PROBABILITY THEORY

Aims - you will be able to

- characterize probability theory
- explain general notions of probability theory
- explain and use general relations between events
- explain a notion of probability
- define probability by basic axioms
- define properties of probability function
- use a conditional probability
- explain theorem of total probability and Bayes theorem

Basic notions

Probability Theory is the deductive part of statistics.

Its purpose is to give a precise mathematical definition or structure to what has been thus far an intuitive notion of randomness. Making randomness more precise will allow us to make exact probability statements.

Probability theory is mathematical branch whose logical structure is created axiomatically.

Random experiment is every finite process whose result is not given in advance by conditions upon whose is run.

Sample space Ω is a set of all possibly results of the experiment.

- **General notions of the probability theory**

Definition of Set - set A is a collection of elements. Elements are basic intuitive mathematically undefined entities. To define a set, it is necessary to be able to determine whether any element is included or not included in the set. The notion of inclusion is also an intuitive undefined concept.

Definition of Elementary Events - In the case of probability theory, the elements of sets on which probability measures are defined are called elementary events. In practice, these elementary events may be measurement units, cases, or sample points.

Example:

{reverse, obverse} –when tossing the coin

{1,2,3,4,5,6} – when tossing the dice

We denote a set of all results Ω. This set we call **sample space** (of the elementary events)**.** The elementary event $\{\omega\}$ is a subset of the Ω set which contains one element ω from Ω set, $\omega \in \Omega$.

Then the event A will be an arbitrary subset of Ω, $A \subset \Omega$.

What are types of the elementary events?

If the elementary event $\omega \in \Omega$ ($\omega \in A$) came then we can say that an event A came with the experiment realization. We denote this result $\omega \in A$ as **result favourable to the event A**.

Certain event

- is the event which become with every realization of the experiment. It is equivalent with the Ω set.

Certain event is for example: *we toss one of these numbers 1,2,3,4,5,6* (while tossing a dice)

Impossible event

- is the event which can never become in the experiment. We will denote it as $\varnothing$.

Impossible event is for example: *we toss number 8* (while tossing a dice).

What are relations between events?

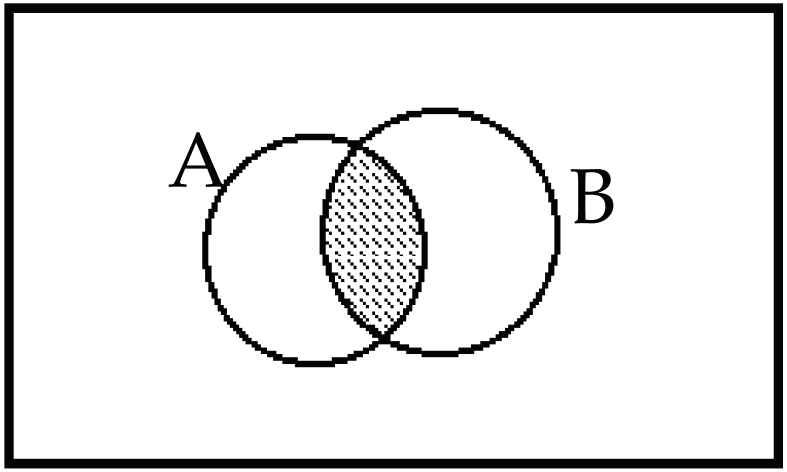
Operations on Sets - The operations of union, intersection, complementation (negation), subtraction, the concept of subset, and the null set and universal set or sample space make up the algebra of sets.

Intersection A ∩ B

The set of all elements that are both in A and in B.

