PROBABILITY THEORY I

As a subject, Statistics builds upon the foundation of probability theory. Historically probability theory was a part of Mathematics and it was developed as a mathematical formulation of gambling.

Here we will concentrate on the basic ideas relating to probability theory that are fundamental to the study of Statistics.

Set theory $\longrightarrow$ Probability theory $\longrightarrow$ Statistics

Briefly one can say that probability is the degree of belief that an event will occur OR probability is a measurement (or scaling on a closed continuous space) of chance to occur an event.

Set theory:

- Sets and elements
- Identity and cardinality

  Identity: Two sets A and B are identical if and only if they have exactly the same numbers, i.e., $A = B$ iff for every x, $x \in A \Leftrightarrow x \in B$.

  Cardinality: The number of elements in a set A is called the cardinality of A, and it's denoted as $|A|$.
  Remark: Infinite sets also have cardinalities but they are not natural numbers.

- Subset, Power set
- Different operations on sets: union, intersection, complement, difference
- Different Laws: Commutative, Associative, Distributive, De-Morgan's, Idempotent

Probability theory always start with some random experiment. Random experiment is such an experiment (performed in real life) for which outcome may not be certain. When such an experiment is performed, then a set of all possible outcomes is associated with that random experiment. This set is known as sample space of the random experiment and it is usually denoted ~~as~~ by Ω. Each outcome in a sample space Ω is called sample point.

Random Experiment:

A random (or statistical) experiment is an experiment in which:

(a) All outcomes of the experiment are known in advance

(b) Any performance of the experiment results in an outcome that is not known in advance.

(c) The experiment can be repeated under identical condition.

Examples: (1) Coin tossing twice, (2) Life-time of an electric bulb.

Sample Space:

The sample space, denoted as Ω, of a random experiment is the set of all possible outcomes.

Examples:

(1) Related to Random experiment (1) : $\Omega = \{HH, TH, HT, TT\}$

(2) " " " (2) : $\Omega = \{t : t \geq 0\} = [0, \infty)$

Now, in probability theory, we are specifically concern about the degree of belief (or probability) of a statement which is based on a random experiment. This statement of interest is referred as event. Hence, we always find the probability of an event. Two examples are..

(1) One Head & one Tail ~~will~~ appear in the random expt. (1).

(2) Life-time of the electric bulb is not more than 72 hrs.

Scanned by CamScanner

Event:

An event is any collection of possible outcomes of a random experiment. So it is any subset of sample space.

Remarks: (1) Event is a collection of some sample points belongs to Ω.

(2) Sample points are also known as elementary events as that cannot be decomposed further.

Hence, in probability theory an event E means it is a set contained in ($\subseteq$) Ω, the sample space (universal set of E). E consists some elementary events or sample points i.e.

$$E = \{x : x \text{ is elementary event } \in E\}.$$

So all types of set theoretic operations can be performed between events.

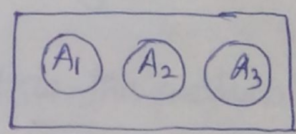
Different types of Events:

- Certain event : $E = \Omega$
- Impossible event : $E = \phi$, the empty set (as $\phi \subset \Omega$).
- Mutually Exclusive Events:

The events in the sequence $A_1, A_2, \ldots$ are said to be mutually exclusive, if

$$A_i \cap A_j = \phi, \text{ for all } i \neq j,$$

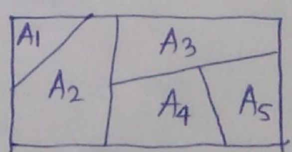
where ϕ represents the empty set.



- Exhaustive Events:

The events $A_1, A_2, \ldots, A_n$, for $n \geq 2$, are said to be exhaustive if

$$A_1 \cup A_2 \cup \ldots \cup A_n = \Omega$$



Here $n=5$.

