Probability space

The notion of probability space is at the heart of the mathematical treatment of probability. From the point of view of this course, you can think of it as a unifying idea: it unifies discrete and continuous probability.

A **probability space** is a triple $(\Omega, \mathcal{A}, P)$ where Ω is the sample space, $\mathcal{A}$ is a σ–field of subsets of Ω and P is a probability measure on Ω.

Understanding this requires two definitions.

Definition A nonempty collection of subsets $\mathcal{A}$ of a set Ω is a $\boldsymbol{\sigma}$**–field** if it has the following two properties.

(i) If $A \in \mathcal{A}$ then $\bar{A} = \Omega \setminus A \in \mathcal{A}$.

(ii) If $A_n \in \mathcal{A},\ n = 1, 2, \ldots,$ then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

Note that (i) and (ii) imply that $\bigcap_{n=1}^{\infty} A_n \in \mathcal{A}$ since $\bigcap_{n=1}^{\infty} A_n = \Omega \setminus \left(\bigcup_{n=1}^{\infty}(\Omega \setminus A_n)\right)$.

$\boldsymbol{\sigma}$**–algebra** is another name for σ–field.

Definition A **probability measure** P on a σ–field $\mathcal{A}$ is a real valued function having domain $\mathcal{A}$ satisfying the following properties:

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

(ii) $P(A) \geq 0$ for all $A \in \mathcal{A}$.

(iii) If $A_n,\ n = 1, 2, 3, \ldots,$ are pairwise disjoint sets in $\mathcal{A}$ (so $A_i \cap A_j = \emptyset$ if $i \neq j$) then

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

Example 1 If Ω is a discrete set sample set with a random variable X and a distribution function m for X then we can take $\mathcal{A}$ to be the collection of all subsets of Ω and the probability measure P to be given by

$$P(E) = \sum_{\omega \in E} m(\omega).$$

Example 2 If $\Omega \subset \mathbf{R}^n$ (e.g., Ω is an interval or a nice region in the plane) and X is a continuous real–valued random variable on Ω with a density function f then

$$P(E) = P(X \in E) = \int_E f(x)dx$$

defines a probability measure on any σ–field of nice subsets of Ω.

I don't want to get into technicalities about what subsets are nice, but I will note that if Ω is an interval in $\mathbf{R}$ we can take the σ–field to be the **Borel sets** $\mathcal{B}$, the smallest σ–field containing all the open subintervals of Ω. It is clear what $P(E)$ means if E is an interval. The hard work is to show that $P(E)$ can be defined for any Borel set E.

In this example it is not possible to take $\mathcal{A}$ to be the σ–field of all subsets of Ω.