Graphic example:

$$A \bigcap B = \{\omega \mid \omega \in A \wedge \omega \in \mathrm{B}\}$$



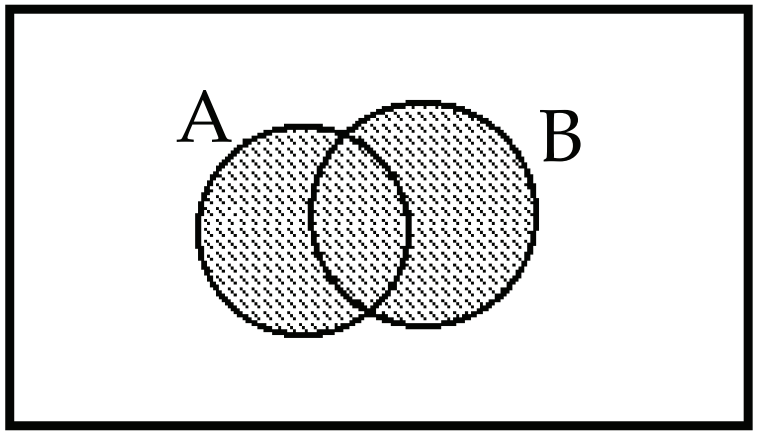
Example – tossing a dice: event A – we toss a number 2 ,3 or 4 and event B – we toss a even number. It is obvious that A ∩ B ={2,4}.

Union A ∪ B

The set of all elements that are either in A or in B.

Graphic example:

$$A \bigcup B = \{\omega \mid \omega \in A \vee \omega \in B\}$$



Example – tossing a dice: Event A = {1,3,4} and event B is when we toss even number. It's obvious that A ∪ B ={1,2,3,4,6}.

Disjoint events $A \cap B = \varnothing$

Two events A and B can't become together. They have none common result. (sometimes called *exclusive*).

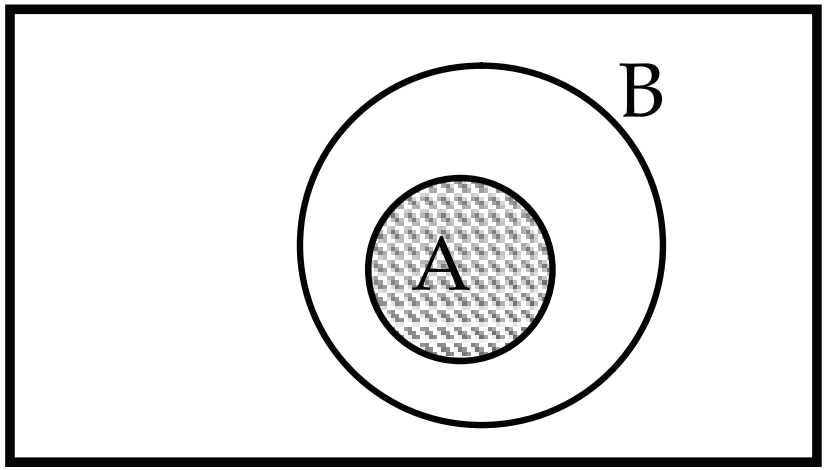
Example – tossing a dice: Event A – we toss even number and even B – we toss odd number. These events never have a same result. If event A become than event B can't become.

Subsets (Subevent) $A \subset B$

A is a subset of B if every element of A is also an element of B. It's mean if event A become than event B become too.

Graphic example:

$$A \subset B \Leftrightarrow \{\omega \in \mathrm{A} \Rightarrow \omega \in B\}$$



Example – tossing a dice: Event A – we toss number 2 and event B – we toss even number. The event A is subevent of the event B.

Events A and B are equivalent $A = B$ if $A \subset B$ and at the same time $B \subset A$.

Example – tossing a dice: Event A – we toss even number, event B – we toss number what is divide of number 2. These events are equivalent.

