Counterexamples of Probability

Seoncheol Park

2019-05-20

Contents

1 Prerequisites

This is a *sample* book written in **Markdown**. You can use anything that Pandoc's Markdown supports, e.g., a math equation $a^2 + b^2 = c^2$.

The **bookdown** package can be installed from CRAN or Github:

```
install.packages("bookdown")
# or the development version
# devtools::install_github("rstudio/bookdown")
```

Remember each Rmd file contains one and only one chapter, and a chapter is defined by the first-level heading `#`.

To compile this example to PDF, you need XeLaTeX. You are recommended to install TinyTeX (which includes XeLaTeX): https://yihui.name/tinytex/.

2 Introduction

You can label chapter and section titles using `{#label}` after them, e.g., we can reference Chapter 2. If you do not manually label them, there will be automatic labels anyway, e.g., Chapter **??**.

Figures and tables with captions will be placed in `figure` and `table` environments, respectively.

```
par(mar = c(4, 4, .1, .1))
plot(pressure, type = 'b', pch = 19)
```

Reference a figure by its code chunk label with the `fig:` prefix, e.g., see Figure 2.1. Similarly, you can reference tables generated from `knitr::kable()`, e.g., see Table 2.1.

```
knitr::kable(
  head(iris, 20), caption = 'Here is a nice table!',
  booktabs = TRUE
)
```

You can write citations, too. For example, we are using the **bookdown** package (**?**) in this sample book, which was built on top of R Markdown and **knitr** (Xie, 2015).

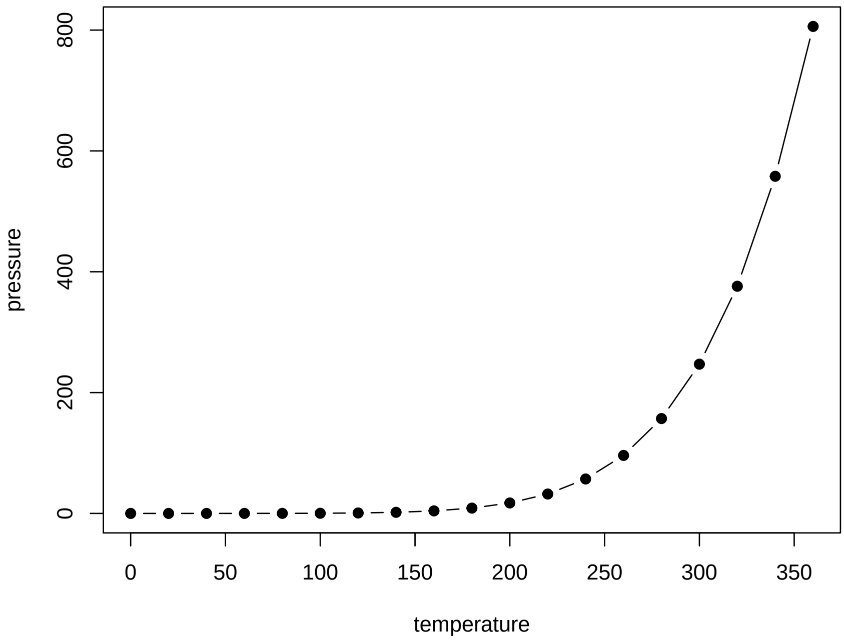

Figure 2.1: Here is a nice figure!

Table 2.1: Here is a nice table!

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
5.1	3.5	1.4	0.2	setosa
4.9	3.0	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
4.6	3.1	1.5	0.2	setosa
5.0	3.6	1.4	0.2	setosa
5.4	3.9	1.7	0.4	setosa
4.6	3.4	1.4	0.3	setosa
5.0	3.4	1.5	0.2	setosa
4.4	2.9	1.4	0.2	setosa
4.9	3.1	1.5	0.1	setosa
5.4	3.7	1.5	0.2	setosa
4.8	3.4	1.6	0.2	setosa
4.8	3.0	1.4	0.1	setosa
4.3	3.0	1.1	0.1	setosa
5.8	4.0	1.2	0.2	setosa
5.7	4.4	1.5	0.4	setosa
5.4	3.9	1.3	0.4	setosa
5.1	3.5	1.4	0.3	setosa
5.7	3.8	1.7	0.3	setosa
5.1	3.8	1.5	0.3	setosa

3 Classes of Random Events

3.1 Sample space and Event

Let Ω (the **sample space** of all possible outcomes of an experiments) be an arbitrary non-empty set. Its elements, denoted by ω, will be interpreted as **outcomes (results)** of some experiment. These outcomes sometimes called **simple events** or **elements** becaus they cannot be decomposed further. Other events (subsets of Ω) are **compound events** because they are composed of two or more simple events. An **event** A occurs if and only if the randomly selected ω belongs to A. As usual, we use $A \cup B$ and $A \cap B$ to represent the union and the intersection of any two subsets A and B of Ω, respectively. Also, A^c is the complement of $A \subset \Omega$. In particular, $\Omega^c = \theta$, where θ is the empty set.

