HSO - 201 : Probability and Statistics

Lecture 1: Random Experiments and Probability

1. INTRO: There are lot of things around us about which we cannot make any precise prediction. For example the weather, the outcome of a coin toss, the roll of a die, etc. Such processes are often referred to as "random experiments". So outcomes of random experiments are uncertain. However in most of these situations we know what are the possible outcomes. Each such outcome is called an elementary event. The set of all elementary events is called the sample space often denoted as Ω.

Example 1: If we toss a coin it is not sure in which way it will turn up when it hits the ground. One may say apart from head or tail one may argue that on a field for example a coin may get stuck and remain vertical. Thus in order to build a theory of probability we need make more idealistic assumption. So we talk of a fair coin, which can either be head or tail, when it falls to the ground. Thus

$$\Omega = \{H, T\}, \quad \because H \equiv \text{Head} \quad T \equiv \text{Tail}.$$

Example 2: If we roll a fair die, then the sample space is given as,

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

Example 3: Now let us toss two coins simultaneously. What is Ω? Note that for each outcome of the first coin, there are two outcomes of the second coin, i.e.

Coin 1 ← H → H, T (Coin 2)

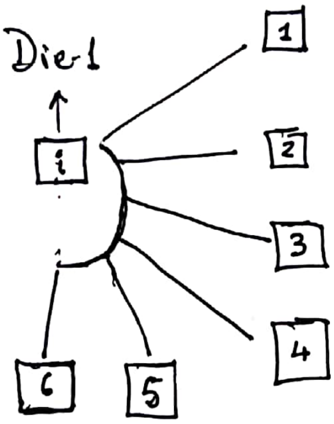
Coin 1 ← T → H, T (Coin 2)

Thus

$$\Omega = \{HH, HT, TH, TT\}.$$

Example 4: If we role a fair ~~dy.~~ die, then ~~role~~ roll, another fair die ~~So~~ then there are 36 possible outcomes, since for each of the six outcomes of the first die, there corresponds six outcomes of the second die. This is given as

$i = 1, 2, 3, 4, 5, 6$



$$\Omega = \left\{ \begin{matrix} (1,1), & (1,2), & (1,3), & (1,4), & (1,5), & (1,6) \\ (2,1), & (2,2), & (2,3), & (2,4), & (2,5), & (2,6) \\ (3,1), & (3,2), & (3,3), & (3,4), & (3,5), & (3,6) \\ (4,1), & (4,2), & (4,3), & (4,4), & (4,5), & (4,6) \\ (5,1), & (5,2), & (5,3), & (5,4), & (5,5), & (5,6) \\ (6,1), & (6,2), & (6,3), & (6,4), & (6,5), & (6,6) \end{matrix} \right\}$$ $\square$

All our examples above may make us feel that for any random experiments the number of all possible outcomes is finite. But these need not be the case. For example consider the number of car accidents that may happen in Kanpur in the next 6 ~~next~~ months. In fact theoretically any non-negative integer would do, i.e.

$$\Omega = \{0, 1, 2, 3, 4 \ldots\ldots\}$$

As we will see later that there can be a sample space with uncountably infinite ~~number o~~ sample points. We shall however in this chapter consider only sample spaces with finite number of sample points.

2. Events and Probability.

To be precise, for a finite sample space Ω, we can define any subset of Ω to be an event. Thus the power set of Ω, ie 2^{Ω} is called the set of all events associated with random experiment with sample space Ω.

Remember the outcome of a ~~rat~~ trial of a random experiment is some $\omega \in \Omega$. If $\omega \in A \subset \Omega$, then we say that the event A has occured

For example let us consider the rolling of a fair die and let us say A be the event that an odd number appears. Then A is said to happen if the face of the die shows 1, 3 or 5. Thus $A = \{1, 3, 5\}$, which is a subset of $\Omega = \{1, 2, 3, 4, 5, 6\}$.

We say that two events A and B are mutually exclusive if $A \cap B = \phi$. We say events $A_1\ A_2 \ldots A_k$ are mutually exclusive if $A_i \cap A_j = \phi$, $i, j = 1 \ldots k$ & $i \neq j$. ϕ or the empty-set denotes the null-event, som very often referred to as the impossible event.

Let us consider a random experiment which has N outcomes, ~~Com~~ i.e. $\Omega = \{\omega_1, \omega_2, \ldots \omega_N\}$, i.e. Card$(\Omega) = N$.