Remarks: (1) $A_1 \cup A_2 \cup \ldots \cup A_n$ means at least one of the n events occurs

(2) $A_1 \cap A_2 \cap \ldots \cap A_n$ means all of the n events occur.

(3) A' OR A^c means A does not occur.

Interpretation of Probability

Aim: To define a measure (i.e. to quantify) of the likelihood or the chances that an event E will occur. Probability of an event E is denoted as P(E).

There are three interpretations of P(E) or we can say there are three approaches to define probability:

1. Equally likely Model $\Leftrightarrow$ classical approach
2. Relative frequency Model
3. Subjective Probability $\Leftrightarrow$ Axiomatic approach

- Classical Approach:

Let us consider a random expt. with finite sample space Ω i.e number of elementary events in Ω is finite and each elementary event has same chance to occur (assumption of equally likely).

Then, probability that an event E will occur is

$$P(E) = \frac{\#\text{ favourable outcomes}}{\#\text{ Total possible outcomes}} = \frac{|E|}{|\Omega|}$$

where $|B| = \#$ elementary events in B for any $B \subseteq \Omega$.

Example: Suppose a die is thrown twice. What is the probability that the ~~sum~~ total number appearing uppermost is 7?

Solution: If a die is thrown twice then $6^2 = 36$ possible outcomes will be there. Then $|\Omega| = 36$. Out of them only 6 possibilities can make the total 7 and these are

$E = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}$

So, $|E| = 6$

Thus $P(E) = \frac{6}{36} = \frac{1}{6}$

Limitations of this approach:

- For infinite Ω, this approach is not applicable
- It is not always possible to have equally likely outcomes
- To calculate geometrical probability, classical approach fails.

Relative Frequency Approach

This approach is based on the idea that chance behaviour is unpredictable in the short run, but has a regular and predictable pattern in the long run. Thus probability of any event is the proportion of times the outcomes, favourable to that event, would occur in a very long series of repetitions.

Definition: Consider a random experiment with a sample space Ω. We repeat the expt. n times. The probability that the event E will occur is

$$P(E) = \lim_{n \to \infty} \frac{f_n(E)}{n}$$

where $f_n(E)$ = number of times that event E occurs among n trials of the expt.

Advantages:

- Here we do not require only finite Ω.
- Equally likely assumption is also relaxed here.

Limitation:

- One serious limitation is that we must repeat the expt. an infinite number (or very large) of times to obtain a probability. This is not at all realistic.

Remark: James Bernoulli has shown that if outcomes are equally likely, then classical approach and relative frequency approach are equivalent. This is known as Law of Large Numbers (LLN).

Axiomatic Approach:

The modern probability theory is based on the three fundamental principles that are called the axioms of probability theory and probability of an event is defined by a function satisfying those axioms. The first two approaches satisfy these axioms. Thus, the axiomatic approach is more general and any consequences of the axioms are also true for both the previous models.

Before defining probability as a function, we need to understand the domain and range of that function. Range will be clearly understood by the axioms. Since every event is nothing but a subset of Ω, so the domain of this function is the collection of all subsets of Ω, popularly known as sigma-field or Borel-field or sigma-algebra.

σ-field or Borel-field:

A collection of subsets of Ω is called sigma(σ)-field or Borel-field, denoted by $\mathcal{B}$, if it satisfies the following properties:

(i) $\phi \in \mathcal{B}$

(ii) If $A \in \mathcal{B}$, then $A^c \in \mathcal{B}$ ($\mathcal{B}$ is closed under complementation)

(iii) If $A_1, A_2, \ldots \in \mathcal{B}$, then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{B}$ ($\mathcal{B}$ is closed under countable unions)

Remark: 1) Property (ii) and (iii) imply that $\mathcal{B}$ is also closed under countable ~~unions~~ intersections i.e. if $A_1, A_2, \ldots \in \mathcal{B}$ then $\bigcap_{i=1}^{\infty} A_i \in \mathcal{B}$.

