Weekly Notes for EE2012 2014/15 – Week 2

T. J. Lim

January 21, 2015

Book sections covered this week: 2.2 and 2.3

1 Probability Laws for Discrete and Continuous Sample Spaces

While the Axioms of Probability give us the properties that a probability mapping (or law) needs to have, it is still not clear how to go about assigning probability values to every event in $\mathcal{F}$, given that $\mathcal{F}$ can be very large.

In this section, we will discuss countable or discrete sample spaces separately from uncountable or continuous ones. A countable set is one whose elements can be "counted" – to be precise, a set whose elements can be mapped one-one to the set of counting numbers $\{1, 2, \ldots\}$. All other sets are uncountable.

Example 1: The following sets are countable –

- The set of all integers, $\mathbb{Z}$;
- The set of prime numbers;
- The set of even numbers;
- Any set with a finite number of elements, e.g. $\{1, 2, 3\}$, $\{x, y, z\}$.

The following sets are uncountable –

- The set of complex numbers, $\mathbb{C}$;
- All real numbers in a finite range, e.g. $(1, 2]$, $(0, 1) \cup (5, 7)$, etc.;
- All points within the unit disc, i.e. $\{(x, y) : x^2 + y^2 \leq 1\}$;
- All vectors within a linear vector space, e.g. $\{\mathbf{z} : \mathbf{z} = a\mathbf{x} + b\mathbf{y}, a, b \in \mathbb{R}\}$ where $\mathbf{x}$ and $\mathbf{y}$ are given.

1.1 Probability Laws for Discrete Sample Spaces

Let S be a discrete sample space, and let all subsets of S be events in $\mathcal{F}$. If we denote the elements of S as a_k, $k = 1, 2, \ldots, |S|$, then $\{a_1\}, \{a_2\}, \ldots, \{a_{|S|}\}$ are the "elementary events", i.e. events that cannot be further divided into the union of smaller events. Suppose we assign probability values to every elementary event, $p_k = P[\{a_k\}]$, $k = 1, 2, \ldots, |S|$. Note that any event $A \in \mathcal{F}$ comprises a set of outcomes, e.g. $A = \{a_1, a_2\} = \{a_1\} \cup \{a_2\}$, therefore by Axiom III the probability of A is the sum of the probabilities of its component elementary events e.g.

$$A = \{a_1, a_2\} = \{a_1\} \cup \{a_2\} \Rightarrow P[A] = p_1 + p_2. \tag{1}$$

A probability law for a discrete sample space is therefore obtained once all elementary event probabilities are determined or assigned.

Example 2: Toss a coin three times and note the sequence of heads and tails. The elementary events are $\{HHH\}$, $\{HHT\}$, etc. A valid probability law is to have equal probability values ($= 1/8$) for each of them. This happens to also be the right model if the coin is fair. Then the probability of any event will be the sum of the elementary event probabilities, e.g.

$$P[\text{“Two heads”}] = P[\{HHT, HTH, THH\}] = \frac{3}{8} \tag{2}$$

$$P[\text{“At least two heads”}] = P[\{HHT, HTH, THH, HHH\}] = \frac{1}{2}. \tag{3}$$

For equi-probable outcomes, the probability of an event A is computed as $|A|/|S|$, where $|A|$ is the cardinality of (or number of elements in) A. This is why it is often important to be able to count the number of outcomes belonging to an event. ■

1.2 Probability Laws for Continuous Sample Spaces

When S is continuous, it is usually a subset of $\mathbb{R}^n$, the n-dimensional real space consisting of all vectors $(x_1, x_2, \ldots, x_n)$, $x_i \in \mathbb{R}$. For now, we consider only the $n = 1$ case, i.e. $S \subset \mathbb{R}$.

Unlike the discrete case, almost all[1] elementary events $\{a\}$ where $a \in S$ have zero probability due to the richness of the sample space. Therefore it is insufficient (and also impractical) for us to specify the probabilities of elementary events when S is continuous in creating a probability law. We also need to restrict ourselves to event fields containing unions, intersections and complements of $(a, b]$, where $a \leq b$. This "restriction" is of no practical consequence since the only sets excluded are those that do not exist in the real world, only in the mathematical one.

If we can find a function

$$F(x) = P((-\infty, x]) \tag{4}$$

[1]This is a technicality that will be explained soon, after we introduce random variables. Some elementary events in a continuous S can have non-zero probability.

or

$$F^c(x) = P([x, \infty)) \tag{5}$$

then all event probabilities can be found, since $P[(a,b]] = F(b) - F(a)$. Therefore the above functions, respectively called the cumulative distribution function (CDF) and the complementary CDF, are of fundamental importance. We will say a lot more about the CDF in future classes.