Note $\{\omega_1\}, \{\omega_2\}, \ldots \{\omega_N\}$ forms a set of mutually exclusive events. For example if we toss a "fair" coin, then either "head" or "tail" appears. Both cannot appear at the same time. However we also "feel" that "head" or "tail" must have an equal chance of ~~appearin~~ materializing. This is what one says as "Head & Tail are equally likely events". In fact each $\omega_i \in \Omega$, is called an elementary event. So we are now ready to give the "Classical" ~~De~~ definition of probability.

Classical ~~events~~ Definition of Probability

Let a random experiment result in N, mutually exclusive, equally likely and exhaustive, elementary events. Let A be an event such that card $(A) = n_A$. Then probability of the occurence of the event A, is the number $P(A)$ given as

$$\boxed{P(A) = \frac{n_A}{N}}$$

* Exhaustive means that apart from these N outcomes the

Certain facts are immediately clear.

i) $0 \leq P(A) \leq 1$

ii) $P(\Omega) = \frac{N}{N} = 1,$

iii) $P(\phi) = \frac{0}{N} = 0$

iv) If card $(A) = n_A$ and Card $(B) = n_B$, then

$$P(A \cup B) = \frac{\text{Card}(A \cup B)}{N} = \frac{n_A + n_B - n_{A \cap B}}{N}$$

$$\therefore P(A \cup B) = \frac{n_A}{N} + \frac{n_B}{N} - \frac{n_{A \cap B}}{N}$$

$$= P(A) + P(B) - P(A \cap B)$$

v) If $A \cap B = \phi \Rightarrow P(A \cup B) = P(A) + P(B)$.

vi) A^c ~~is called~~ denotes the complement of the event A; (ie. NOT A)

In fact Card $(A^c) = N - n_A$. Thus

$$P(A^c) = \frac{N - n_A}{N} = 1 - \frac{n_A}{N}$$

$$\therefore \boxed{P(A^c) = 1 - \frac{n_A}{N}}$$

$$\therefore \boxed{P(A^c) = 1 - P(A)}$$

Another way to look at the last inequality is the following

$$A \cup A^c = \Omega,$$

Since $A \cap A^c = \phi$, $\Rightarrow P(\Omega) = P(A) + P(A^c)$

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

<u>Drawback:</u> One might note that the word "equally likely" must refer to the fact that each elementary event has an equal "probability" of occurence, or that they are '<u>equiprobable</u>'. Thus the definition might appear circular. In the next chapter we shall see how ~~Kolm~~ Andrei Kolmogorov, the great Russian mathematician provided a solid mathematical foundation to probability.

The term equiprobable means: $P(\{\omega_i\}) = \frac{1}{N}$, for $i = 1 \dots N$.

Thus when you toss a coin: $P(\text{Head}) = 0.5$, $P(\text{Tail}) = 0.5$

Thus when you roll a die: $P(i) = \frac{1}{6}$, $i = 1, 2, \dots, 6$.

Infact if you repeatedly toss a coin you will see after a large number of tosses, the number of heads and tails are almost equal. Thus using these kinds of experimentation one may have a feel that the classical definition is not bad at all.

We will soon see its power, but let us just begin with a story. <u>Chevalier de Mere</u>, ~~pose~~ who was a French gambler posed this problem to his mathematician friend <u>Blaise Pascal</u>

<u>Which is more likelier:</u>

Rolling at least one six in four throws of a die

OR

Getting a double six in 24 throws of a pair of die

The gambler reasoned as follows :

He felt that the average number of succesful roll is same in both cases:

Case-I: Probability of getting six in a single throw is $\frac{1}{6}$. So the expected value in 4 throws is $4 \times \frac{1}{6} = \frac{2}{3}$

Case-II: Probability of getting a double six when a double die is rolled is $\frac{1}{36}$, so the average in 24 throws is $24 \times \frac{1}{36} = \frac{2}{3}$.

But 'Chevalier de Mere' observed that he lost frequently when he used the second gamble. This problem excited Blaise Pascal, and he wrote about this to his friend Pierre de Fermat about this issue, and between them they formed the mathematical edifice of probability theory. We shall solve 'Chevalier de Mere''s problem in Lecture 3.