2) For finite or countable Ω with m elements, i.e. $|\Omega| = m$, then there are 2^m sets in $\mathcal{B}$ i.e. $|\mathcal{B}| = 2^m$.

Axiomatic definition of probability (by A. Kolmogorov):

Given a sample space Ω and an associated sigma-field $\mathcal{B}$, a set function P with domain $\mathcal{B}$ is called a probability if the following axioms are satisfied:

I. $P(E) \geq 0$ for all $E \in \mathcal{B}$,

II. $P(\Omega) = 1$,

III. If $E_1, E_2, \ldots \in \mathcal{B}$ are mutually exclusive events, then

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

Theorem P is monotone and subtractive, i.e., if $A, B \in \mathcal{B}$ and $A \subseteq B$, then $P(A) \leq P(B)$ and $P(B-A) = P(B) - P(A)$.

Proof: If $A \subseteq B$, then we can write

$$B = (A \cap B) + (B - A) \quad (\because A \cap B = A)$$

$$\Rightarrow \quad P(B) = P(A) + P(B-A)$$

$$= P(A) + P(B \cap A^c)$$

$$\Rightarrow \quad P(B) \geq P(A) \quad \text{and} \quad P(B-A) = P(B) - P(A).$$

Theorem (Principle of Exclusion-Inclusion)

Let $A_1, A_2, \ldots, A_n \in \mathcal{B}$. Then for $n \geq 2$,

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = S_1 - S_2 + S_3 - \ldots + (-1)^{n-1} S_n$$

where $S_1 = \sum_{i \geq 1}^{n} P(A_i)$, $S_2 = \sum_{i<j} P(A_i \cap A_j)$, $S_3 = \sum_{i<j<k} P(A_i \cap A_j \cap A_k)$ and so on

Proof: Suppose $n=2$ at first.

$$P(A_1 \cup A_2) = P(A_1 \cap A_2^c) + P(A_1 \cap A_2) + P(A_1^c \cap A_2)$$

$(\because (A_1 \cap A_2^c), (A_1 \cap A_2)$ & $(A_1^c \cap A_2)$ are mutually exclusive events$)$

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

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

So theorem holds for $n=2$. Let the theorem be true for $n = t (\geq 2)$. We have to show that the theorem will be true for $n = t+1$.

$$P(A_1 \cup A_2 \cup \ldots \cup A_{t+1}) = P(A_1 \cup A_2 \cup \ldots \cup A_t) + P(A_{t+1}) - P\left(\left(\bigcup_{i\geq 1}^{t} A_i\right) \cap A_{t+1}\right)$$

$$= \sum_{i \geq 1}^{t} P(A_i) - \sum_{\substack{i,j \geq 1 \\ i<j}}^{t} P(A_i \cap A_j) + \sum_{\substack{i,j,k \geq 1 \\ i<j<k}}^{t} P(A_i \cap A_j \cap A_k) - \ldots$$

$$\ldots + (-1)^{t-1} P\left(\bigcap_{i\geq 1}^{t} A_i\right) + P(A_{t+1}) - P\left(\bigcup_{i \geq 1}^{t} (A_i \cap A_{t+1})\right)$$

$$= \sum_{i\geq 1}^{t+1} P(A_i) - \sum_{i<j}^{t} P(A_i \cap A_j) + \sum_{i<j<k}^{t} P(A_i \cap A_j \cap A_k) - \ldots$$

$$\ldots + (-1)^{t-1} P\left(\bigcap_{i\geq 1}^{t} A_i\right) - \left\{ \sum_{i \geq 1}^{t} P(A_i \cap A_{t+1}) - \sum_{i<j}^{t} P(A_i \cap A_j \cap A_{t+1}) \right.$$

$$\left. + \sum_{i<j<k}^{t} P(A_i \cap A_j \cap A_k \cap A_{t+1}) - \ldots + (-1)^{t-1} P\left(\bigcap_{i\geq 1}^{t+1} A_i\right) \right\}$$

$$= \sum_{i=1}^{t+1} P(A_i) - \sum_{i<j}^{t+1} P(A_i \cap A_j) + \sum_{i<j<k}^{t+1} P(A_i \cap A_j \cap A_k) - \cdots + (-1)^{t} P\left(\bigcap_{i=1}^{t+1} A_i\right)$$

Thus the theorem holds for $n = t+1$ when it holds for $n = t$.
So, the theorem is true for $n = 2, 3, 4, \cdots$ or $n \geq 2$.

