MA20104 Probability & Statistics

Summary of lectures in the first week.

(Instructor: Swanand Khare)

Random Experiment: (R)

An experiment whose outcomes are not deterministic (fixed) is called a random experiment.

There are a lot of real life examples of the random experiments which lead to the development of probability theory.

(various examples of random experiment are discussed in the class).

Sample space : (Ω) Set of all possible outcomes of a random experiment.

Event : Any subset of the sample space.

The objective is to be able to assign probabilities to outcomes/events of a random experiment.

Types of sample spaces:

Ω : 1) discrete/denumerable/countable

2) Continuous

3) mixed

Probability assignment:

Case 1 : The case of equally likely outcomes (relative frequency apporach).

- ~~Conduct~~ Define random experiment. (R)
- List all possible outcomes (Ω) and make sure that all outcomes are equally likely.
- Event E is collection of favourable outcomes.

Then $P(E) = \dfrac{\text{no. of favourable outcomes}}{\text{total no. of outcomes}}$

$$\boxed{P(E) = \frac{\# E}{\# \Omega}}$$

<u>Examples</u> : Tossing a fair coin, Rolling a fair die etc.

<u>Case 2</u> : The case when the outcomes in a sample space may <u>NOT</u> be equally likely.

Let the sample space be finite.

In this case, we can repeat the random experiment (R) 'n' number of times. Then count how many number of times a particular outcome appears (n_f). Then the ratio ~~of th~~

$$\frac{n_f}{n}$$ is a very good "<u>estimate</u>" of

the probability of occurrence of that particular outcome.

<u>Simulation exercise</u> : Simulate the scenario of throwing a fair die in a computer. Then for various values of $n = 10, 50, 100, 1000, 10000, \ldots$ compute estimates of $P(\{1\})$, $P(\{3\})$. How does the "<u>quality</u>" of these estimates change with varying n?? We will revisit this example again in greater details in the later part of this course.

In this case, let $\Omega = \{s_1, s_2, \dots, s_k\}$. Then using the procedure described above, we can assign probabilities to the outcomes $s_1, s_2, \dots, s_k$ looking at the estimates. However, intuitively we would follow following rules in this probability assignment.

$$P(\Omega) = P(s_1) + P(s_2) + \dots + P(s_k) = 1.$$

Further, we would expect that $P(\{s_1, s_2\}) = P(\{s_1\}) + P(\{s_2\})$ and for any event $E \subseteq \Omega$, $P(E^c) = 1 - P(E)$.

In fact, these expectations can be formalised in a mathematical way to give a rigorous definition of a probability space and probability measure.

Probability Space

Consider a random experiment ~~R~~ R with sample space Ω. We are interested in building a collection of events, denoted as $\mathcal{F}$, so that we can assign probabilities to every member of $\mathcal{F}$. Then obviously we would expect that $\mathcal{F}$ satisfies some properties observed earlier. The most trivial requirement of $\mathcal{F}$ is that it contains Ω. Further for any event $A \subseteq \Omega$, we expect $A^c \in \mathcal{F}$. Also, for a sequence of events $A_1, A_2, \ldots, A_n, \ldots$ we expect $\bigcup_{i=1}^{\infty} A_i$ and $\bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$. We formally define such $\mathcal{F}$ now.

<u>Definition</u>: Let R be a random experiment with sample space Ω. A non-empty collection $\mathcal{F}$ of subsets of Ω is called a σ-field of Ω if the following properties hold:

i) $\Omega \in \mathcal{F}$

ii) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$

iii) $A_n \in \mathcal{F}$ for $n = 1, 2, \ldots$, then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$

and $\bigcap_{n=1}^{\infty} A_n \in \mathcal{F}$.

<u>Example</u>: Let R be a random experiment with a finite sample space Ω. Then $\mathcal{F} = 2^{\Omega}$, power set of Ω. (Power set is set of all subsets).

In a fair coin tossing expt., $\mathcal{F} = \{\phi, \{H\}, \{T\}, \{H,T\}\}$.

With the definition of σ-field $\mathcal{F}$ in place, now we can give definition of probability measure (we have termed this probability assignment / probability model for the random experiment) which is also called an axiomatic definition.