Example 3.1 (an example of sample space and event)**.** When drawing randomly from $[0, 1]$ the sample space is $\Omega = [0, 1]$, each $\omega \in [0, 1]$ is a sample event, and the interval $A = (\frac{1}{2}, 1]$ is an example of a compound event.

3.2 Field and sigma-field

Definition 3.1 (Field)**.** The class $\mathcal{A}$ of subsets of Ω is called a **field** if it contains Ω and is closed under the formation of complements and finite unions, that is if:

1. $\Omega \in \mathcal{A}$;
2. $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$;
3. $A_1, A_2 \in \mathcal{A} \Rightarrow A_1 \cup A_2 \in \mathcal{A}$.

Proposition 3.1 (Definition of a field)**.** *The condition 3 of the definition of a field can be replaced by*

3. $A_1, A_2 \in \mathcal{A} \Rightarrow A_1 \cap A_2 \in \mathcal{A}$.

That means, $\mathcal{A}$ is closed under finite intersections.

Proof. By de Morgan laws, $(A_1 \cap A_2)^C = A_1^C \cup A_2^C$ and $(A_1 \cup A_2)^c = A_1^C \cap A_2^C$. By the condition 2 of the definition of a field, we can get the desired result. $\square$

Exercise 3.1. Which of the followings are σ-fields of subsets of $\Omega = \{1,2,3,4,5,6\}$?

a) $\{\{1,2,\},\{1,2,3,\},\{4,5,6\},\emptyset,\Omega\}$.

b) $\{\{1,2\},\{3,4,5,6\},\emptyset,\Omega\}$.

c) $\{\{1\},\{3,4,5,6\},\{1,3,4,5,6\},\Omega\}$.

Solution. b.

Definition 3.2 (Semi-field)**.** Let Ω be an arbitrary set. A non-empty class $\mathcal{J}$ of subsets of Ω is called a **semi-field** if $\Omega = \mathcal{J}, \emptyset \in \mathcal{J}$, $\mathcal{J}$ is closed under the formation of finite intersections, and the complement of any set in $\mathcal{J}$ is a finite sum of disjoint sets of $\mathcal{J}$.

It is easy to see that any field of subsets of Ω is also a semi-field. However, the following simple examples show that the converse is not true.

Example 3.2 (A class of events which is a semi-field but not a field)**.** 1. Let $\Omega = [-\infty,\infty)$ and $\mathcal{J}_1$, contain Ω, $\{\infty\}$ and all intervals of the type $[a,b)$ where $-\infty < a \leq b \leq \infty$. Then, $\emptyset \in \mathcal{J}_1$, $\Omega \in \mathcal{J}_1$, $[a_1,b_1) \cap [a_2,b_2) = [a_1 \vee a_2, b_1 \wedge b_2) \in \mathcal{J}_1$, and $[a,b)^C = [-\infty,a) \cup [b,\infty)$. So, $\mathcal{J}_1$, is a semi-field. Obviously $\mathcal{J}_1$ is not a field (Check 3.3).

2. Take $\Omega = \mathbb{R}^1$ and denote by $\mathcal{J}_2$ the class of all subsets of the form $A \cap B$ where A is closed and B is an open set in Ω. Then again, $\mathcal{J}_2$ is a semi-field but not a field.

Definition 3.3 (Sigma-field)**.** The class $\mathcal{F}$ of subsets of Ω is called a σ**-field** if it is a field and closed under the formation of countable unions, that is if:

4. $A_1, A_2, \dots, \in \mathcal{F} \Rightarrow \cup_{n=1}^{\infty} A_n \in \mathcal{F}$.

Proposition 3.2 (Definition of a sigma-field)**.** *The condition 4 of the definition of a sigma-field can be replaced by*

4. $A_1, A_2, \dots, \in \mathcal{F} \Rightarrow \cap_{n=1}^{\infty} A_n \in \mathcal{F}$.

That means, σ-field $\mathcal{F}$ is closed under countable intersections because $\bigcap_{n=1}^{\infty} A_n = (\cup_{n=1}^{\infty} A_n^C)^C$ *and each A_n^C belongs to $\mathcal{F}$.*

Proof. By de Morgan laws and condition 2 of the definition of a field, $(\cap_{n=1}^{\infty} A_n)^C = \cup_{n=1}^{\infty} A_n^C \in \mathcal{F}$ and $(\cup_{n=1}^{\infty} A_n)^C = \cap_{n=1}^{\infty} A_n^C \in \mathcal{F}$. So, we can change the condition 4. $\square$

Usually, $\mathcal{F}$ is the set of events (subsets of Ω) we are allowed to consider.

Definition 3.4 (Random events (events))**.** The elements of any field of σ-field are called **random events** (or simply, **events**).

Example 3.3 (A class of events which is a field but not a sigma-field)**.** Let $\Omega = [0, \infty)$ and $\mathcal{F}_1$ be the class of all intervals of the type $[a, b)$ or $[a, \infty)$, where $0 \leq a < b < \infty$. Let $\mathcal{F}_2$ be the class of all finite sums of intervals of $\mathcal{F}_1$. Then $\mathcal{F}_1$ is not a field, and $\mathcal{F}_2$ is a field but not a σ-field.