Theorem (Bonferroni's Inequality)

Let $A_1, A_2, \cdots, A_n \in \mathcal{B}$. Then for $n \geq 2$,

$$\sum_{i=1}^{n} P(A_i) - \sum_{i<j}^{n} P(A_i \cap A_j) \leq P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i)$$

Proof: Let $n=2$ at first.

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

$$\leq \sum_{i=1}^{2} P(A_i)$$

Let the above result is true for $n = t$. Then we have to show it will be true for $n = t+1$.

$$P\left(\bigcup_{i=1}^{t+1} A_i\right) = P\left(\left(\bigcup_{i=1}^{t} A_i\right) \cup A_{t+1}\right)$$

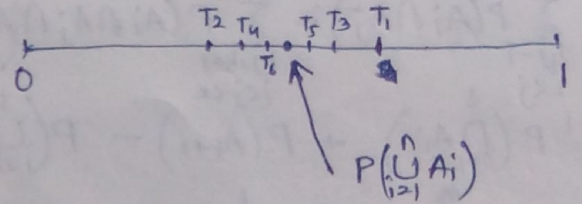
$$= P\left(\bigcup_{i=1}^{t} A_i\right) + P(A_{t+1}) - P\left(\left(\bigcup_{i=1}^{t} A_i\right) \cap A_{t+1}\right)$$

$$\geq \sum_{i=1}^{t} P(A_i) - \sum_{i<j}^{t} P(A_i \cap A_j) + P(A_{t+1}) - P\left(\bigcup_{i=1}^{t} (A_i \cap A_{t+1})\right)$$

$$\geq \sum_{i=1}^{t+1} P(A_i) - \sum_{i<j}^{t} P(A_i \cap A_j) - \sum_{i=1}^{t} P(A_i \cap A_{t+1})$$

$$= \sum_{i=1}^{t+1} P(A_i) - \sum_{i<j}^{t+1} P(A_i \cap A_j)$$

Remark:



where $T_i = \sum_{j=1}^{i} (-1)^{j-1} S_j$.

Another Version of Bonferroni's Inequality:

$$P\left(\bigcap_{i=1}^{n} A_i\right) \geq \sum_{i=1}^{n} P(A_i) - (n-1)$$

[Hint: Replace A_i by A_i^c $\forall i$ in the above Theorem]

Theorem (Boole's Inequality)

For any two events A and B, $P(A \cap B) \geq 1 - (P(A^c) + P(B^c))$

Proof:
$$P(A^c) + P(B^c) = P(A^c \cup B^c) + P(A^c \cap B^c)$$
$$\geq P((A \cap B)^c)$$
$$= 1 - P(A \cap B)$$
$$\Rightarrow P(A \cap B) \geq 1 - (P(A^c) + P(B^c))$$

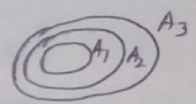
Generalization: Let $\{A_j ; j = 1, 2, 3, \ldots\}$ be a countable sequence of events, then $P(\bigcap_j A_j) \geq 1 - \sum_j P(A_j^c)$

[Proof: Hint: Consider $A = A_1$ and $B = \bigcap_{j=2}^{\infty} A_j$.]