Let us now use the classical definition to solve an interesting problem, Let there be N people in ~~this~~ a party. What is probability that at least two of them have the same birthday. Let us ignore leap years. This problem is solved by first computing the ~~complement~~ probability of the complementary event, i.e. no two people have their birthdays on the same day of the year. Let A^c denote that event

Now $\text{card}(\Omega) = 365^N$ in this case.

$$P(A^c) = \frac{365 \times 364 \times 363 \times \cdots \times (365-(N-1))}{365^N}$$

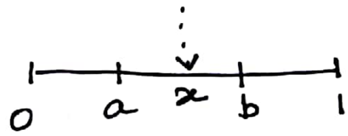
The first person can have his birthday on any of the 365 days and second one on the remaining 364 days, the third on the remaining 363 days and so on. Thus our event A has the probability

$$P(A) = 1 - \left(\frac{365 \times 364 \times 363 \times \cdots \times (365-N+1)}{365^N} \right)$$

If $N = 25$, then one can show that $P(A^c) \approx 0.40$, showing that $P(A) \approx 0.60$, thus the probability of the birthdays of two people matching ~~be~~ increases beyond 0.5 if $N \geq 25$. So if N is large such an event is equally like .

3. Infinite sample Spaces & Bertrand's Paradox

Consider the following problem. Let us be given the line $[0,1]$

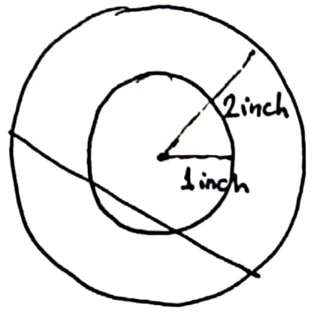


We ~~be~~ throw a stone or a pin on the line. We would like to Know if our pin falls in the interval (a, b), $a<b$, & $0<a<b<1$, ~~So our sample space~~ and what is the probability that ~~he~~ it will ~~do so be so~~ fall there.
It is intutively clear that $\Omega = [0,1]$. One way of calculating the probability of the event $A = \{\omega : a < \omega < b\}$, is to consider ~~or~~ the lengths as a kind of "cardinality" of these intervals.
Thus we may write

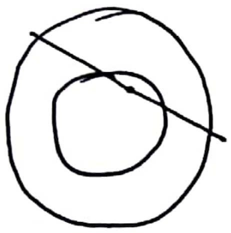
$$P(A) = \frac{\text{length}\,(a,b)}{\text{length}\,[0,1]} = b-a .$$

However observe that if this is the way we compute the probability in the above case then if $\tilde{A} = \{\omega : \omega = a\}$, then $P(\tilde{A}) = 0$. Thus it is impossible to through a pin at an exact point on $[0,1]$.
However such approaches have its own perils while we are handling the ~~inf~~ case of infinite sample spaces. This is aptly demonstrated by the Bertrand's paradox, where in the same event is shown to have to different probabilities.

Bertrand's Paradox:



Consider a ~~circler~~ circle of radius 2 inch and a concentric circle of radius 1 inch. A chord is drawn at random in the circle of radius 2inch. What is the probability of the chord cutting through ~~it~~ the circle of radius one inch?

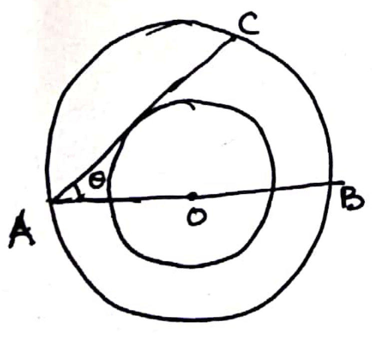


Solution: Approach 1: The chord of course is unique if it is not a diameter, in the sense that it can be identified through its mid-point. The chord to cut through the smaller concentric circle the mid point of the chord must lie in the circular disc formed by the smaller circle. Let us denote by A as the event that the chord cuts through the smaller circle. Here the sample space is the circular disc of radius 2inch. We shall get the probability of A as a ratio of areas

i.e

$$P(A) = \frac{\text{Area of Smaller circle}}{\text{Area of bigger circle}} = \frac{\pi}{\pi 4.} = \frac{1}{4}$$

Solution: Approach 2



Fix a diameter ~~and draw~~ AB of the bigger circle and observe that any chord AC starting at A will vary from 0 to π ~~whic~~ while it varies from 0 to $\pm \frac{\pi}{6}$ to cut through the Smaller circle. So, we have

$$P(A) = \frac{\frac{2\pi}{6}}{\pi} = \frac{1}{3}.$$

Note that for any chord AC we actually draw a horizontal diameter AB and argue as above.

You can see the contradiction!!