Indian Institute of Technology Bombay
Department of Electrical Engineering

**Handout 8** EE 325 Probability and Random Processes
Lecture Notes 5 August 7, 2014

# 1 Constructing Probability Spaces

We will use our axioms and set properties to paint a larger picture now, i.e. to construct probability spaces in an unambiguous manner. The axioms and properties are kind of the building blocks and glue, magnificent structures can be constructed by proper placement. Here is an algorithmic view of what we are about to do.

1. Have the set of axioms ready as a toolbox.
2. Take an initial seed $C$ of events. The seed is some set of events for which one can unambiguously associate probability values. Many a times this set is a very natural or intuitive choice.
3. Apply the axioms to extend the probability measure defined on the seed to all countable unions and complements of these events.

In the above, the probability values are specified (associated) so as to truthfully model a system we wish to examine. We do not overly worry about the statistical learning and inference done in the past to arrive at a meaningful probability association. Again, what we suggested above is a recipe to construct a probability space. The taste of our preparation will depend on the ingradients used. In particular, a proper choice of the raw ingradient, i.e. seed sets will turn out of atmost importance to get a consistent definition which is stable over allowed operations.

The probability space construction that we describe is similar to building some shape with LEGOs. In LEGO, one needs to be alert about a few things. Firstly, if the LEGO has too many blocks, all the time will be spend in finding the correct pieces, can be a big problem when the blocks are *countably infinite* or more. In fact, the idea of the seed set was to tackle this. We kind of choose the *blocks* which appear *useful* in our construction and start with those. The second problem occurs when some necessary shape cannot be made out of the seed set, in other words, an inadequate seed. So our seed set has to be sufficiently large to accommodate our designs, but not too large that we fail to manage it effectively. One may feel that what we talk is more *managementt* than mathematics, with lot of dangling English words. But this is the real story behind it, and mathematics plugs the gaps and presents these in a coherent manner.

While the modern theory builds the spaces in a systematic fashion, with overheads required to avoid pit-falls, we can compare this first to a naive way of constructing probability spaces, which is kind of classical. This naive way works well when the sample-space $\Omega$ is finite or countable.

## 1.1 Naive Construction

Given a countable $\Omega$ one can in principle specify a probability for every singleton event corresponding to some outcome. Thus one can specify

$$P(\{\omega\}) = p_\omega, \forall \omega \in \Omega.$$

The measure $P(\cdot)$, as mentioned earlier, is determined by the statistics or prior knowledge of the system.

**Example 1** *For a biased coin* $\Omega = \{H, T\}$, *and we can specify*

$$P(H) = p, \, P(T) = 1 - p.$$

*Clearly, we are dealing with the sigma field of* $\mathcal{F} = \{\Omega, \{H\}, \{T\}, \varnothing\}$, *and for each event in* $\mathcal{F}$, *the probability is unambiguous in light of the axioms.*

■

**Example 2** *Consider a biased die, where* $\Omega = \{1, 2, 3, 4, 5, 6\}$. *We can specify*

$$P(\{\omega\}) = p_\omega, \, \omega \in \Omega.$$

*This will again associate probabilities to events in the power-set.*

■

One drawback of this naive association is that we may overdo things, creating unnecessary overheads. For example, if our interest is in just known whether the outcome of a rolled die is even or odd, we need not specify probabilities for every event in the power-set. Rather the four element sigma-field $\mathcal{F} = \{\Omega, A, A^c, \varnothing\}$ will do, where $A = \{1, 3, 5\}$. So we have to specify only $P(A)$. This is a good saving. Imagine when our sample-space is $\mathcal{N}$, and our interest is to know whether the outcome is odd or not. There is immense saving, and many a times easier to specify probability for a small set of events.

So in the modern approach, we will start with a reasonable seed of events for which a probability measure can be associated in a natural and often intuitive way.

## 1.2 Modern Approach

The axiomatic approach starts with a minimal set for which probabilities can be associated, and extends this to relevant events in a sigma-field of interest. The idea is the following.

> "If the seed set $C$ is properly chosen, then one can unambiguously construct the probability space $(\Omega, \mathcal{F}, P)$.