Take arbitrary numbers a and b, $0 < a < b < \infty$. Then $A = [a, b) \in \mathcal{F}_1$. However, $A^C = [0, a) \cup [b, \infty) \neq \mathcal{F}_1$ and thus $\mathcal{F}_1$ is not a field.

It is easy to see that

1. the finite union of finite sums of intervals (of $\mathcal{F}_1$) is again a sum of intervals;
2. the complement of a finite sum of intervals is also a sum of intervals. This means that $\mathcal{F}_2$ is a field.

However, $\mathcal{F}_2$ is not a σ-field because, for example, the set $A_n = [0, \frac{1}{n}) \in \mathcal{F}_1$ for each $n = 1, 2, \dots$, and the intersection $\bigcap_{n=1}^{\infty} A_n = \{0\}$ does not belong to $\mathcal{F}_1$.

Let us look at two additional cases.

1. Let $\Omega = \mathbb{R}^1$ and $\mathcal{F}$ be the class of all finite sums of intervals of the type $(-∞, a], $(b, c]$ and (d, ∞). Then $\mathcal{F}$ is a field. But the intersection $\cap_{n=1}^{\infty}(b - \frac{1}{n}, c]$ is equal to $[b, c]$ which does not belong to $\mathcal{F}$. Hence, the field $\mathcal{F}$ is not a σ-field.
2. Let Ω be any infinite set and $\mathcal{A}$ the collection of all subsets $A \in \Omega$ such that either A or A^c is finite. Then it is easy to see that $\mathcal{A}$ is a field but not a σ-field.

Example 3.4 (A class of events can be closed under finite unions and finite intersections but not under complements)**.** Let $\Omega = \mathbb{R}$ and the class $\mathcal{A}$ consist of intervals of the type (x, ∞), $x \in \Omega$. Then using the notations $u = x \wedge y := \min\{x, y\}$ and $v = x \vee y := \max\{x, y\}$ we have:

$$(x, \infty) \cup (y, \infty) = (u, \infty) \in \mathcal{A}$$

$$(x, \infty) \cap (y, \infty) = (v, \infty) \in \mathcal{A}.$$

However, $(x, \infty)^C = (-\infty, x] \notin \mathcal{A}$.

Remark. Because any field or σ-field $\mathcal{F}$ is nonempty, it contains a set A, and therefore A^C. Consequently, any field or σ-field contains $\Omega = A \cup A^C$ and $\emptyset = \Omega^C$.

A σ-field is more restrictive than a field. Because we want to be able to consider countable unions, we require the allowable sets $\mathcal{F}$ to be a σ-field. Therefore, it is understood in probability theory and throughout the remainder we are allowed to consider only the events belonging to the σ-field $\mathcal{F}$. We will see that fields also play an important role in probability because a common technique is first to define a probability measure on a field and then extend it to a σ-field.

Definition 3.5 (D-system)**.** A system $\mathcal{D}$ of subsets of a given set Ω is called a **D-system** (Dynkin system) in Ω if the following three conditions hold:

1. $\Omega \in \mathcal{D}$;
2. $A, B \in \mathcal{D}$ and $A \subset B \Rightarrow B \backslash A \in \mathcal{D}$;
3. $A_n \in \mathcal{D}, n = 1, 2, \dots .$ and $A_1 \subset A_2 \subset \dots \Rightarrow \cup_{n=1}^{\infty} A_n \in \mathcal{D}$.

It is obvious that every σ-field is a D-system, but the converse may not be true.

Example 3.5 (Every sigma-field of events is a D-system, but the converse does not always hold)**.** Take $\Omega = \{\omega_1, \omega_2, \dots, \omega_{2n}\}, n \in \mathbb{N}$. Denote by $\mathcal{D}_e$ the collection of all subsets $D \in \Omega$ consisting of an even number of elements. Conditions 1, 2, and 3 above are satisfied, and hence $\mathcal{D}_e$ is a D-system. However, if $n < 1$ and we take $A = \{\omega_1, \omega_2\}$ and $B = \{\omega_2, \omega_3\}$, we see that $A \in \mathcal{D}_e, B \in \mathcal{D}_e$ and $A \cap B = \{\omega_2\} \notin \mathcal{D}_e$. Thus $\mathcal{D}_e$ is not even a field and hence not a σ-field.

Note that a D-system $\mathcal{D}$ is a σ-field if and only if the intersection of any two sets in $\mathcal{D}$ is again in $\mathcal{D}$.

Example 3.6 (Trivial and total sigma-fields)**.** For any sample space Ω, the smallest and largest σ-fields are the **trivial σ-field** $2^{\Omega} = \{\emptyset, \Omega\}$ and the **total σ-field** (also called the power σ-field) $(\Omega) = \{\text{all subsets of } \Omega\}$, respectively.

Remember that we are allowed to determine only whether or not event A occured for $A \in \mathcal{F}$, not for sets outside $\mathcal{F}$. The trivial σ-field is not very useful because the only events that we are allowed to consider are the entire sample space and the empty set. In other words, we can determine only whether something or nothing happened. With the total σ-field we can determine, for each subset $A \subset \Omega$, whether or not event A occurred. The trouble with 2^{Ω} is that it may be too large to ensure that the properties we desire for a probability measure hold for all $A \in \mathcal{F}$.

