DEPARTMENT OF MATHEMATICS
Indian Institute of Technology Guwahati
Probability and Random Processes
Notes-1

The classical definition of Probability was introduced formally by James Bernoulli (1654-1705) where he defines the probability of an event as the ratio of the number of outcomes favorable to an event to the total number of all possible equally likely outcomes. There were obvious difficulties with this definition.
Richard von Mises proposed a definition of probability which was based on empirical observations. We should have observations obtained either by repeating an experiment a large number of times, or from some mass-phenomena.
If A is an event and if the experiment is performed n times of which $N_n(A)$ is the number of times A occurred, then he claimed that the relative frequency of occurrence of A given by $\frac{N_n(A)}{n}$ tends to some constant value as n grows larger and larger, which is the probability of A.
That is, $P(A) = \lim_{n\to\infty} \frac{N_n(A)}{n}$.

The axiomatic definition was given by A.N. Kolmogorov.

Axiomatic definition of probability: Initial attempt
Given a nonempty set Ω and the power set of Ω (denoted by say P^{Ω}), a function $P : P^{\Omega} \to \mathbb{R}$ is called a probability measure if it satisfies the following axioms.

1. $P(A) \geq 0$ for all $A \subseteq \Omega$.
2. $P(\Omega) = 1$.
3. If A_i for $i = 1, 2.....$ is a sequence of mutually exclusive (disjoint) subsets of Ω then $P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$.

The following example illustrates that defining probability for all subsets of Ω may not always be possible or desirable.
Example 1: Let $\Omega = \{1, 2, ..., 60\}$.
Take $A = \{\omega \in \Omega : \omega \text{ is a multiple of 3}\}$ and $B = \{\omega \in \Omega : \omega \text{ is a multiple of 4}\}$.
Define $P(A) = \frac{N_{60}(A)}{60}$, where $N_{60}(A)$ is the number of times A occurs in Ω.
Hence $P(A)$ is the proportion of the number of elements which belong to A to the total number of elements in Ω.
Similarly for any $S \subseteq \Omega$, we can define $P(S) = \frac{N_{60}(S)}{60}$.
Now consider the changed problem where $\Omega = \mathbb{N}$, the set of natural numbers.
Let us see if we can use the above definition of P to get a probability measure (that is a P satisfying all the three axioms) for each and every subset of Ω.
For any $S \subseteq \Omega$, let $P(S) = \lim_{n\to\infty} \frac{N_n(S)}{n}$,
where $N_n(S)$ is the number of times S occurs in the first n natural numbers.
If A is defined as before then note that
$\frac{N_n(A)}{n} = \frac{m}{3m}$ if $n = 3m$.
$= \frac{m}{3m+1}$ if $n = 3m + 1$.
$= \frac{m}{3m+2}$ if $n = 3m + 2$.
Hence for all $n \in \mathbb{N}$, $\frac{1}{3+\frac{6}{n-2}} \leq \frac{N_n(A)}{n} \leq \frac{1}{3}$.

Hence $P(A) = \lim_{n\to\infty} \frac{N_n(A)}{n} = \frac{1}{3}$.
Similarly $P(B) = \lim_{n\to\infty} \frac{N_n(B)}{n} = \frac{1}{4}$.
Let $C = \{2\}$.
Then $\frac{N_n(C)}{n} = \frac{0}{n}$ for $n = 1$
$= \frac{1}{n}$ for $n \geq 2$.
Hence $P(C) = \lim_{n\to\infty} \frac{N_n(C)}{n} = 0$.
Hence $P(C) = 0$ for any singleton set C.
But $\Omega = \mathbb{N} = \cup_{i\in\mathbb{N}}\{i\}$, hence if P satisfies the 3rd axiom then $P(\Omega) = \sum_i P(\{i\}) = 0 \neq 1$, which contradicts the 2nd axiom.
Hence this P defined on the power set of Ω does not satisfy all the three axioms but this P gives meaningful probabilities for sets like A and B. That is this P works well for certain subsets of Ω but it cannot be defined for all subsets of Ω.
Hence as this example suggests, depending on our objective we may have to choose from the set of all subsets of Ω, certain subsets (not all) of Ω on which to define a probability measure P. Let us denote this collection of subsets of Ω by $\mathcal{A}$.