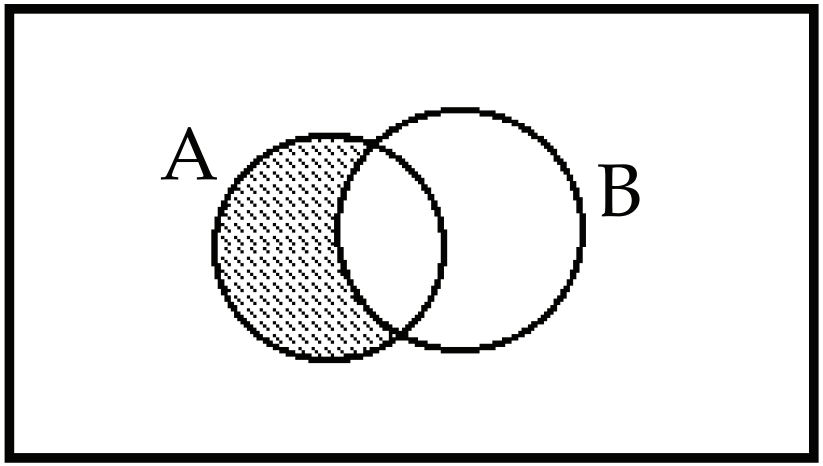
Subtraction A-B

The set of all elements that are in A but not in B

Graphic example:

$$A - \mathrm{B} = \mathrm{A} \cap \overline{\mathrm{B}}$$

$$A - \mathrm{B} = \{\omega \,|\, \omega \in \mathrm{A} \wedge \omega \notin \mathrm{B}\}$$



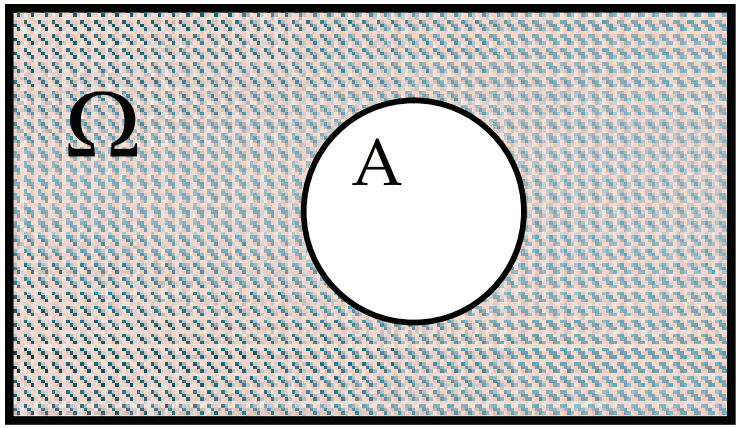
Example – tossing a dice: Event A – we toss a number greater than two and event B – we toss an even number. Subtraction of the events A and B is an event A – B ={3,5}.

Complement of the event A (opposite event)

The set of all elements that are not in A.

Graphic example:

$$\overline{A} = \{\omega \mid \omega \notin A\}$$



Example – tossing a dice: Event A – we toss an even number, then an event $\overline{A}$ - we toss an odd number.

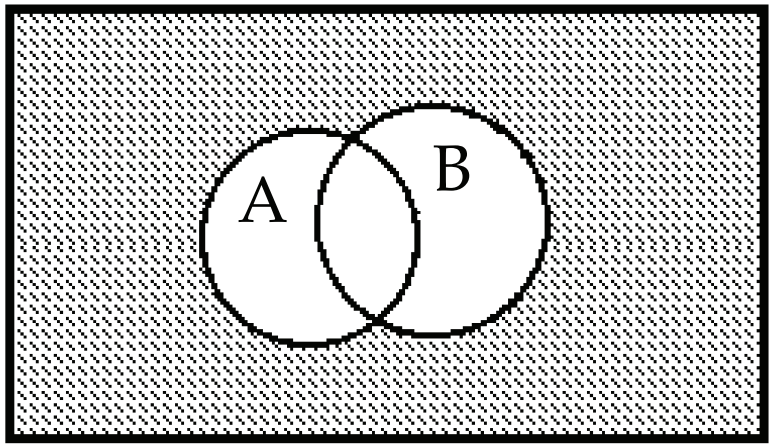
DeMorgan's Laws

- DeMorgan's Laws are logical consequences of the fundamental concepts and basic operations of sets

1. law

The set of all elements that are neither in A nor in B.