Therefore, we want $\mathcal{F}$ to be large, but not too large.

Example 3.7 (For countable sample space, the sigma-field should be the total sigma-field)**.** Suppose Ω is a countable set like $\{0, 1, 2, \dots, \}$. We certainly want $\mathcal{F}$ to contain each singleton. But any σ-field that contains each singleton ω must contain all subsets of Ω. This follows from the fact that any subset A of Ω is a countable union $\cup_{n=1}^{\infty} \{\omega_i\}$ of singletons, and σ-fields are closed under countable unions. Therefore, when the sample space is countable, we should always use the total σ-field.

Proposition 3.3 (The intersection of sigma-fields are sigma-fields)**.** *The intersection $\bigcap_t \mathcal{F}_t$ of any collection of σ-fields of Ω is a σ-field.*

Proof. Note that $\mathscr{F}$ is nonempty, because $\emptyset \in \mathscr{F}_t$ and $\Omega \in \mathscr{F}_t$ for each t, and therefore $\emptyset$ and Ω belong to $\mathscr{F} = \bigcap_t \mathscr{F}$. Also, if $A \in \mathscr{F}$, then A belongs to each $\mathscr{F}_t$. Thus, A^C belongs to each $\mathscr{F}_t$, and therefore $\mathscr{F} = \bigcap_t \mathscr{F}_t$. Similarly, if $A_1, A_2, ...$ belong to each $\mathscr{F}_t$, then $A = \bigcup_{i=1}^{\infty} A_i$ also belongs to each $\mathscr{F}_t$, so A belongs to $\mathscr{F} = \bigcap_t \mathscr{F}_t$. It follows that $\mathscr{F}$ is a σ-field. □

Example 3.8 (The union of a sequence of sigma-fields need not be a sigma-field)**.** Let $\mathscr{F}_1, \mathscr{F}_2, ...$ be a sequence of σ-fields of subsets of the set Ω. Then their intersection $\bigcap_{n=1}^{\infty} \mathscr{F}_n$ is always a σ-field and it is natural to ask whether the union $\bigcup_{n=1}^{\infty} \mathscr{F}_n$ is a σ-field. We shall now show that the answer to this question is negative.

Consider the set $\Omega = \{\omega_1, \omega_2, \omega_3\}$ and the following two classes of its subsets: $\mathscr{F}_1 = \{\emptyset, \{\omega_1\}, \{\omega_2, \omega_3\}, \Omega\}$, $\mathscr{F}_2 = \{\emptyset, \{\omega_2\}, \{\omega_1, \omega_3\}, \Omega\}$. Then $\mathscr{F}_1$ and $\mathscr{F}_2$ are fields and hence σ-fields. Obviously the intersection $\mathscr{F}_1 \cap \mathscr{F}_2 = \{\emptyset, \Omega\}$, the trivial σ-field. However, the union

$$
\begin{aligned}
\mathscr{F} &= \mathscr{F}_1 \cup \mathscr{F}_2 \\
&= \{\emptyset, \{\omega_1\}, \{\omega_2\}, \{\omega_2, \omega_3\}, \{\omega_1, \omega_3\}, \Omega\}
\end{aligned}
$$

is not a field, and hence not a σ-field because the element $\{\omega_1\} \cup \{\omega_2\} = \{\omega_1, \omega_2\} \notin \mathscr{F}$.

3.3 Sigma-fields in real line

We next consider σ-fields when Ω is the uncountable set $\mathbb{R}$ (or an interval). Again, any useful σ-field must contain each singleton ω.

Definition 3.6 (Co-finite and co-countable)**.** The set $A \in \Omega$ is called **co-finite** set if its complement A^C is finnite. Further, the set $A \in \Omega$ is called a **co-countable** set if A^C is countable.

Let $\mathscr{F}_1$ consist of the finite and co-finite subsets of Ω. Then, $\mathscr{F}_1$ is a field. It is a σ-field if and only if Ω is finite. Let $\mathscr{F}_2$ consist of the countable and the co-countable subsets of Ω. Then it is easy to check that $\mathscr{F}_2$ is a σ-field.

Exercise 3.2. Prove that the collection of countable and co-countable sets is itself a σ-field.

Example 3.9 (A sigma-field of subsets of sample space need not contain all subsets of sample space)**.** Suppose that Ω is uncountable. Then Ω contains a subset A such that A and A^C are both uncountable. This show that in general a σ-field of Ω need not contain all subsets of Ω and need not be closed under the formation of arbitrary uncountable unions.

Example 3.10 (Sigma-field of countable and co-countable sets: too small)**.** Suppose that Ω is uncountable set like $\mathbb{R}$ or $[0,1]$. If we want $\mathcal{F}$ to include each singleton ω, then $\mathcal{F}$ must contain each countable subset A. It must also contain the complement of each countable set. Therefore, any σ-field containing the singletons must contain all countable and co-countable sets. Unfortunately, this σ-field is not large enough to be useful. For instance, if $\Omega = [0,1]$, then we are not allowed to consider events like $[0, \frac{1}{2}]$ because it is neither countable nor co-countable.