- **Non-decreasing sequence of events:**

Let $\{A_n\}$ be a sequence of events in $\mathcal{B}$ and $\{A_n\}$ be a non-decreasing sequence, that means

$$A_n \supseteq A_{n-1} \quad \text{for } n = 2, 3, \ldots$$

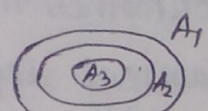


then
$$\lim_{n \to \infty} P(A_n) = P\left(\lim_{n \to \infty} A_n\right) = P\left(\bigcup_{n=1}^{\infty} A_n\right) \quad \text{———} (*)$$

- **Non-increasing sequence of events:**

Let $\{A_n\}$ be a sequence of non-increasing events in $\mathcal{B}$, that means,

$$A_n \subseteq A_{n-1} \quad \text{for } n = 2, 3, 4, \ldots$$



then
$$\lim_{n \to \infty} P(A_n) = P\left(\lim_{n \to \infty} A_n\right) = P\left(\bigcap_{n=1}^{\infty} A_n\right) \quad \text{———} (**)$$

Remark: From the above two properties (*) & (**), we can say that the set function $P(\cdot)$ is continuous from above and below. (*) is "continuity of $P(\cdot)$ from below and (**) is "continuity of $P(\cdot)$" from above.

Combinatorics : Probability on finite sample space

Here we talk only about the sample spaces that have atmost a finite number of elements i.e.,

$$\Omega = \{\omega_1, \omega_2, \ldots\ldots, \omega_n\}$$

- The term "equally likely", which is used in classical definition of probability, is meaning full only in this finite sample space set up and it means

  When Ω contains n elementary events ω_j, $j=1(1)n$ then $P(\{\omega_j\}) = \frac{1}{n}$ $\forall j=1(1)n$

- Different types of sampling schemes from finite sample space (or finite population):

  Number of possible arrangements of size r from n objects in Ω.

	With replacement	Without replacement
Ordered	n^r	$^nP_r = \frac{n!}{(n-r)!}$
Unordered	$\binom{n+r-1}{r}$	$^nC_r = \binom{n}{r} = \frac{n!}{r!\,(n-r)!}$

- If all outcomes in Ω are equally likely, then Axiomatic definition and classical definition of probability give same formula.

  For that situation. Axiom 3 says that for any event A,

$$P(A) = \sum_{\omega_j \in A} P(\{\omega_j\}) = \sum_{\omega_j \in A} \frac{1}{n} = \frac{\#\text{ elements in } A}{\#\text{ elements in } \Omega}$$

Conditional Probability and Independence

So far we have discussed unconditional probabilities. Why unconditional? Because a sample space was defined and all probabilities were calculated w.r.t that sample space. No condition (or restriction) has been imposed on that sample space.

When we have some information prior to obtain the $P(A)$ for an arbitrary event $A \in \mathcal{B}$, associated with sample space Ω, we may want to update the sample space based on the new information.

~~Example: Random experiment: Tossing a coin thrice.~~

~~Sample space: $\Omega = \{(HHH), (HHT), (HTT), (TTT), (TTH), (THH), (THT), (HTH)\}$~~

~~An event of interest: A = Head will ap~~

Example: Random experiment: Throwing a die twice

Sample space : $\Omega = \{(i,j) \; ; \; i = 1(1)6, \; j = 1(1)6\}$

An event of interest : A = sum of the upper face value is 5.

Then set of elementary events favourable to A is

$$F_A = \{\omega_j \in A : \omega_j \in \Omega \;\; \forall j = 1(1)36\}$$

$$= \{(1,4), (4,1), (2,3), (3,2)\}$$

Then unconditional probability of A, denoted as $P(A)$ is $\frac{4}{36} = \frac{1}{9}$

Now, let us consider one information that the upper face value in second throwing appeared as 3 and denote this event as B.

Then updated sample space will be

$$\Omega^{u} = \{(1,3), (2,3), (3,3), (4,3), (5,3), (6,3)\}$$

Then set of elementary events favourable to A, conditional to the given information, is $F_A^u = \{(2,3)\}$

Then conditional probability of A given B, denoted as $P(A|B)$ is $\frac{1}{6}$.