Definition: A nonempty collection of subsets $\mathcal{A}$ of a (nonempty set) Ω is called a **sigma field** of subsets of Ω if it satisfies the following properties:

1. If $A \in \mathcal{A}$ then $A^c \in \mathcal{A}$.

2. If $A_i \in \mathcal{A}$ for all $i = 1, 2....$ then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$.

Definition: A nonempty collection of subsets $\mathcal{A}$ of a (nonempty set) Ω is called a **field** of subsets of Ω if it satisfies the following properties:

1. If $A \in \mathcal{A}$ then $A^c \in \mathcal{A}$.

2. If $A_i \in \mathcal{A}$ for $i = 1, 2....n$, $n \in \mathbb{N}$ then $\bigcup_{i=1}^{n} A_i \in \mathcal{A}$.

Exercise: Note that condition (2) in the definition of a **field** is equivalent to the condition that for any two sets $A_1, A_2 \in \mathcal{A}$, $A_1 \cup A_2 \in \mathcal{A}$. Use induction to see this.

Example: Consider the example of tossing a coin twice successively.

1. $\mathcal{A}_1 = \{\phi, \Omega, \{HT\}, \{TH, TT, HH\}\}$

2. $\mathcal{A}_2 = \{\phi, \Omega, \{HT\}, \{TH\}, \{TH, TT, HH\}, \{HT, TT, TH\}, \{HT, TH\}, \{TT, HH\}\}$

are all fields and sigma fields of subsets of Ω.
Remark: Note that a sigma field is always a field but the converse is not true.
Example: Let $\Omega = \mathbb{N}$ and let $\mathcal{A} = \{A \subseteq \Omega : \text{ either } A \text{ or } A^c \text{ is finite } \}$.
For all $n \in \mathbb{N}$, let $A_n = \{2n\}$, then $A_n \in \mathcal{A}$ for all n.
If $\mathcal{A}$ is a sigma field, $\bigcup_n A_n = \{\omega \in \Omega : \omega = 2n \text{ for some } n \in \mathbb{N}\}$ should also belong to $\mathcal{A}$.
But since neither $\bigcup_n A_n$ nor $(\bigcup_n A_n)^c = \{\omega \in \Omega : \omega = 2n - 1, \text{ for some } n \in \mathbb{N}\}$ is finite $\bigcup_n A_n$ does not belong to $\mathcal{A}$. So $\mathcal{A}$ is not a sigma field.
However $\mathcal{A}$ is a field.
$\mathcal{A} \neq \phi$ is obvious.
Let us assume $A \in \mathcal{A}$ then either $A = (A^c)^c$ or A^c is finite so $A^c \in \mathcal{A}$.
Let $A, B \in \mathcal{A}$.
If both A and B are finite then $A \cup B$ is also finite hence $A \cup B \in \mathcal{A}$.
If both A and B are not finite or at least one of them is infinite (countably infinite) then since $A, B \in \mathcal{A}$ at least one of A^c or B^c is finite, hence $A^c \cap B^c = (A \cup B)^c$ is finite, hence $(A \cup B) \in \mathcal{A}$.
$\Rightarrow \mathcal{A}$ is a field.

Hence the definition of probability.

Axiomatic definition of probability: Given a nonempty set Ω and a sigma field of subsets of Ω denoted by $\mathcal{A}$, a function $P : \mathcal{A} \to \mathbb{R}$ is called a probability measure if it satisfies the following axioms.

1. $P(A) \geq 0$ for all $A \in \mathcal{A}$
2. $P(\Omega) = 1$
3. If A_i for $i = 1, 2, ...$ is a sequence of mutually exclusive (disjoint) sets in $\mathcal{A}$ then $P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$.

Exercise: For the function P defined in **Example 1** can you suggest a sigma field of subsets of $\Omega = \mathbb{N}$ on which to define P such that the P will be a probability measure and $P(A)$, $P(B)$ will be defined.