Module 3

A MEASURE OF UNCERTAINTY

Example 1:

- An agricultural scientist wishes to study the effect of a fertilizer on the yield of certain crop;
- To get the information about the effect of fertilizer, one needs to perform experiment (say, fertilizer is applied on 100 homogenous plots of an agricultural land and output of crop is recorded in each plot);
- Here each experiment terminates in an outcome (say crop output per square yard of land in a given plot), which can not be predicted with certainty in advance.

Some Definitions:

Definition 1: A *random experiment* is an experiment

(a) in which all possible outcomes of the experiment are known in advance;

(b) outcome of a particular trial of the experiment cannot be predicted in advance;

(c) the experiment can be repeated under identical conditions.

Definition 2: The collection of all possible outcomes of a random experiment is called a *sample space* (to be denoted by Ω).

Definition 3: An *event* is a subset of sample space.

- If the outcome of an experiment is a member (element) of $A \subseteq \Omega$, we say that event A has occurred.

Example 2:

- Random experiment ($\mathcal{E}$): Casting one red die and one white die;
-

$$
\begin{aligned}
\text{Sample space } \Omega &= \{(i,j) : i \text{ is the number of spots on the red die and} \\
&\qquad j \text{ is the number of spots on the white die}\} \\
&= \{(i,j) : i,j \in \{1,2,\ldots,6\}\} \\
&= \{(1,1),\ldots,(1,6),(2,1),\ldots,(2,6),\ldots,(6,1),\ldots(6,6)\}\,;
\end{aligned}
$$

- $A = \{(i,j) : i \in \{1,3,5\}\,, j \in \{2,4,6\}\}$ is the event of getting an odd number of spots on red die and even number of spots on white die.

Some More Definitions

Definition 4: A set of sets (usually denoted by script letters $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}, \ldots$) will be called a *class of sets*. For example $\mathcal{A} = \{\{1\}, \{1,3\}, \{1,2,3,4\}\}$ is a class of sets.

Definition 5: The class of all subsets (including empty set ϕ) of a given set A is called the *power set* of A (denoted by $\mathcal{P}(A)$).

Remark 1: In mathematical treatment of probability theory, to avoid certain mathematical inconsistencies, not all subsets of sample space are considered to be events (especially when the sample space Ω contains an interval (or elements of Ω can be put in one-one correspondence with an interval)). In such situations some weird sets (which will seldom occur in real life situations) are excluded from the event space. In this course we will consider any subset of sample space to be an event and take our event space to be power set of Ω.

Relative Frequency Approach to Assigning Probability:

- $\mathcal{E}$: a random experiment;
- Ω: Sample space associated with $\mathcal{E}$;
- $\mathcal{P}(\Omega)$ (power set of Ω) : event space associated with $\mathcal{E}$;
- Let $E \in \mathcal{P}(\Omega)$ be a given event;
- To assign probability to event E, we repeat $\mathcal{E}$ N (a large number) of times;
- $f_N(E)$: number of times event E occurred in first N performances of $\mathcal{E}$ (frequency of event E in first N trials);
- $r_N(E) = \frac{f_N(E)}{N}$: relative frequency of E in first N performances of $\mathcal{E}$;

- Assign probability to event E as

$$\begin{aligned} P(E) &= \lim_{N\to\infty} r_N(E) \quad \text{(part of probability modelling)} \\ &\approx r_N(E) \quad \text{(for large N).} \end{aligned}$$

- Existence of above limit will be justified (in certain sense), later in the course.

- Clearly,

  (i) $P(E) \geq 0, \qquad \forall E \in \mathcal{P}(\Omega);$

  (ii) $P(\Omega) = \lim_{N\to\infty} \frac{f_N(\Omega)}{N} = 1;$

  (iii) if E_1 and E_2 are disjoint events, then

$$\begin{aligned} P(E_1 \cup E_2) &= \lim_{N\to\infty} \frac{f_N(E_1 \cup E_2)}{N} \\ &= \lim_{N\to\infty} \frac{f_N(E_1) + f_N(E_2)}{N} \\ &= \lim_{N\to\infty} \left[\frac{f_N(E_1)}{N} + \frac{f_N(E_2)}{N}\right] \\ &= P(E_1) + P(E_2). \end{aligned}$$

Some Definitions and Notations

- **Set Function:** An extended real valued function whose domain is a class of sets;

- **Countable Set:** A set that is either finite (has finite number of elements) or whose elements can be put in $1-1$ correspondence with the set of natural numbers $\mathbb{N}$. Clearly if A is a countable set then we can write $A = \{a_i : i \in S\}$, for some $S \subseteq \mathbb{N}$.

- $\mathbb{R}$: real line (i.e., $\mathbb{R} = (-\infty, \infty)$).

Motivated by properties (i)-(iii) of relative frequency assignment of probability, we provide the following definition.

Definition 6: Given a random experiment $\mathcal{E}$ and the associated sample space Ω, a probability function P is a set function defined on event space $\mathcal{P}(\Omega)$ satisfying the following three axioms:

Axiom 1: $P(E) \geq 0, \forall E \in \mathcal{P}(\Omega)$;

Axiom 2: if $\{E_i : i \in S\}$ is a countable collection of disjoint events, then

$$P\left(\bigcup_{i \in S} E_i\right) = \sum_{i \in S} P(E_i);$$

Axiom 3: $P(\Omega) = 1$.

Definition 7: The triplet $(\Omega, \mathcal{P}(\Omega), P)$ is called a probability space.

Example 3:

- Random experiment ($\mathcal{E}$): Casting a red and a white die;
- Sample space: $\Omega = \{(i, j) : i, j \in \{1, 2, ..., 6\}\}$;
- $\mathcal{P}(\Omega) =$ event space;
- For any event $E \in \mathcal{P}(\Omega)$, define

$$P(E) = \frac{|E|}{|\Omega|} = \frac{|E|}{36}, \quad \text{(part of probability modelling)}$$

where for an event A, $|A|$ denotes the number of elements in A (or number of cases favorable to A). Verify that the set function P defined on event space $\mathcal{P}(\Omega)$ is a probability function.

Abstract of Module 4

Using three axioms of probability function we will derive various properties of probability function.

Thank you for your patience