The word *the* is used above to stress the unambiguous nature of the mentioned space. However, this can also lead to a bit of confusion about which $\mathcal{F}$ we are talking about. There are many possible choices for $\mathcal{F}$, but the above procedure is about constructing a sigma-field which contains the seed-set $C$. More precisely, we will construct the smallest sigma-field containing $C$, denoted as $\sigma(C)$, which is also known as the sigma-field generated by $C$.

### 1.2.1 $\sigma-$ field generated by a set

The previous examples show that there are possibly many $\sigma-$fields which contain an event $A$. In practice, our interest may only be in certain set of events, however the axiomatic approach forces us to work with $\sigma-$fields. If your set of interest is $C$, we can indeed choose the *smallest* [1] $\sigma-$field which contains $C$. This gives us a convenient $\sigma-$field with the minimum amount of *book-keeping*. We denote this $\sigma-$field as $\sigma(C)$ and call this the **$\sigma-$field generated by** $C$. What guarantees that there is such a $\sigma-$field?

[1] here small is defined by set-inclusion

**Theorem 1** *There exists a unique $\sigma(C)$.*

*Proof:* Let $\sigma_1(C)$ and $\sigma_2(C)$ are two $\sigma-$fields containing $C$. Then, $\sigma_1(C) \bigcap \sigma_2(C)$ is also a $\sigma-$field. Consider a countable collection of sets $A_i, i \in S$ such that $A_i \in \sigma_1(C) \bigcap \sigma_2(C), \forall i \in S$. Clearly $A_i \in \sigma_1(C), \forall i$ and $A_i \in \sigma_2(C), \forall i$. Thus $\bigcup A_i$ is in $\sigma_1(C)$ as well as $\sigma_2(C)$, as they are sigma-fields by definition. Thus the intersection of sigma-fields is a sigma-field, and $\sigma(C)$ is the intersection of all $\sigma-$fields containing $C$.

■

In order to list the benefits of working with seed-sets which are natural and easy to handle, let us consider the infinite coin toss example of a biased coin.

**Example 3** *Construct a meaningful probability space for the countable toss of a biased coin. Different tosses are done independently*[2]*. It is also given that $P(HEADs) = p$.*

**Solution:** Let us first recall the sample-space.

$$\Omega = \{(\omega_1, \omega_2, \cdots) : \omega_i \in \{H, T\}\}.$$

Let us generate a seed-set. A very natural choice to look at events in the first $k$ tosses. For any $k-$length sequence $(t_1, t_2, \cdots, t_k)$, $t_i \in \{H, T\}$,

$$E_k^t = \{(\omega_1, \omega_2, \cdots) : (\omega_1, \cdots, \omega_k) = (t_1, \cdots, t_k)\}.$$

Here, the superscript $t$ denotes that it is the set of all outcomes of $\Omega$ such that the first $k$ tosses equal $t = (t_1, \cdots, t_k)$. Let us assign a probability measure to event $E_k$. Clearly, from our 'experience', the probability that the outcome is a sequence of all HEADs is

$$P(H, H, \cdots, H) = p^k.$$

For sequence $t$ with some HEADs as well as TAILs,

$$P(E_k) = p^{N_H^t}(1-p)^{N_T^t},$$

where $N_H^t$ counts the number of HEADs in the sequence $(t_1, t_2, \cdots, t_k)$ and $N_T^t = k - N_H^t$. A concise way to represent $N_H^t$ is

$$N_H^t = \sum_{i=1}^{k} \mathbb{1}_{\{t_i = H\}}.$$

For convenience let us call $\mathbb{1}_{\{t_i=H\}}$ as $d_i$.

