Lecture Notes 2: Probability and Statistics

In the previous lecture note we have seen that a probability function is a set function and it is defined on the subsets of a sample space. Further, it satisfies three properties. We have also seen that if the number of elements in a sample space is either finite or countable, then how a probability function can always be defined on all subsets of the sample space. But the same may not be true if the sample space is not either finite or countable. It means suppose we have the sample space $\Omega = [0, 1]$, and we want to define a probability function on all subsets of Ω,it may not be always possible. In fact it can be shown (let us accept this fact in this course, because the actual illustration is not very easy) that suppose we want to define an uniform probability function, i.e. $P([a, b]) = b - a$, for any $0 \leq a < b \leq 1$, then it is not possible to extend it to all subsets of $\Omega = [0, 1]$.

Due to this reason we would like to introduce a new class of subsets of Ω on which it is always possible to define a set function so that it satisfies the properties of a probability function. This class is known as σ-field.

Definition: A class of subsets of Ω is said to be a σ-field, and it will be denoted by $\mathcal{F}$, if it satisfies the following three properties.

1. $\Omega \in \mathcal{F}$.
2. If $A \subset \Omega$ and $A \in \mathcal{F}$, then $A' \in \mathcal{F}$.
3. If $A_1, A_2, \ldots$ are subsets of Ω and $A_i \in \mathcal{F}$, for all $i = 1, 2, \ldots$, then $\displaystyle\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Note that while we are defining the σ-field there is any notion of probability is needed. But it will help us in defining probability function later. Some of the examples of σ-field are as follows:

Example 1: The trivial σ-field is $\mathcal{F} = \{\Omega, \phi\}$. Sometimes it is known as the smallest σ-field.

Example 2: The class of all subsets of Ω is clearly a σ-field. It is also known as the largest σ-field.

Example 3: If A is any subset of Ω, then $\mathcal{F} = \{\Omega, \phi, A, A'\}$ is a σ-field.

Exercise Please verify that the above three $\mathcal{F}$'s are σ-field.

We need one more definition. Suppose $\mathcal{G}$ is a class of subsets of Ω, not necessarily a σ-field, then $\mathcal{F}$ is called the minimum σ-field contains $\mathcal{G}$ if it satisfies the following properties:

1. $\mathcal{F}$ is a σ-field and $\mathcal{G} \subset \mathcal{F}$.
2. $\mathcal{F}_s$ is any other σ-field such that $\mathcal{G} \subset \mathcal{F}_s$, then $\mathcal{F} \subset \mathcal{F}_s$.

Example: Suppose $\mathcal{G} = \{A, \phi\}$. Clearly $\mathcal{G}$ is NOT a σ-field. The minimum σ-field contains $\mathcal{G}$ is $\mathcal{F} = \{\Omega, \phi, A, A'\}$.

Comment: Note that if $\mathcal{G}$ is a σ-field, then the minimum σ-field contains $\mathcal{G}$ is $\mathcal{G}$ itself.

Now we will define one important σ-field which plays a very important role in defining probability. It is known as Borel σ-field, and it can be described as follows. Suppose $\Omega = (-\infty, \infty)$, and $\mathcal{G} = \{(a, b); -\infty < a < b < \infty\}$. Then the minimum σ-field which contains $\mathcal{G}$ is known as the Borel σ-field. Note that the Borel σ-field contains all the intervals of the type (a, b), all the union of intervals, their complements and many such sets. Now the most general way we can define a probability function when Ω is $(-\infty, \infty)$ is to define the set function only on $\{\mathcal{G}\}$ and make sure it satisfies all the properties of a probability function if A_i's are elements of $\mathcal{G}$. Then we have a very interesting result which tells us how it can be extended for all the elements of a Borel σ-field. Since it is not possible to demonstrate it in this course we will avoid it. Now we will show how we can define probability function on the elements of $\mathcal{G}$.