<u>Definition</u>: Let R be the random experiment with sample space Ω and associated σ-algebra $\mathcal{F}$. Then a set function $P : \mathcal{F} \to \mathbb{R}$ is called probability measure if

i) for $A \in \mathcal{F}$, $0 \le P(A) \le 1$.

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

iii) $A_1, A_2, \cdots, A_n, \ldots \in \mathcal{F}$ is a collection of mutually disjoint events, then

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

<u>Example</u>: R: Tossing a ~~fa~~ coin.

$$\Omega = \{H, T\}. \qquad \mathcal{F} = \{ \phi, \{H\}, \{T\}, \{H, T\} \}$$

Note that it is enough to specify $P(\{H\})$ alone here. Then all other probability assignments follow from the postulates (i), (ii), (iii) in the definition.

Example: R: Throwing a die

$$\Omega = \{1, 2, 3, 4, 5, 6\} \quad ; \mathcal{F} = 2^{\Omega}$$

Note that here also it is enough to assign probabilities $P\{(1)\}$, $P\{(2)\}$, $P(\{3\})$, $P(\{4\})$, $P(\{5\})$. Then using the axioms in the definition, one can assign the other probabilities easily.

Example: Uniform probability spaces

We have done this in class, though I did not mention the name. One of the most important problems in probability theory is to pick a point "at random" from a set Ω. Then intuition suggests that for two subsets A & B of Ω with "equal size", the chance of picking a point from A should be same as that from B. If Ω is finite, then one can measure size of a set by its cardinality. Thus sets A & B are of same size if they have same cardinality. The corresponding probability space in this case is then: Ω – sample space, $\mathcal{F}$: power set of Ω and for any $A \in \mathcal{F}$, $P(A) = \dfrac{\# A}{\# \Omega}$.

Note that, we have already seen two examples of this scenario: (1) Tossing a fair coin and (2) Rolling a fair die. The condition of fairness plays a key role in making the corresponding probability spaces uniform.

Another interesting case arises here when $\Omega = [0,1]$. Here, measure of "size" of any subset $A = (a,b)$ of Ω can be considered to be its length $b-a$. Two sets are of equal size if they have same length. Now the challenge here is to construct σ-algebra $\mathcal{F}$. It is shown (see the book Probability & Measure by P. Billingsley) that $\mathcal{F}$ is a Borel sigma field. (In loose terms, all sets in $\mathcal{F}$ can be constructed by countable union and intersection of open intervals). The probability measure is $P(A) = b-a$ when $A = [a,b]$ with $0 \le a \le b \le 1$. It is to be noted that all the usual sets one would encounter in real life applications are generally present in the collection $\mathcal{F}$ and we can assign the probability to these sets (events).

Some consequences of definition of probability space.

Thm: If $A \in \mathcal{F}$, then $P(A^c) = 1 - P(A)$

Pf: Note $\Omega = A \cup A^c$ is a representation of Ω as union of disjoint events.

$$1 = P(\Omega) = P(A \cup A^c) = P(A) + P(A^c)$$

$$\Rightarrow \quad P(A^c) = 1 - P(A).$$

Thm: $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$

Pf: $A_1 \cup A_2 = A_1 \cup (A_2 \cap A_1^c)$ ... disjoint union.

$$\Rightarrow P(A_1 \cup A_2) = P(A_1) + P(A_2 \cap A_1^c) \quad \text{—(1)}$$

Further, $A_2 = (A_1 \cap A_2) \cup (A_2 \cap A_1^c)$ ... disjoint union.

$$\Rightarrow P(A_2) = P(A_1 \cap A_2) + P(A_2 \cap A_1^c) \quad \text{—(2)}$$

From (1) & (2)

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

Remark: To better understand, draw Venn diagrams.

Remark: Generalize above theorem to the case of union of events $A_1, A_2, \ldots, A_n$.

Remark: From the above theorem, it is clear that

$$P(A_1 \cup A_2) \leq P(A_1) + P(A_2)$$

Law of total probability.