Does the above probability measure yield $P(\Omega) = 1$, this is the first check that you should do once it is clear that the measure is always non-negative. Notice that $\Omega = \bigcup E_k^t$, where the union is over all $k-$length sequences $t$. Indeed any outcome in $\Omega$ has to start with some $k-$length sequence, and thus will belong to the corresponding $E_k^t$. Observe also that for any given $k$, $E_k^t$ are mutually disjoint sets.

$$P(\bigcup_{t \in \{H,T\}^k} E_k^t) = \sum_t P(E_k) = \sum_t p^{N_H^t}(1-p)^{N_T^t} \tag{1}$$

$$= \sum_t p^{\sum_{i=1}^k d_i}(1-p)^{\sum_{i=1}^k (1-d_i)} \tag{2}$$

$$= \sum_t \prod_{i=1}^{k} p^{d_i}(1-p)^{1-d_i} \tag{3}$$

[2] rigorous idea to be introduced later

The above summation can be split into two sums, where the first one runs over all sequences of length $k-1$, which we denote as $t-1$. The second sum is over values of $t_k$ which can be either $H$ or $T$.

$$\sum_t \prod_{i=1}^{k} p^{d_i}(1-p)^{1-d_i} = \sum_{t-1}\sum_{t_k}\prod_{i=1}^{k} p^{d_i}(1-p)^{1-d_i} \tag{4}$$

$$= \sum_{t-1}\sum_{t_k}\prod_{i=1}^{k-1} p^{d_i}(1-p)^{1-d_i} p^{d_k}(1-p)^{1-d_k} \tag{5}$$

$$= \sum_{t-1}\prod_{i=1}^{k-1} p^{d_i}(1-p)^{1-d_i} \sum_{t_k} p^{d_k}(1-p)^{1-d_k} \tag{6}$$

There is no magic in any of the steps above. We have moved the summation to the appropriate variables, that is it. Continuing,

$$\sum_t \prod_{i=1}^{k} p^{d_i}(1-p)^{1-d_i} = \sum_{t-1}\prod_{i=1}^{k-1} p^{d_i}(1-p)^{1-d_i}(p^1(1-p)^0 + p^0(1-p)^1) \tag{7}$$

$$= \sum_{t-1}\prod_{i=1}^{k-1} p^{d_i}(1-p)^{1-d_i}(p+1-p) \tag{8}$$

$$= \sum_{t-1}\prod_{i=1}^{k-1} p^{d_i}(1-p)^{1-d_i} \tag{9}$$

Repeating the same steps as above, the sequence length is decremented in each step, and this shows that $P(\Omega) = 1$.

■

This is a very useful probability measure and can model a wide variety of experiments. Students are advised to go through this in full detail. Do not fail to notice the fact that the above probability association did not specify probabilities for any outcome in $\Omega$, which is actually an infinite sequence of Heads and Tails. Nevertheless, it will turn out that our specification for finite $k$ has all the ingredients and recipe to answer almost all questions related to the coin toss experiment.

**Exercise 1** *Induct the argument in the previous example to show that indeed $P(\Omega) = 1$.*

The above example underlines the enormous ease that the concept of seed-sets will provide, if chosen wisely. That a wise choice of the starting see is important will be exemplified in the next subsection. Recall that our job is to construct a probability measure valid for every event of $\sigma(C)$. If the collection $C$ is not adequate enough we may not be able to unambiguously specify $P(A)$ for all $A \in \sigma(C)$.

# 2 Uniqueness of Measure

So far our strategy was to associate probabilities to a sensible set $C$ and then extend it using the axioms. How do we know that our starting seed-set $C$ will not lead to any ambiguity. For example, a probability measure defined on $C$ need not uniquely specify a probability measure on $\sigma(C)$. In that case, our strategy of extension from the seed-set fails. More precisely, there can be multiple probability measures on $\sigma(C)$ which will give identical probability association for elements in $C$. Let us illustrate it by a popular example (Fristedt and Gray, *A Modern Approach to Probability*, Birkauser 1997).