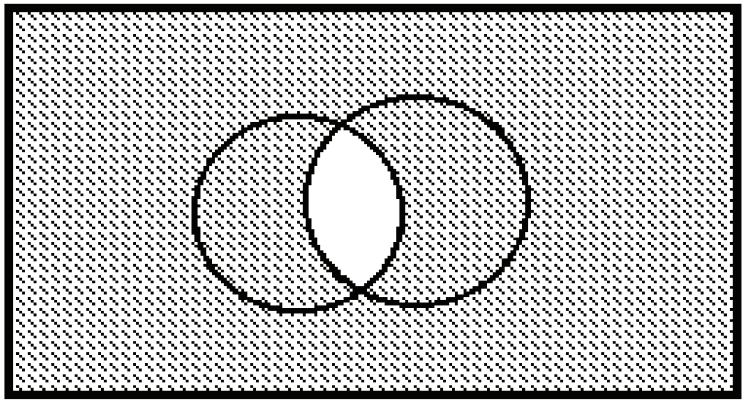
$$\overline{A \cup B} = \overline{A} \cap \overline{B}$$



2.Law:

The set of all elements that are either not in A or not in B.

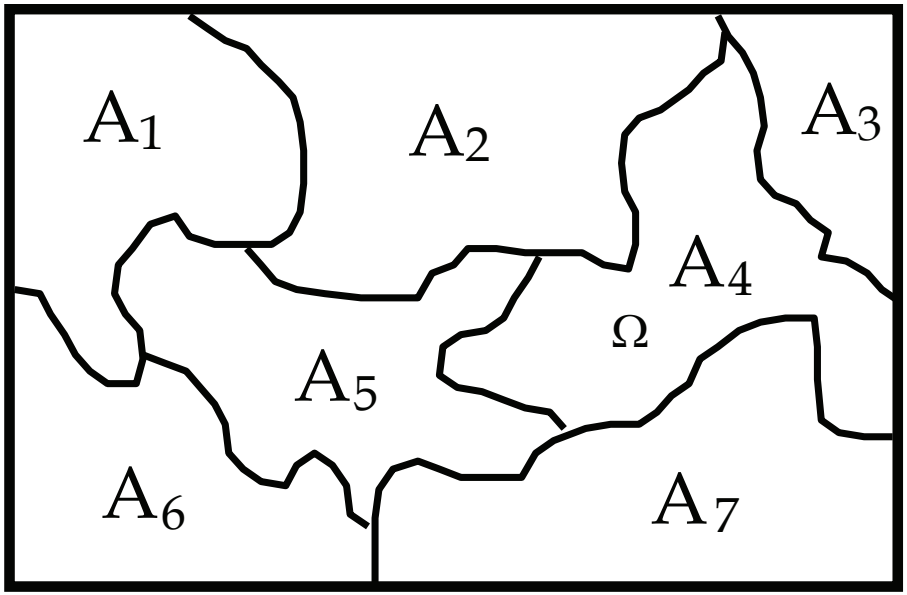
$$\overline{A \cap B} = \overline{A} \cup \overline{B}$$



Mutually disjoint sets and partitioning the sample space

The collection of sets $\{A_1, A_2, A_3, \ldots\}$ partition the sample space Ω:

- $A_i \bigcap A_j = \varnothing$, for $i \neq j$
- $\Omega = \bigcup_{i=1}^{n} A_i$



Probabbility theory

Aim

- notion of probability
- basic theorems and axioms of probability
- types of probability
- conditional probability
- theorem of total probability and Bayes theorem

Notion of probability

Probability of the event A is a number P(A) which has a property that a relative frequency of the event A with increase realizations number is approaching to the number P(A).

This probability definition is known as *classic probability definition.*

Axiomatic probability definition

Probability space is a triad (Ω, S, P) where

(i) Ω is sample space (elements of Ω are elementary events)

(ii) S is a set of subsets of Ω that it holds:

a) $\Omega \in S$;

b) if $A \in S$, then $\overline{A} = \Omega - A \in S$;

c) if $A_1, A_2, A_3, \ldots \in S$, then $\bigcup_{i=1}^{\infty} A_i \in S$

Elements of S we denote as **random events**.