We continue our quest for the smalles σ-field that contains all of the "useful" subsets of the uncountable set $\mathbb{R}$. We have seen that the collection of countable and co-countable sets is not a good choice because it excludes common sets like $[0, \frac{1}{2}]$. We certainly want $\mathcal{F}$ to include all intervals. But remember that we do not want $\mathcal{F}$ to be too large, else it might contain "pathological sets" that cause probability to behave strangely. Therefore, we would like to find the **smallest** σ-field $\mathcal{F}$ that contains all of the intervals. That is, we want $\mathcal{F}$ to contain all intervals, and if $\mathcal{G}$ is any other σ-field containing the intervals, then $\mathcal{F} \subset \mathcal{G}$. The smallest σ-field containing a collection $\mathcal{A}$, denoted $\sigma(\mathcal{A})$, is called the **sigma-field generated by** $\mathcal{A}$. We want to find σ(intervals). Any σ-field containing the intervals must also contain all finite unions of disjoint sets of the form $(a,b]$, $(-\infty, a]$, or (b, ∞). Therefore, if this latter collection is a σ-field, it must be σ(intervals). Unfortunately, it is not a σ-field.

Exercise 3.3. If $\mathcal{A}$ is a σ-field, then $\sigma(\mathcal{A}) = \mathcal{A}$.

Example 3.11 (Very important field)**.** If $\Omega = \mathbb{R}$, the collection $\mathcal{F}_0$ of finite unions of disjoint sets of the form $(a,b]$, $(-\infty, a]$, or (b, ∞) is a field, but not a σ-field. It generates σ(intervals).

Exercise 3.4. Prove the above example.

Continuing aour search for σ(intervals), consider the collection $\{\mathcal{F}_t\}_{t\in T}$ of all σ-fields containing the intervals. There are at least one, namely the total σ-field $2^{\mathbb{R}}$ of all subsets of the line. Let $\mathcal{F} = \bigcap_t \mathcal{F}_t$. We claim that $\mathcal{F} = \sigma$(intervals). We need only show that $\mathcal{F}$ is a σ-field, because any other σ-field containing the intervals clearly contains $\mathcal{F}$.

3.4 Borel sigma-field

Definition 3.7 (1-dimensional Borel sigma-field)**.** The smallest σ-field containing the intervals of $\mathbb{R}$, namely the intersection of all σ-fields containing the intervals, is called the **Borel sigma-field** and denoted $\mathcal{B}$ or $\mathcal{B}^1$. The sets in $\mathcal{B}$ are called **Borel sets**. The Borel sets of $\mathbb{R}$ that are subsets of $[0,1]$ are denoted $\mathcal{B}_{[0,1]}$.

The definition of the Borel σ-field is somewhat dissatisfying because it is hard to picture the intersection of all σ-fields containing the intervals. It would be more pleasing if we could begin with the intervals $\mathcal{I}_1$, throw in all complements and countable unions of sets in $\mathcal{I}_1$ to create $\mathcal{I}_2$, then throw in all complements and countable unions of sets in $\mathcal{I}_2$ to create $\mathcal{I}_3$, etc. It would still be difficult to picture all of the sets, but at least we could imagine how to construct them. Unfortunately, this process must be continued indefinitely. In fact, even $\bigcup_{i=1}^{\infty} \mathcal{I}_i$ does not produce all of the Borel sets. (Billingsley, 2012)

Remark (The Bore sigma-field is very broad). The Borel σ-field contains many sets other than intervals, including, for example:

1. Countable unions of intervals, and therefore all open sets by (Proposition A.15.)
2. All closed sets by part 1 and the fact that complements of sets in $\mathcal{B}$ are in $\mathcal{B}$.
3. The rational numbers by part 2 because each rational r is a closed set (it is easy to see that its complement is open), and the set of rational numbers is a countable union of these closed sets.
4. The irrational numbers, being the complement of the Borel set of rational numbers.

It is actually difficult to construct a set that is **not a Borel set**. Nevertheless, we will see later that an even larger σ-field, called the **Lebesgue sets**, is also useful in probability and measure theory.

It is sometimes helpful to consider the events $\{-\infty\}$ or $\{+\infty\}$. These sets are not in the Borel σ-field because they are not subsets of $\mathbb{R} = (-\infty, \infty)$. To consider them as events, we must extend the real line by augmenting it with $\{-\infty, +\infty\}$.

Definition 3.8 (Extended line and extended intervals)**.** The **extended line** $\bar{\mathbb{R}}$ is $[-\infty, \infty]$, meaning $(-\infty, \infty) \cup \{-\infty, \infty\}$. An extended interval is of the form $[x, y]$, $[x, y)$, $(x, y]$, or (x, y), where x or y can be numbers or $\pm\infty$; $[a, \infty]$ means $[a, \infty) \cup \{\infty\}$, $[-\infty, b]$ means $(-\infty, b] \cup \{-\infty\}$, etc.