**Example 4** *Let $\Omega = \{a, b, c, d\}$. Consider $C = \{C_1, C_2\}$ where $C_1 = \{a, b\}$ and $C_2 = \{b, c\}$. It can be easily verified that $\sigma(C) = \mathcal{P}_\Omega$, the power-set. Associate probabilities to $C$ and check whether there is a unique extension to $\sigma(C)$.*

**Solution:** Consider a probability $\hat{Q}$ defined on $\hat{Q}(\Omega)$ such that $\hat{Q}(\{b\}) = \hat{Q}(\{d\}) = \frac{1}{2}$. Certainly this is a valid association, and will yield

$$\hat{Q}(C_1) = \hat{Q}(C_2) = \frac{1}{2}.$$

Now consider an alternate probability measure $Q(\cdot)$ defined on $\Omega$ such that $Q(\{a\}) = Q(\{c\}) = \frac{1}{2}$. Using this,

$$Q(C_1) = Q(C_2) = \frac{1}{2}.$$

Thus both $\hat{Q}$ and $Q$ defined on $\mathcal{P}_\Omega$ has identical restriction to the set $C$. Conversely, associating $C_1$ as well as $C_2$ with an identical probability of $\frac{1}{2}$ is not enough and can yield ambiguous probabilities for events in $\sigma(C)$. As an example

$$Q(\{a, c\}) = 1 \text{ and } \hat{Q}(\{a, c\}) = 0.$$

■

The whole purpose of our axiomatic upbringing was to avoid these kind of ambiguities. Loosing it because of a bad choice of the seed-set $C$ is like *breaking the pot after filling it.* Fortunately, the **uniqueness of measure theorem** along with an extension tells us a simple, but powerful way to avoid these paradoxes.

**Theorem 2** *(Extension of Uniqueness of Measure) A probability association defined on $C$ has a unique extension to $\sigma(C)$, if $C$ is closed under pairwise intersections and complements.*

> While the uniqueness of measure theorem does not need complements, the existence of an extension requires the presence of complementation.

For your reference, I am pasting the so called Extension Theorem from measure theory below. Recall that $C$ is a field, if it is closed under pairwise intersections and complementation.

**Theorem 3** *Let $C$ be a field of subsets of a space $\Omega$ and $Q$ a non-negative countably additive function defined on $C$ such that $Q(\Omega) = 1$. Then there exists a unique probability measure $P$ defined on $\sigma(C)$ such that $P(A) = Q(A)$ for every $A \in C$.*

The theorem as such does not state the necessity of the mentioned conditions. In particular, it can be relaxed in several ways. In the class, we stated that it is sufficient for $C$ to be a semi-algebra.

**Definition 1** *A collection of sets $C$ defined on $\Omega$ is a semi-algebra, if*

- *$C$ is closed w.r.t. pairwise intersections.*
- *For every $A \in C$, the complement $A^c$ is a finite union of disjoint events in $C$.*
- *$C$ contains $\Omega$.*

It is perhaps true that many students felt the requirement of $C$ being a field is very restrictive. Recall that we only demanded the probability measure to be specified for a field $C$. This does not mean we have to enumerate for every element in $C$. There are known ways of starting from a manageable level of seedsets and then extending the probability measure to $C$. In the upcoming example in next section, you will realize that choosing a good $C$ for many cases of interest to us is not as complicated as it may sound. After specifying the probability on $C$, one then extends it to $\sigma(C)$. Please take any standard book on real analysis, our objective was to tell you the essence of the approach, than being exhaustive (S. Resnick, *A Probability Path*, Birkhauser, has accessible explanations).

# 3 Uniform on the Square

Let us consider the real interval in $\mathbb{R}^2$, which is an exciting space for many applications. The uniform choice on $\Omega = [0,1]^2$ may seem trivial, but you have to go through it in a step by step fashion, clearly understanding each step. This procedure is insisted because certain key features of this problem are immensely useful in building more complex systems.

**Example 5** *Consider a square placed on the non-negative quadrant with one corner at the origin. If a point is uniformly chosen, compute the probability that the chosen point $(x,y)$ satisfies $x+y \leq 1$.*

**Solution:** Surely this is a joke, we are asking for a point being in the diagonal half of the square, what else other than $\frac{1}{2}$ could the probability be. Sure, but are there Bertrand's paradoxes? Since any point will *never* be taken, saying a point to be uniformly taken itself can be ambiguous. Let us walk the steps than taking the elevator. The sample-space is

$$\Omega = \{(x,y) \in \mathbb{R}^2 : 0 \leq x \leq 1; 0 \leq y \leq 1\}.$$

What events can you choose for a seed set? Since we are dealing with $\mathbb{R}^2$, it is natural to look for two-dimensional sets, perhaps the *simplest*. Seriously, it is not a circle, or an ellipse, or a corrugated cloud. It indeed is a rectangle, For any rectangle $R_{abcd}$ of diagonal coordinates $(a,c)$ and $(b,d)$ with $d \geq c$, we can associate a probability in a very sensible fashion.