Definition: If A and B are events in Ω and $P(B) > 0$, then the conditional probability of A given B, denoted as $P(A|B)$, is

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Note: If two events A & B are disjoint, then $P(A \cap B) = 0$. Then $P(A|B) = P(B|A) = 0$.

From the definition one can write
$$P(A \cap B) = P(A|B)\,P(B) = P(B|A)\,P(A)$$

Then
$$P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A)} \quad \text{--------- } (***)$$

Theorem (Multiplication Rule)

Let $(\Omega, \mathcal{B}, P)$ be a probability space and $A_1, A_2, \ldots, A_m \in \mathcal{B}$, with $P\left(\bigcap_{j=1}^{m-1} A_j\right) > 0$, then

$$P\left(\bigcap_{j=1}^{m} A_j\right) = P(A_1)\,P(A_2|A_1)\,P(A_3|A_1 \cap A_2) \cdots\cdots P\left(A_m \Big| \bigcap_{j=1}^{m-1} A_j\right)$$

Proof: Proof is simple, try yourself.

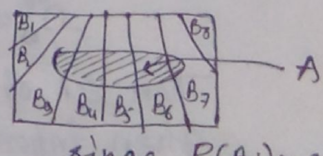
Theorem of Total Probability

Suppose the events $B_1, B_2, \ldots, B_m, \ldots$ are exhaustive and mutually exclusive, i.e., $B_i \cap B_j = \phi \;\; \forall\, i \neq j$ and $\sum_{i=1}^{\infty} B_i = \Omega$. In that case if $P(B_i) > 0 \;\; \forall\, i$, then

$$P(A) = \sum_{i=1}^{\infty} P(A|B_i)\,P(B_i)$$

Proof: The events $A \cap B_i \;\; \forall\, i$ are mutually exclusive as $B_i \;\forall\, i$ are mutually exclusive. and

$$A = \bigcup_{i=1}^{\infty} (A \cap B_i)$$



Then $P(A) = \sum_{i=1}^{\infty} P(A \cap B_i) = \sum_{i=1}^{\infty} P(A|B_i)\,P(B_i)$, since $P(B_i) > 0 \;\forall\, i$

Theorem (Bayes' Rule)

Let $B_1, B_2, \ldots$ be a disjoint sequence of events such that $P(B_i) > 0$, $i = 1, 2, \ldots$, and $\sum_{i=1}^{\infty} B_i = \Omega$. Let $A \in \Omega$ with $P(A) > 0$. Then, for each $i = 1, 2, \ldots$

$$P(B_i | A) = \frac{P(A|B_i) P(B_i)}{\sum_{i=1}^{\infty} P(A|B_i) P(B_i)}$$

Proof: From the definition of conditional probability we have

$$P(B_i | A) = \frac{P(A|B_i) \cdot P(B_i)}{P(A)}$$

Then using the result of "Theorem of Total Probability", we can complete the prove as

$$P(A) = \sum_{i=1}^{\infty} P(A|B_i) P(B_i).$$ (Proof is required here).

Remark: The above rule is a general form of (***). Sometimes we know $P(A|B)$, but our interest is on $P(B|A)$. Then Bayes rule can help us.

Notion of Statistical Independence:

In some situation it may happen that the occurrence of a particular event B, has no effect on the probability of another event A. Notationally we can write

$$P(A|B) = P(A)$$

Then using Bayes' rule also we have $P(B|A) = P(B)$, that means occurrence of A also has no effect on B. This situation implies that

$$P(A \cap B) = P(A) \cdot P(B)$$

Definition: Two events, A and B, are said to be statistically independent if and only if $P(A \cap B) = P(A) \cdot P(B)$.