When considering $\bar{\mathbb{R}}$ instead of $\mathbb{R}$, we must remember to take complements with respect to $\mathbb{R}$. For instance, the complement of $(0,1]$ is $\bar{\mathbb{R}} \backslash (0,1] = [-\infty, 0] \cup (1, \infty]$, and the complement of $\mathbb{R} = (-\infty, \infty)$ is $\bar{\mathbb{R}} \backslash \mathbb{R} = \{-\infty, \infty\}$.

Definition 3.9 (Extended Borel sigma-field)**.** The extended Borel σ-field $\bar{\mathcal{B}}$ is the smallest σ-field of subsets of $\bar{\mathbb{R}}$ that contains all of the extended intervals.

Proposition 3.4 (Equivalent formulations of extended Borel sigma-field)**.** *The flollowing are equivalent.*

1. *$\bar{\mathcal{B}}$ is the extended Borel σ-field.*
2. *$\bar{\mathcal{B}}$ is the smallest σ-field that contains the two sets $\{-\infty\}$ and $\{+\infty\}$ and all Borel subsets of $\mathbb{R}$.*
3. *$\bar{\mathcal{B}}$ consists of all Borel subsets $\mathcal{B}$ of $\mathbb{R}$, together with all augumentations $\bar{mathcalB}$ of $\mathcal{B}$ by $-\infty$, ∞, or both.*

Exercise 3.5. Prove the above proposition.

3.5 High-dimensional sigma fields

3.6 Infinitely often event

IF $\{A_n\} \in \mathcal{F}$, we can define other events involving the A_n. We have seen, for example, that $B = \cup_{n=1}^{\infty} A_n$ is the event that at least one A_n occurs, while $C = \cap_{n=1}^{\infty} A_n$ is the event that all of the A_n occur. Sometimes there is interest in seeing whether A_n occurs for infinitely many n. For instance, in a coin-flipping experiment, let A_n denote the event that the n-th coin flip is heads. We might be interested in whether infinitely many flips are heads. We now show that for general $A_1, A_2, \dots$, the event that A_n occurs for infinitely many n is a honest (bona fide) event in the sense that it is in $\mathcal{F}$. Note that $\omega \in A_n$ for infinitely many n if and only if no matter how large N i, $\omega \in \cap_{N=1}^{\infty} \cup_{n \geq N} A_n = \cap_{N=1}^{\infty} B_N$, where $B_N = \cup_{n \geq N} A_n$. Because B_N is a countalbe union of $\mathcal{F}$-sets, $B_N \in \mathcal{F}$. This being true for each N, $\cap_{N=1}^{\infty} B_N \in \mathcal{F}$ because $\mathcal{F}$ is closed under countable intersections. (refer to Prop 3.2) Therefore, the event that A_n occurs for infinitely many n is in $\mathcal{F}$. We denote this event by $\{A_n \text{ i.o}\}$, which stands for A_n occurs infinitely often (in n).

Proposition 3.5 (The infinitely often event is in sigma-field.)**.** *If $A_n \in \mathcal{F}$ for $n = 1, 2, \dots$, the event $\{A_n$ i.o$\}$, namely $\cap_{N=1}^{\infty} \cup_{n \geq N} A_n$ is in $\mathcal{F}$.*

Remark. The complement of $\{A_n \text{ i.o}\}$, namely that A_n occurs only finitely often, is also in $\mathcal{F}$ because $\mathcal{F}$ is closed under complementation. This event is, by De Morgans laws, $\{\cap_{N=1}^{\infty} \cup_{n \geq N} A_n\}^C = \cup_{N=1}^{\infty} B_N^C = \cup_{N=1}^{\infty} \cap_{n \geq N} A_n^C$.

This expression can also be deduced from general principles because if ω is in only finitely many of the A_n, then there must be an N for which ω is not in A_n for any $n \geq N$. That is, ω must be in A_n^C for all $n \geq N$. This must occur for at least on N, which translates into a union. Thus again we arrive at the expression $\cup_{N=1}^{\infty} \cap_{n \geq N} A_n^C$ for the event that A_n occurs only finitely often.

4 Measures and Probability Measures

A probability measure P is a function that assigns a probability to the allowable sets $\mathcal{F}$ and that satisfies certain conditions. Ideally, we would like P to satisfiy $P(\Omega) = 1$ and $P(\cup_t A_t) = \sum_t P(A_t)$ for every collection of disjoint sets A_t, even an uncountable collection. Unfortunately, this **uncountable additivity** requirement is too much to ask because it precludes the existence of a "uniform" probability measure P on $\Omega = [0, 1]$. To see this, note that for such a measure, each $\omega = \Omega$ has the same probability p, so if P were uncountably additive, then $1 = P(\Omega) = P([0, 1]) = \sum_\Omega p$. But this produces a contradiction because $\sum_\Omega p$ is 0 if $p = 0$ and ∞ if $p > 0$. Therefore, uncountable additivity is too strong a condition to require. It turns out the strongest property of this type that we con impose is **countable additivity**.