$$P(R_{abcd}) = (b-a)(d-c).$$

If the area of the sample-space is not unity, the above probability can be divided by the total area covered by $\Omega$. Now let $T$ be the bottom triangle obtained by diagonally cutting the square from top-left to right-bottom. Our event of interest is whether the outcome is in $T$ or not. To compute $P(T)$, we have to express $T$ as a countable sum of events of the form $R_{abcd}$. There is nothing holy about this, we know how to handle rectangles, their unions and intersections, very very well. In fact, in this case, we can write $T$ as a union of mutually disjoint rectangles. The idea is to divide $T$ first into a square and two identical triangles as shown in Figure 1. The new square will have half the area of the triangle, and each small triangle will have half the area of the square.

We will recurse this procedure to divide every remaining triangle in the picture, with the larger triangles broken first. The procedure will allow the rectangles to approximate the square, see figure below for the $4^{th}$ stage (in blue, more precisely, when each triangle present after this stage has already undergone 4 divisions, or it is the $4^{th}$ generation child triangles) and the $6^{th}$ stage(red) of triangle division.

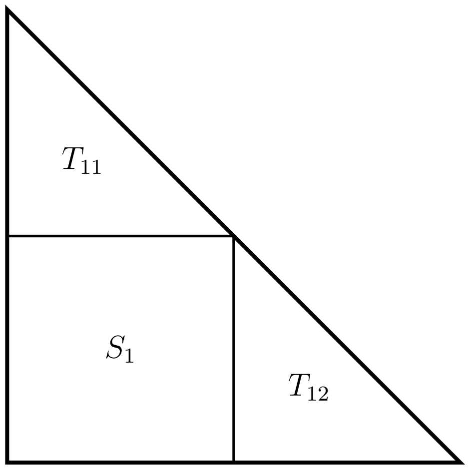


Figure 1: Dividing the triangle to a square $S_1$ and two smaller triangles ($T_{11}$ and $T_{12}$)

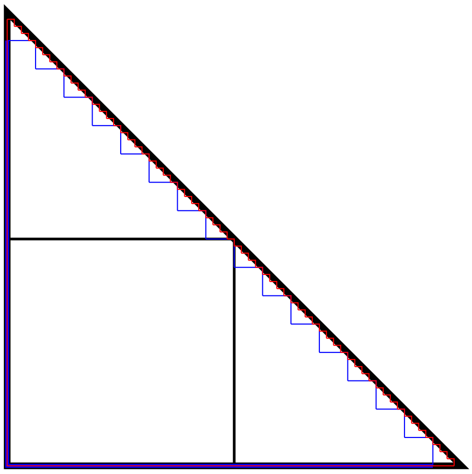

Figure 2: Approximating a triangle by squares

Thus, the triangle is a union of a countable number of rectangles. By the second probability axiom,

$$P(T) = \lim_{n\to\infty} \frac{1}{2^n} \sum_{i=1}^{2^n-1} \frac{i}{2^n} \quad (10)$$

$$= \frac{1}{2}. \quad (11)$$

While we have summed the rectangles generated by the staircase waveform shown in Figure 2, one can alternately sum over appropriate squares contained inside the triangle.

# 4 Continuity of Measure

Our last example showed how to construct a triangle by countable union of rectangles. The strategy there was to consider non-overlapping rectangles and add the corresponding well-defined probabilities. In fact, at each stage of our division strategy, a new staircase waveform (see last figure in lecture notes 4b) which approximates the triangle better is generated. As the enclosure made by the the stair-case and the quadrants tends to a triangle, the probability measure of the stair-case tends progressively to that of the triangle. For those who have done analysis, this is in fact similar to the notion of continuity. We will write $A_n \uparrow A$ to denote the limit of a sequence of sets where $A_{n-1} \subseteq A_n, \forall n$. Often, we term such a sequence as *monotone*.