Example: Let $\Omega = (0, 1)$, and $\mathcal{G} = \{(a, b); 0 < a < b < 1\}$. Let us define $P((a, b)) = (b - a)$ for any $0 < a < b < 1$, and if $I_1, I_2, \dots$ are disjoint intervals of $(0,1)$, then $P\left(\bigcup_{I=1}^{\infty} I_i\right) = \sum_{i=1}^{\infty} P(\text{length of the interval } I_i)$. Since the P satisfies all the properties of a probability function on the elements of $\mathcal{G}$, therefore it is a probability function on the Borel σ-field.

Example: Let $\Omega = (0,2)$, and $\mathcal{G} = \{(a,b); 0 < a < b < 2\}$. Let us define $P((a,b)) = \frac{(b-a)}{2}$ for any $0 < a < b < 2$, and if $I_1, I_2, \ldots$ are disjoint intervals of $(0,2)$, then $P\left(\bigcup_{I=1}^{\infty} I_i\right) = \sum_{i=1}^{\infty} P(\text{length of the interval } I_i)$. Since the above P satisfies all the properties of a probability function on the elements of $\mathcal{G}$, therefore it is a probability function on the Borel σ-field.

Counter Example: Let $\Omega = (0,2)$, and $\mathcal{G} = \{(a,b); 0 < a < b < 2\}$. Let us define $P((a,b)) = \sqrt{\frac{(b-a)}{2}}$ for any $0 < a < b < 2$, and if $I_1, I_2, \ldots$ are disjoint intervals of $[0,2]$, then $P\left(\bigcup_{I=1}^{\infty} I_i\right) = \sum_{i=1}^{\infty} P(\text{length of the interval } I_i)$. Clearly this is NOT a probability function as in one way $P((0,2)) = 1$ and if we write $(0,2) = (0,1) \cup (1,2)$, then $P(0,2) = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}$. Therefore, the above P is NOT a probability function.

Example: Let $\Omega = (0,2)$, and $\mathcal{G} = \{(a,b); 0 < a < b < 2\}$. Let us define

$$P((a,b)) = \int_a^b x dx,$$

for any $0 < a < b < 2$, and if $I_1, I_2, \ldots$ are disjoint intervals of $(0,2)$, then

$$P\left(\bigcup_{I=1}^{\infty} I_i\right) = \int_{I_1 \cup I_2 \ldots} x dx.$$

Since the above P satisfies all the properties of a probability function on the elements of $\mathcal{G}$, therefore it is a probability function on the Borel σ-field.

Example: Let $\Omega = (0, \infty)$, and $\mathcal{G} = \{(a,b); 0 < a < b < \infty\}$. Let us define

$$P((a,b)) = \int_a^b e^{-x} dx,$$

for any $0 < a < b < \infty$, and if $I_1, I_2, \ldots$ are disjoint intervals of $[0,2]$, then

$$P\left(\bigcup_{I=1}^{\infty} I_i\right) = \int_{I_1 \cup I_2 \ldots} e^{-x} dx.$$

Since the above P satisfies all the properties of a probability function on the elements of $\mathcal{G}$, therefore it is a probability function on the Borel σ-field.

Example: Suppose $\Omega = \{1,2,3,4,5\}$, $A = \{1,2\}$ and $B = \{3,4\}$. Then let us look at the minimum σ-field which contains the sets A and B. $\mathcal{F} = \{\Omega, \phi, \{1,2\}, \{3,4\}, \{3,4,5\}, \{1,2,5\}, \{1,2,3,4\}, \{5\}\}$. Suppose we want to define the probability function on $\mathcal{F}$, note that it is enough to do it on A and B. Say we define $P(A_1) = p_1$, $P(A_2) = p_2$, where $0 \le p_1, p_2, p_1 + p_2 \le 1$, then it is uniquely define on all the elements of $\mathcal{F}$. We have $P(\Omega) = 1$, $P(\phi) = 0$, $P(\{1,2\}) = p_1$, $P(, \{3,4\}) = p_2$, $P(\{3,4,5\}) = 1 - P(\{1,2\}) = 1 - p_1$, $P(\{1,2,5\}) = 1 - p_2$, $P(\{1,2,3,4\}) = p_1 + p_2$.