(iii) P is function from S to $< 0,1 >$ such that it holds:

a) $P(A) \geq 0$ for $A \in S$, ... **P is non negative function**

b) $P(\Omega) = 1$ … probabilities are scaled to lie in the interval [0,1];

c) For a collection of mutually disjoint sets, the probability of their union is equal to the sum of their probabilities.

$\{A1, A2, A3, \ldots\} \in S$

$$A_i \cap A_j = \emptyset, \ \forall i, j \in N; \ i \neq j$$

$$\Rightarrow P\{\bigcup_{i=1}^{\infty} A_i\} = \sum_{i=1}^{\infty} P\{A_i\}$$

Function P is called **probability measure** or shorter **probability**.

Especially for two mutually disjoint events A, B:

$$A \cap B = \varnothing, \text{ then}$$

$$P\{A \cup B\} = P\{A\} + P\{B\}$$

Basic properties of probability function

1. If for two events A,B holds: $B \subset A$, then $P\{B\} \leq P\{A\}$
- note that A is partitioned by B and its complement, and hence $P\{A\}$ is sum of these two parts

2. $P(\overline{A}) = 1 - P(A)$ for every $A \in S$;
   - the union of the two sets is the sample space, the intersection is the null sets

3. It holds: $P\{\varnothing\} = 0$
4. It holds: $P\{B - A\} = P\{B\} - P\{B \bigcap A\}$
   - note that B-A and B intersection A are two disjoint sets whose union is B

5. Especially if $A \subset B$, then $P\{B - A\} = P\{B\} - P\{A\}$

6. For arbitrary events A,B holds:

$$P\{A \cup B\} = P\{A\} + P\{B\} - P\{A \cap B\}$$

7. Follows from de Morgan's laws:

$$\begin{aligned} P\{A \cup B\} &= 1 - P\{\overline{A \cup B}\} \\ &= 1 - P\{\overline{A} \cap \overline{B}\} \end{aligned}$$

Definition of Conditional Probability

The definition of conditional probability determines how probabilities adjust to changing conditions. When we say that the condition B applies, we mean that the set B is known to have occurred and therefore the rest of the sample space in the complement of B has zero probability. Under these new circumstances, the revised probability of any other event, A, can be determined from the following definition of conditional probability:

Conditional probability

$$P\{A|B\}$$

$$P\{A|B\} = \frac{P\{A \cap B\}}{P\{B\}}, \quad P\{B\} \neq 0$$

Definition of Independence

If the condition that B has occurred does not affect the probability of A, then we say that A is independent of B.

$$P\{A|B\} = P\{A\},$$

From the definition of conditional probability, this implies

$$P\{A \cap B\} = P\{A\}.P\{B\}$$

That is because

$$P\{A|B\} = \frac{P\{A \cap B\}}{P\{B\}} = \frac{P\{A\}.P\{B\}}{P\{B\}} = P\{A\}, \quad P\{B\} \neq 0$$

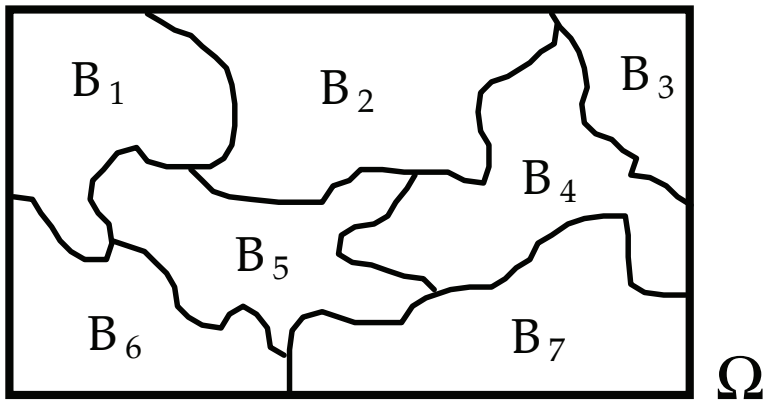
Total Probability Theorem

If a collection of sets $\{B_1, B_2, B_3, \ldots, B_n\}$ partition the sample space Ω, that is,