**Property 1** *(Continuity of Measure) Let* $(\Omega, \mathcal{F}, P)$ *be a probability space. If* $A_n \uparrow A, A_n \in \mathcal{F}, \forall n$ *then* $P(A_n) \uparrow P(A)$.

**Proof:** Before proving, notice that the above property can be rephrased as "the measure $P(\cdot)$ is continuous for monotone sequences $A_n$". The proof uses the idea that we had for the last example, write $A_n$ as a union of disjoint sets. Since the sequence is monotone, $\forall n$

$$A_1 \subseteq A_2 \subseteq \ldots \subseteq A_n.$$

Let $B_i = A_i \setminus A_{i-1}$, then $A_n = \bigcup_{i=1}^n B_i$. We have,

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

# 5 Some Extra Lessons

So far, our strategy of choosing simple seed-sets have given us rich dividends. The probability measure that we specify for the seed-set $C$ can be extended (by axioms) to countable unions and complements of the events in $C$, see Figure 3.

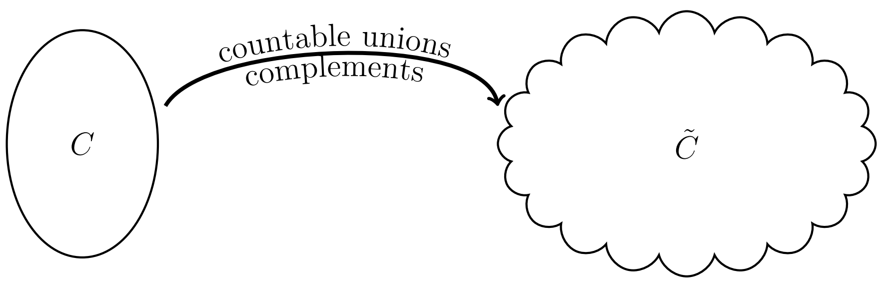


Figure 3: Extension of $C$ by Axioms may not be enough

However, the second entity of our space needs to be a sigma-field for the robustness of construction. Will countable unions and complements of $C$ generate a sigma-field? While this works when $\Omega$ has finite cardinality, it turns out to be insufficient in general. So we have to add more objects to obtain a sigma field. Unfortunately, there is no constructive procedure to add the missing objects, we kind of fill it to the extend possible by known manipulations and leave the rest.

In any case, it is natural to consider the minimal sigma-field, denoted as $\sigma(C)$, also termed as the $\sigma$−field generated by $C$. Why take extra load? To our advantage, it can be shown that there is a unique minimal sigma-field for any set $C$.

**Proposition 1** *For any collection of subsets $C$, there exists a unique sigma field $\sigma(C)$ such that if $\mathcal{G}$ is any sigma-field containing $C$, then $\sigma(C) \subseteq \mathcal{G}$.*

We have already proved the proposition in Lecture Notes 3. Thus,

$$\sigma(C) = \bigcap_{t \in \Theta} \sigma_t(C),$$

where $\Theta$ indexes all sigma-fields $\sigma_t(C)$ having $C$ inside it.

# 6 Borel Fields

Borel Fields are the minimal sigma-fields containing all generalized rectangles in $\mathbb{R}^n$. Recall the previous example, where $\Omega$ was a square in $\mathbb{R}^2$. There we took the sub-rectangles and associated a probability to each and every rectangle. The measure then was extended to triangles, the same can be done for hexagons, circles or any other sensible shapes.