A collection $\{B_1, B_2, \cdots, B_k\}$ is said to be a partition of the sample space Ω if (i) $B_i \subseteq \Omega$, $i = 1, 2, \cdots, k$ (ii) B_i's are mutually disjoint and (iii) $\Omega = \bigcup_{i=1}^{k} B_i$.

Then for any event $A \subseteq \Omega$, we can write

$A = \bigcup_{i=1}^{k} (A \cap B_i)$ which is a disjoint union of events. Thus $\boxed{P(A) = \sum_{i=1}^{k} P(A \cap B_i)}$.

Special Type of chains of events.

Thm: Let $A_1, A_2, \cdots, A_n, \cdots$ be an increasing sequence of events of Ω, that is,

$$A_1 \subseteq A_2 \subseteq \cdots \subseteq A_n \subseteq \cdots \subseteq A \subseteq \Omega.$$

Then $\lim_{n \to \infty} P(A_n) = P(A)$.

Proof: Define $B_1 = A_1$, $B_2 = A_2 \cap A_1^c$, ...

In general, $B_k = A_k \cap A_{k-1}^c$ for $k = 2, 3, \cdots$.

Then $A = \bigcup_{i=1}^{\infty} B_i$ where B_i are mutually disjoint. (Check this with Venn diagram).

Then. $P(A) = \sum_{i=1}^{\infty} P(B_i) = \lim_{n\to\infty} \sum_{i=1}^{n} P(B_i)$

... from definition of limit of series

$= \lim_{n\to\infty} P(B_1 \cup B_2 \cup \dots \cup B_n)$

... B_i's are mutually disjoint

$$\boxed{P(A) = \lim_{n\to\infty} P(A_n)}$$

... See Venn diagram.

<u>Theorem</u>: Let $A_1, A_2, \dots, A_n, \dots$ be a decreasing sequence of events of Ω, that is,

$$A_1 \supseteq A_2 \supseteq \dots \supseteq A_n \supseteq \dots \supseteq A$$

Then $\lim_{n\to\infty} P(A_n) = P(A)$.

<u>Proof</u>: $A_1 \supseteq A_2 \supseteq \dots \supseteq A_n \supseteq \dots \supseteq A$

$\Rightarrow A_1^c \subseteq A_2^c \subseteq \dots \subseteq A_n^c \subseteq \dots \subseteq A^c$

$\Rightarrow \lim_{n\to\infty} P(A_n^c) = P(A^c)$ — by previous theorem

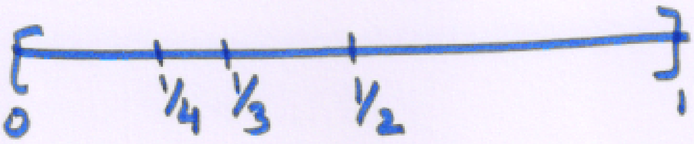
$\Rightarrow \lim_{n\to\infty} [1 - P(A_n)] = 1 - P(A)$

$\Rightarrow \boxed{\lim_{n\to\infty} P(A_n) = P(A)}$

<u>Example</u>: Consider the uniform probability space with $\Omega = [0,1]$. Remember, here $\mathcal{F}$ is a Borel-sigma algebra which contains subsets of $[0,1]$ which can be constructed from countable unions and intersections of intervals. We illustrate the use of previous theorem to compute certain probabilities.

Let $A_n = [0, \frac{1}{n}]$ for $n \in \mathbb{N}$.

Then clearly, $A_1 \supseteq A_2 \supseteq \cdots \supseteq A_n \supseteq \cdots \supseteq A = \{0\}$



$$P(A_n) = \frac{1}{n} - 0 = \frac{1}{n}$$

According to the previous theorem,

$$P(\{0\}) = P(A) = \lim_{n\to\infty} P(A_n) = \lim_{n\to\infty} \frac{1}{n}$$

Thus $\boxed{P(\{0\}) = 0}$

We can similarly deduce that for any $\{x\}$ with $x \in [0,1]$, $P(\{x\}) = 0$. Then using, the third postulate in the probability space definition, we conclude

$P(\mathbb{Q} \cap [0,1]) = 0$ where $\mathbb{Q} \cap [0,1]$ is the set of all rational numbers in $[0,1]$.