$$B_i \cap B_j = \varnothing; \forall \; i \neq j$$

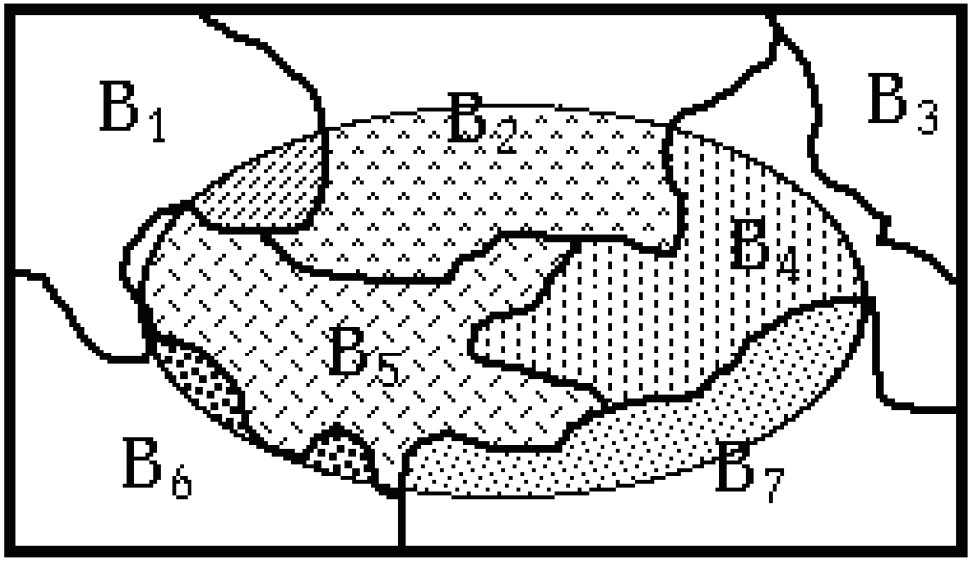
$$\bigcup_{i=1}^{n} B_i = \Omega \qquad ,$$

For example for *n=7:*



then for any set A ($P\{A\}\neq 0$) in the sample space Ω,

$$P\{A\} = \sum_{i=1}^{n} P\{A|B_i\}\, P\{B_i\} = \sum_{i=1}^{n} P\{A \cap B_i\}$$



Proof:

Since the collection of sets $\{B_1, B_2, B_3, \ldots, B_n\}$ partitions the sample space Ω,

$$P\{A\} = \sum_{i=1}^{n} P\{A \cap B_i\} .$$

From the definition of conditional probability

$$P\{A \cap B_i\} = P\{A|B_i\}\, P\{B_i\}$$

Bayes Theorem

If the collection of sets $\{B_1, B_2, B_3, \ldots, B_n\}$ partitions the sample space Ω, then:

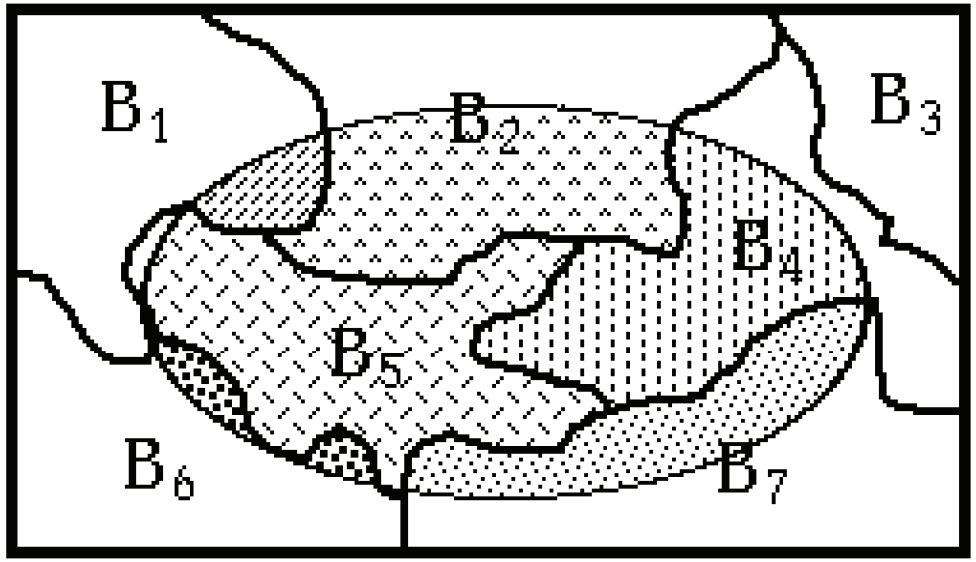
$$P\{B_k|A\} = \frac{P\{A|B_k\}\,P\{B_k\}}{\sum_{i=1}^{n} P\{A|B_i\}\,P\{B_i\}}$$

Proof: From the definition of conditional probability,

$$P\{B_k|A\} = \frac{P\{B_k \cap A\}}{P\{A\}} = \frac{P\{A|B_k\}\,P\{B_k\}}{P\{A\}}$$

Substituting for $P\{A\}$ from the Theorem of Total Probability, the proof follows.

Graphical representation of Bayes theorem :

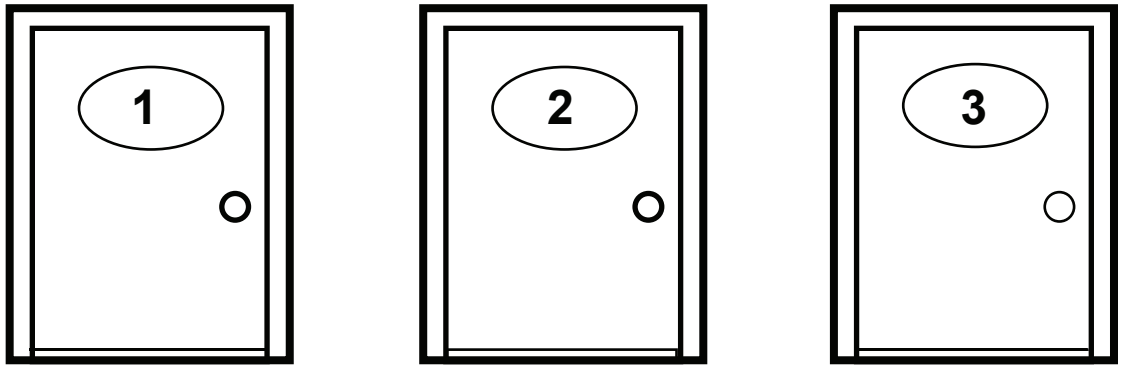


$$P\{B_k | A\} = \frac{P\{B_k \cap A\}}{P\{A\}}$$

Example of an Application of Bayes Theorem

In a famous television game show, the winner of the preliminary round is given the opportunity to enhance his winnings. The contestant is presented with three closed doors and told that behind one of the doors is a new automobile while behind the other two doors are goats. If the contestant correctly selects the door which conceals the automobile, he will win the automobile.

The game show host asks the contestant to make a preliminary selection, after which the host opens one of the other two doors to reveal a goat. The contestant is then given the option of switching his choice to the other door which remains closed. Should he change his choice ?



Solution:

The sample space consists of three possible arrangements {AGG, GAG, GGA}

Assume that each of the three arrangements have the following probabilities:

p_1 = Prob [AGG] p_2 = Prob [GAG] p_3 = Prob [GGA]

where $p_1 + p_2 + p_3 = 1$