Let us first consider $\mathbb{R}$. A natural choice of seed-set in $\mathbb{R}$ are open sub-intervals of the form $(a,b) : a < b$. We know that a sub-interval $(a,b)$ is also an open-set. The sigma field which contains all the open sets are known as Borel Field, denoted as $\mathcal{B}(\mathbb{R})$. While open-sets is a stable notion than open sub-intervals in general topological spaces, open sub-intervals of the form $(a,b)$ are much more easy to visualize and manipulate in $\mathbb{R}^n$. In fact, for Cartesian coordinates, the generated sigma-fields by both notions can be shown as the same, i.e.

$$\sigma(\text{ open sets }) = \sigma(\{(a,b), -\infty < a \leq b < \infty\}).$$

We do not list a complete proof here, rather use a well known result that an open set $A^o$ can be expressed as a countable union of disjoint open sub-intervals.

$$A^o = \bigcup_{i \geq 1} A_i,$$

where $A_i \bigcup A_j = \emptyset, j \neq i$ and $A_i = (a_i, b_i)$. Thus any open-set has to be present in $\sigma((a,b))$ and consequently

$$\sigma((a,b)) \supseteq \mathcal{B}(\mathbb{R}).$$

But any open-subinterval is an open-set and hence

$$\sigma((a,b)) \subseteq \mathcal{B}(\mathbb{R}).$$

The last two equations show that $\sigma((a,b)) = \mathcal{B}(\mathbb{R})$, which is the desired result.

Similarly, we denote by $\mathcal{B}(\mathbb{R}^n)$ the Borel field generated by open sub-intervals in $\mathcal{B}(\mathbb{R}^n)$. Why we take open-sets, is $\mathcal{B}(\mathbb{R})$ different from the ones generated by closed rectangles? The answer is NO, we will explain this in the following examples.

**Example 6** *Show $\sigma(C) = \mathcal{B}(\mathbb{R})$, when $\Omega = \mathbb{R}$ and $C = \{(a,b], -\infty \leq a \leq b < \infty\}$.*

**Solution:** We already know that $\mathcal{B}(\mathbb{R}) = \sigma((a,b))$. We will show that each open sub-interval can be generated by *allowed* operations on closed sub-intervals and vice-versa. For any open sub-interval $(a,b)$

$$\bigcup_{i=n^*}^{\infty} [a + \frac{1}{n}, b] = (a,b],$$

where $n^* > \frac{1}{b-a}$ ensures that each set in the countable union is contained in $(a,b]$. Thus $\sigma(C) \subseteq \mathcal{B}(\mathbb{R})$. On the other hand,

$$\bigcap_{i\geq 1}(a-\frac{1}{n},b] = [a,b],$$

showing $\mathcal{B}(\mathbb{R} \subseteq \sigma(C)$ as well. Putting these together

$$\sigma(C) = \mathcal{B}(\mathbb{R}).$$

■

**Example 7** *Show $\sigma(C) = \mathcal{B}(\mathbb{R})$, when $\Omega = \mathbb{R}$ and $C = \{[a,b], a \leq b\}$.*

**Solution:** This is identical to the previous example. For any semi-open interval $(a,b]$

$$\bigcup_{i=n^*}^{\infty}(a,b-\frac{1}{n}] = (a,b),$$

where $n^* > \frac{1}{b-a}$ ensures that each set in the countable union is contained in $(a,b]$. Thus, in light of the previous example, $\sigma(C) \subseteq \mathcal{B}(\mathbb{R})$. On the other hand,

$$\bigcap_{i\geq 1}(a-\frac{1}{n},b] = [a,b],$$

showing $\mathcal{B}(\mathbb{R} \subseteq \sigma(C))$ as well. Putting these together

$$\sigma(C) = \mathcal{B}(\mathbb{R}).$$

■

We now show another very important example, that this will be part of our daily staple of probability in future.

**Example 8** *Show $\sigma(C) = \mathcal{B}(\mathbb{R})$, when $\Omega = \mathbb{R}$ and $C = \{(-\infty,x], x \in \mathbb{R}\}$.*

**Solution:** We will show one side of the proof and leave the remaining as an exercise. Observe that

$$(a,b] = (-\infty,b]\bigcap(-\infty,a]^c,$$

where the complement used in the equation gives a semi-infinite open interval, i.e. $(-\infty,a]^c = (a,+\infty)$. Thus

$$\sigma(C) \supseteq \mathcal{B}(\mathbb{R}).$$

See exercise below for completing the proof. ■

**Exercise 2** *Given $C = \{(-\infty,x], x \in \mathbb{R}\}$, show that $\sigma(C) \subseteq \mathcal{B}(\mathbb{R})$.*