4.1 Probability

Definition 4.1 (Measure and probability measure)**.** Consider a measurable space $(\Omega, \mathcal{F})$. A **measure** μ is defined for all events $A \in \mathcal{F}$ and satisfies the following (Kolmogorov) axioms.

1. For any $A \in \mathcal{F}$, there exists a number $\mu(A) \geq 0$; the measure of A.
2. (**countable additivity**) Let $\{A_n, n \geq 1\}$ be disjoint. Then

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

A **probability measure** P is a measure such that

3. $P(\Omega) = 1$.

Definition 4.2 (Probability space)**.** The triple $(\Omega, \mathcal{F}, P)$ is a **probability (measure) space** if

- Ω is the sample space, that is, some (possibly abstract) set;
- $\mathcal{F}$ is a σ-field of sets (events): the measurable subsets of Ω. The "automs", $\{\omega\}$ of Ω, are called elementary events;
- P is a probability measure on $\mathcal{F}$, that is P satisfies conditions of 1, 2, and 3.

Example 4.1 (Counting measure)**.** Let $\Omega = \{x_1, x_2, ...\}$ be a countable set and $\mathcal{F} = 2^\Omega$ be the total σ-field of all subsets of ω. For $A \in \mathcal{F}$, define $\mu(A) =$the number of elements of A. Then μ is clearly nonnegative and countably additive because the number of elements in the union of disjoint sets in the sum of the numbers of elements in the individual sets. Therefore μ is a measure called **counting measure**, on $\mathcal{F}$. Counting measure is not a probability measure if Ω has more than one element because $\mu(\Omega) > 1$.

Example 4.2 (A probability mass function defines a measure)**.** Let Ω and $\mathcal{F}$ be as defined in the Example 4.1. Suppose that $p_i \geq 0$ and $\sum_{i=1}^{\infty} p_i = 1$. For $A \in \mathcal{F}$, define $P(A) = \sum_{x_i \in A} p_i$. This is well defined because the summand is nonnegative, so we get the same value irrespective of the order in which we add. Then P is nonnegative and countably additive because if I_i and I index the elements of E_i and $E = \cup_j E_j$, resepctively, where $E_1, E_2, ...$ are disjoint, then $P(\cup_j E_j) = \sum_{j \in I} p_j = \sum_i \sum_{j \in I_i} p_j$. Also $P(\Omega) = \sum_{j=1}^{\infty} p_j = 1$. Therefore P is a probability measure on $\mathcal{F}$. This example shows that any probability mass function-the binomial, Poisson, hypergeometric, etc.-specifies a probability measure.

Examples 4.1 and 4.2 involved a countable sample space. The most important measure on the uncountable sample space $\mathbb{R}$ is the following;

Example 4.3 (Lebesgue measure)**.** There is a measure μ_L defined on a σ-field $\mathcal{L}$ containing the Borel sets such that $\mu_L(C) = \text{length}(C)$ is an interval: μ_L is called **Lebesgue measure** and $\mathcal{L}$ are called **Lebesgue sets**. If we restrict attention to Lebesgue sets that are subsets of $[0, 1]$, μ_L is a probability measure. We can think of the experiment as picking a number at random from the unit interval. We will see later that Lebesgue measure on $[0, 1]$ is the only probability measure we ever need.

4.2 The existence of Lebesgue measure

Courses in real analysis spend a good deal of effort proving the existence of Lebesgue measure.

1. Define $\mu(I) = \text{length}(I)$ for any interval I.

2. Now extend μ to an arbitrary set A by defining an **outer measure** μ^* by $\mu^*(A) = \inf \sum_n \text{length}(I_n)$, where the infimum is over all countable unions $\cup_n I_n$ of intervals I_n such that $A \subset \cup_n I_n$. The term outer measure stems from the fact that we are approximating the measure of A from the outside, meaning that we consider the measure of a collection of sets that cover A. It is easy to see that μ^* agrees with μ if A is an interval. But μ^* is not necessarily countably additive on **all sets**, so the next step is:

3. Restrict the class of sets to $\mathcal{L} = \{L : \mu^*(A) = \mu^*(L \cap A) + \mu^*(L^C \cap A)$ for every set $A\}$. The collection $\mathcal{L}$, called Lebesgue sets, is then shown to be a σ-field containing all Borel sets. This restriction of μ^* to $\mathcal{L}$ is Lebesgue measure μ_L.

We are often puzzled that each possible outcome ω has probability 0, yet the probability of $[0, 1]$ is 1. Every event is "impossible" (i.e., has probability 0), yet something is sure to happen.

4.3 Properties of probability measures

Proposition 4.1 (Basic properties of probability measures)**.** *If P is a probability measure:*

1. $P(\emptyset) = 0$.
2. *If* $E_1, E_2 \in \mathcal{F}$, $E_1 \subset E_2$, *then* $P(E_1) \leq P(E_2)$.
3. $0 \leq P(E) \leq 1$ *for each* $E \in \mathcal{F}$.
4. *If* $E \in \mathcal{F}$, $P(E^C) = 1 - P(E)$.
5. *If* $E_1, E_2 \in \mathcal{F}, P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$.

