let $P(C) > 0$. Show that $P(A \cup B|C) = P(A|C) + P(B|C) - P(A \cap B|C)$.

Answer:

$$
\begin{aligned}
P(A \cup B|C) &= \frac{P((A \cup B) \cap C)}{P(C)} \\
&= \frac{P((A \cap C) \cup (B \cap C)}{P(C)} \\
&= \frac{P(A \cap C) + P(B \cap C) - P((A \cap C) \cap (B \cap C)}{P(C)} \\
&= \frac{P(A \cap C)}{P(C)} + \frac{P(B \cap C)}{P(C)} - \frac{P((A \cap B) \cap C)}{P(C)} \\
&= P(A|B) + P(B|C) - P(A \cap B|C).
\end{aligned}
$$

Exercise 3.6

Donated blood is screened for AIDS. Suppose the test has 99% accuracy, and that one in ten thousand people in your age group are HIV positive. The test has a 5% false positive rating, as well. Suppose the test screens you as positive. What is the probability you have AIDS?

Answer: Let A be the event that you have AIDS and B the event that you test HIV Positive. The events A, A^c are a finite partition of the probability space. By Baye's Theorem, the probability that you have AIDS given that you have tested positive is

$$
P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)} = \frac{(.99)(.0001)}{(.99)(.0001) + (0.05)(.9999)} \approx 0.001976.
$$

Exercise 3.7

Let $(A_n)_{n\geq 1}, (B_n)_{n\geq 1}$ with $A_n, B_n \in \mathcal{A}$ for each n, $A_n \to A$, $B_n \to B$, $P(B) > 0$, and $P(B_n) > 0$ for each n.

1. $\lim_{n\to\infty} P(A_n|B) = P(A|B)$,
2. $\lim_{n\to\infty} P(A|B_n) = P(A|B)$,
3. $\lim_{n\to\infty} P(A_n|B_n) = P(A|B)$.

Answer:

1. Since $A_n \to A$, $A_n \cap B \to A \cap B$. By Theorem 2.4, $\lim_{n\to\infty} P(A_n \cap B) = P(A \cap B)$.

$$
\lim_{n\to\infty} P(A_n|B) = \lim_{n\to\infty} \frac{P(A_n \cap B)}{P(B)} = \frac{P(A \cap B)}{P(B)}.
$$

2. Since $B_n \to B$, $A \cap B_n \to A \cap B$. By Theorem 2.4, $\lim_{n\to\infty} P(A \cap B_n) = P(A \cap B)$.

$$
\lim_{n\to\infty} P(A|B_n) = \lim_{n\to\infty} \frac{P(A \cap B_n)}{P(B_n)} = \frac{P(A \cap B)}{P(B)}.
$$

3. Since $A_n \to A, B_n \to B$, $A_n \cap B_n \to A \cap B$. By Theorem 2.4, $\lim_{n\to\infty} P(A_n \cap B_n) = P(A \cap B)$.

$$
\lim_{n\to\infty} P(A_n|B_n) = \lim_{n\to\infty} \frac{P(A_n \cap B_n)}{P(B_n)} = \frac{P(A \cap B)}{P(B)}.
$$

Exercise 3.11

(Polya's Urn) An urn contains r red balls and b blue balls. A ball is chosen at random from the urn, its color is noted, and it is returned together with d more balls of the same color. This is repeated indefinitely. What is the probability that

1. The second ball drawn is blue?
2. The first ball drawn is blue given that the second ball drawn is blue?

Answer: Let B_n be the event that the nth ball drawn is blue and R_n the event that the nth ball drawn is red.

1. $P(B_2) = P(B_2|B_1)P(B_1) + P(B_2|R_1)P(R_1) = \frac{b+d}{b+r+d}\frac{b}{b+r} + \frac{b}{b+r+d}\frac{r}{b+r} = \frac{b}{b+r}\left(\frac{b+d+r}{b+r+d}\right) = \frac{b}{b+r}$.
2. $P(B_1|B_2) = \frac{P(B_1 \cap B_2)}{P(B_2)} = \frac{P(B_2|B_1)P(B_1)}{P(B_2)} = \frac{b+d}{b+r+d}\frac{b}{b+r}\frac{b+r}{b} = \frac{b+d}{b+r+d}$

Exercise 3.12

Consider the framework of Exercise 3.11. Let B_n denote the event that the nth ball drawn is blue. Show that $P(B_n) = P(B_1)$ for all $n \geq 1$.

Answer: Prove $P(B_n) = P(B_1)$ for all $n \geq 1$ by induction. Exercise 3.11 showed $P(B_2) = P(B_1) = b/(b+r)$. Assume that $P(B_n) = P(B_1)$ for some $n \geq 1$. Let b_n, r_n stand respectively for the number of blue and red balls in the urn during the nth draw.

$$
\begin{aligned}
P(B_{n+1}) &= P(B_{n+1}|B_n)P(B_n) + P(B_{n+1}|R_n)P(R_n) \\
&= \frac{b_n + d}{b_n + r_n + d}\frac{b_n}{b_n + r_n} + \frac{b_n}{b_n + r_n + d}\frac{r_n}{b_n + r_n} \\
&= \frac{b_n}{b_n + r_n}\left(\frac{b_n + d}{b_n + r_n + d} + \frac{r_n}{b_n + r_n + d}\right) \\
&= \frac{b_n}{b_n + r_n} \\
&= P(B_n) \\
&= P(B_1).
\end{aligned}
$$

Exercise 3.13

Consider the framework of Exercise 3.11. Find the probability that the first ball is blue given that the n subsequent drawn balls are all blue. Find the limit of this probability as $n \to \infty$.
Answer:

$$
\begin{aligned}
P(B_1|B_2\cap\cdots\cap B_{n+1}) &= \frac{P(B_2\cap\cdots\cap B_{n+1}|B_1)P(B_1)}{P(B_2\cap\cdots\cap B_{n+1})}\\
&= \frac{P(B_2\cap\cdots\cap B_{n+1}|B_1)P(B_1)}{P(B_2\cap\cdots\cap B_{n+1}|B_1)P(B_1)+P(B_2\cap\cdots\cap B_{n+1}|R_1)P(R_1)}\\
&= \frac{\frac{b+d}{b+r+d}\,\frac{b+2d}{b+r+2d}\cdots\frac{b+nd}{b+r+nd}\,\frac{b}{b+r}}{\frac{b+d}{b+r+d}\,\frac{b+2d}{b+r+2d}\cdots\frac{b+nd}{b+r+nd}\,\frac{b}{b+r}+\frac{b}{b+r+d}\,\frac{b+d}{b+2r+d}\cdots\frac{b+(n-1)d}{b+r+nd}\,\frac{r}{b+r}}\\
&= \frac{(b+d)(b+2d)\dots(b+nd)b}{[(b+d)(b+2d)\dots(b+nd)b]+[b(b+d)(b+2d)\dots(b+(n-1)d)r]}\\
&= \frac{b+nd}{b+nd+r}\\
&= \frac{b+nd}{b+r+nd}.
\end{aligned}
$$

$$
\lim_{n\to\infty} P(B_1|B_2\cap\cdots\cap B_{n+1}) = \lim_{n\to\infty}\frac{b+nd}{b+r+nd} = \lim_{n\to\infty}\frac{b/n+d}{b/n+r/n+d} = 1.
$$