Generalization of Definition of Statistical Independence:

A collection of events $A_1, A_2, A_3, \ldots, A_n$ are said to be statistically independent if

$$P\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} P(A_i), \text{ for any } n \geq 2$$

- Suppose we have three events A_1, A_2 and A_3. Then if

  (i) $P(A_1 \cap A_2) = P(A_1) \cdot P(A_2)$

  (ii) $P(A_1 \cap A_3) = P(A_1) \cdot P(A_3)$

  (iii) $P(A_2 \cap A_3) = P(A_2) \cdot P(A_3)$ and

  (iv) $P(A_1 \cap A_2 \cap A_3) = P(A_1) \cdot P(A_2) \cdot P(A_3)$.

  hold, then we say that A_1, A_2 and A_3 are mutually independent.

Definition: A collection of events $A_1, \ldots, A_n$ are mutually independent if for any subcollection $A_{i_1}, A_{i_2}, \ldots, A_{i_k}$, we have

$$P\left(\bigcap_{j=1}^{k} A_{i_j}\right) = \prod_{j=1}^{k} P(A_{i_j})$$

Note: Mutually Independence $\Rightarrow$ (Statistical) Independence
But the reverse not true always.

Miscellanea

*1. With Replacement & Unordered sampling from finite population

Suppose $N = 30$ and n, sample size $= 5$

An easiest way to think the problem as placing 5 markers on the 30 numbers. Let 30 numbers define 30 bins as below

```
| M |   |   | MM |   | .... | M |    | M |    |
  1   2   3    4   5   ....  27   28   29   30
```

So, the number of ways to get sample = number of ways that we can put 5 markers into 30 bins. That means ~~the total no. of~~ the arrangement of markers and the walls of the bins.

$\therefore$ Arrangement of 29 walls (30 bins yield 31 walls, but 2 walls play no part) and 5 marker $= (29+5)! = 34!$

By eliminating orderings, we have the answer $= \dfrac{34!}{5!\,29!} = {}^{34}C_5$

*2. Geometric Probability:

In classical approach we have ~~defined~~ calculated probabilities by counting the number of favourable outcomes and dividing that number by the total number of possible outcomes. In this lesson, we will use a geometric measures such as length, area, volume for geometric object instead of counting outcomes.

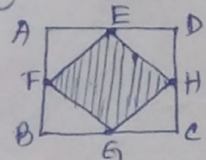
Let $\Omega \subseteq \mathbb{R}^n$ be a given set, and $A \subseteq \Omega$. We ~~are~~ want to find the probability that a randomly chosen point in Ω falls in A. By assuming that both A and Ω have well-defined finite measure (e.g. length, area, volume), we define

$$P(A) = \frac{\text{measure}(A)}{\text{measure}(\Omega)}$$

This probability is called <u>geometric probability</u>.

<u>Example: 1</u> A point is chosen at random from the interval (a,b). Find the probability that it lies in (c,d), $a \le c < d \le b$?
Ans: $(d-c)/(b-a)$.

<u>Example: 2</u> A point is chosen at random in the square of length 4 unit. What is the probability that it lies in EFGH



<u>Note:</u> In advance stage of statistics, we will see that in ~~measure~~ the view of measure theory, we alway define $P(A) = \mu(A)/\mu(\Omega)$, where μ is n-dimensional <u>Lebesgue measure</u>. for usual probability, Lebesgue measure = counting measure, for geometric probability, Lebesgue measure = geometric measure (e.g. ~~a~~ length, area etc.). For length, of course $n=1$; for area, $n=2$ and for volume, $n=3$.

Problems

1. Two friends who take metro to their jobs from the same station arrive to the station uniformly randomly between 9 and 9:20 a.m. They are willing to wait for one another for 5 minutes, after which they take a train whether together or alone. What is the probability of their meeting at the station?

Ans: $7/16$

2. Three points A, B, C are placed at random on a circle of radius 1. What is the probability for ΔABC to be acute?

Ans: $1/4$.