Example 3: Suppose the sample space is $\mathbb{R}$, and we have the distribution function

$$F(x) = \begin{cases} 1 - e^{-x}, & x > 0 \\ 0 & \text{otherwise} \end{cases}. \tag{6}$$

We can find probabilities of the following events:

$$\begin{aligned}
P[(2,\infty)] &= 1 - P[(-\infty, 2]] = e^{-2} \\
P[(1,2]] &= P[(-\infty, 2]] - P[(-\infty, 1]] \\
&= e^{-1} - e^{-2} \\
P[(-2,3)] &= P[(-2,0]] + P[(0,3]] - P[\{3\}] \\
&= 0 + 1 - e^{-3} - 0 \\
&= 1 - e^{-3}.
\end{aligned}$$

2 Counting Methods

In many important discrete scenarios, such as tossing a fair coin, rolling a fair die, picking a sequence of balls from an urn, each elementary event has the same probability of $|S|^{-1}$. In these cases, the probability of an event A is simply

$$P[A] = \frac{|A|}{|S|}.$$

It is therefore important to be able to count the number of elements in a set, and most such exercises can be reduced to variations on an urn experiment. This is simply an experiment in which n balls labelled 1 through n are placed in an urn, and we pick k times from the urn, either with or without replacement after each pick, and with or without noting the order of appearance.

2.1 Sampling With Replacement With Ordering

In this case, each of the k picks can result in any of the n balls. Therefore the set of unique k-tuples is

$$\{(x_1, \ldots, x_k) : x_i \in \{1, 2, \ldots, n\}\} \tag{7}$$

and we can see that there are n^k elements in this set.

2.2 Sampling Without Replacement With Ordering

A tree diagram helps to visualize this experiment, with the i-th level of the tree representing the i-th ball picked from the urn. The set of possible k-tuples excludes repeated entries (and is therefore smaller than n^k), and can be written as

$$\{(x_1, \ldots, x_k) : x_i \in \{1, 2, \ldots, n\}, x_i \neq x_j \text{ for all } i \neq j\}. \tag{8}$$

Counting the number of paths from the root of the tree to the leaves, we see that there are a total of

$$^nP_k = \frac{n!}{(n-k)!}$$

unique k-tuples.

2.3 Sampling Without Replacement Without Ordering

When the order or sequence of picks is not important, only the identities of the balls picked, this is akin to taking the k picked balls and placing them in another container, and looking at the contents of the container at the end of the trial. Then we will count, say, $(1, 2)$ and $(2, 1)$ only one time.

The number of possible ways to fill the second container will be smaller than the number of unique k-tuples in the last section (without replacement with ordering) by a factor of $k!$, the number of ways to arrange k distinct objects. This number is known as the binomial coefficient:

$$\binom{n}{k} = \frac{^nP_k}{k!} = \frac{n!}{(n-k)!k!}. \tag{9}$$

An important application of this result is in counting the number of ways we can "line up" k objects of one type and $n-k$ objects of another type, say black balls and white balls, heads and tails, ones and zeros, etc. Imagine an urn with n labelled balls, and sample from it k times without replacement and without ordering. Place the k Type 1 objects in the positions indicated by the k balls picked. The total number of possible arrangements is therefore the number of ways to pick k balls from n without considering their order of appearance, i.e. $\binom{n}{k}$.

The extension to the scenario of having J different types of objects, Type 1 through Type J, is as follows. Let there be k_i objects of Type i, $i = 1, \ldots, J$, and let $n = \sum_{i=1}^{J} k_i$. An urn experiment to line these objects up is:

1. Pick k_1 balls without replacement from the n that we start with, place the Type 1 objects in those positions indicated by the numbers on the balls;

2. Pick k_2 balls from the remaining $n - k_1$, and place the Type 2 objects in those positions;

3. Carry on until the there are only k_J balls left in the urn.

Now there are $\binom{n}{k_1}$ arrangements arising from the first step, $\binom{n-k_1}{k_2}$ arrangements from the second, etc. In total, there will be

$$\binom{n}{k_1}\binom{n-k_1}{k_2}\cdots\binom{k_J}{k_J} = \frac{n!}{k_1!\cdots k_J!} \tag{10}$$

object arrangements. This is known as the multi-nomial coefficient.

Example 4: What is the probability that among 12 randomly selected people, 5 are women and 7 are men?

Ans: We assume that it is equally likely for a randomly selected person to be male or female. The outcome of this experiment is the 12-tuple $(x_1, \ldots, x_{12})$ where $x_i \in \{\text{M,W}\}$, so the sample space S has cardinality 2^{12}. Each outcome is equally likely. From the discussion in this section, we know that of these 2^{12} outcomes, $\binom{12}{5}$ contain 5 W and 7 M. In other words, the event "5 women and 7 men" contains $\binom{12}{5}$ outcomes. Therefore it has the probability

$$\frac{\binom{12}{5}}{2^{12}} = 0.1933. \tag{11}$$

2.4 Sampling With Replacement Without Ordering

This case tends not to arise too often in our work and so we can ignore it. Interested readers can refer to Section 2.3.5 of the book.