Proof. To see part 1, note that any event E may be written as the disjoint union $E \cup \emptyset$. By countable additivity, $P(E) = P(E) + P(\emptyset)$, from which it follows that $P(\emptyset) = 0$.

To see part 5, note that $E_1 \cup E_2 = E_1 \cup (E_2 \backslash E_1)$. Because E_1 and $E_2 \backslash E_1$ are disjoint,

$$P(E_1 \cup E_2) = P(E_1) + P(E_2 \backslash E_1).$$

Also, E_2 is the union of the disjoint sets $E_1 \cap E_2$ and $E_2 \backslash E_1$, so $P(E_1 \cap E_2) + P(E_2 \backslash E_1) = P(E_2)$. Thus, $P(E_2 \backslash E_1) = P(E_2) - P(E_1 \cap E_2)$. Substituting this expression into the above equation gives $P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$. □

The following result that any countable union can be written as a countable union of disjoint sets is extremely useful in probability theory.

Proposition 4.2 (Any countable union can be written as a countable union of disjoint sets)**.** *Let* $E_1, E_2, ... \in \mathcal{F}$ *and define* $D_1 = E_1, D_2 = E_2 \backslash E_1, D_3 = E_3 \backslash (E_1 \cup E_2), ...$ *Then* $\{D_i\}$ *is a collection of disjoint sets and* $\cup_{i=1}^n E_i = \cup_{i=1}^n D_i$ *for* n *any positive integer or* $+\infty$.

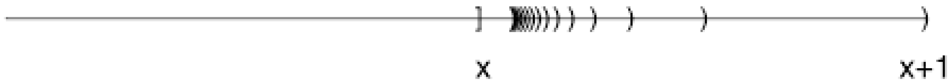

Figure 4.1: An example of descreasing set.

Proof. It is clear that the D_i defined above are disjoint. We will prove that each $\omega \in \cup_{i=1}^n E_i$ is in $\cup_{i=1}^n D_i$ and vice versa. Note that each step in the following proof is valid whether n is a positive integer or $+\infty$. If $\omega \in \cup_{i=1}^n E_i$, then $\omega \in E_i$ for some finite $i \leq n$. Let I be the smallest $i \leq n$ for which $\omega \in E_i$. Then by definition, $\omega \in E_t$ but not in any of the previous E_i. That is, $\omega \in D_I$, so $\omega \in \cup_{i=1}^n D_i$. Thus, if $\omega \in \cup_{i=1}^n E_i$, then $\omega \in \cup_{i=1}^n D_i$.

To prove the converse, note that if $\omega \in \cup_{i=1}^n D_i$, then $\omega \in D_i$ for some finite $i \leq n$. But then $\omega \in E_i$, and hence $\omega \in \cup_{i=1}^n E_i$. We have shown that each $\omega \in \cup_{i=1}^n E_i$ is in $\cup_{i=1}^n D_i$ and vice versa, and therefore $\cup_{i=1}^n E_i = \cup_{i=1}^n D_i$. □

One consequence of Proposition 4.2 is the following, which is called **countable subadditivity**. Ii is also called Boole's inequality, or, when there are only finitely many sets, Bonferroni's ineqaulity.

Proposition 4.3 (Countable subadditivity (also called Boole's or Bonferroni's inequality)**.** *If $E_1, E_2, ...$ are any sets (not necessarily disjoint), then* $P(\cup_i E_i) \leq \sum_i P(E_i)$.

Exercise 4.1. Prove the above proposition.

The importance of Proposition 4.3 in probability and statistics cannot be overstated.

Another consequence of Proposition 4.2 is a certain continuity property of probability, namely that if a sequence of sets $E_1, E_2, ...$ "gets closer and closer" to a set E, then the probability of E_i also gets "closer and closer" to the probability of E. But what does it mean for a sequence of sets $(E_i)_{i=1}^\infty$ to ge "closer and closer" to E?

Definition 4.3 (Increasing and decreasing sets)**.** We say that a sequence of sets $(E_i)_{i=1}^\infty$ increase to $E<$ written $E_i \uparrow E$, if $E_1 \subset E_2 \subset E_3 \subset ...$ and $\cup_{i=1}^\infty E_i = E$. Similarly, $(E_i)_{i=1}^\infty$ decreases to E, written $E_i \downarrow E$, if $E_1 \supset E_2 \supset E_3 \supset ...$ and $\cap_{i=1}^\infty E_i = E$.

For example, $E_i = (-\infty, i] \uparrow E = (-\infty, \infty)$ and $E_i = (-\infty, x + \frac{1}{i}) \downarrow E = (-\infty, x]$ (see Figure 4.1).

5 Applications

Some *significant* applications are demonstrated in this chapter.

5.1 Example one

5.2 Example two

6 Final Words

We have finished a nice book.

Bibliography

Billingsley, P. (2012). *Probability and Measure*. John Wiley and Sons, 3rd edition. ISBN 978-1118341919.

Xie, Y. (2015). *Dynamic Documents with R and knitr*. Chapman and Hall/CRC, Boca Raton, Florida, 2nd edition. ISBN 978-1498716963.