CHAPTER 1 – PROBABILITY

2.1 – Introduction

- Why is statistics needed for?

  Give support and recommendations to decision making process with uncertainty.(is likely to be, is probable that…)

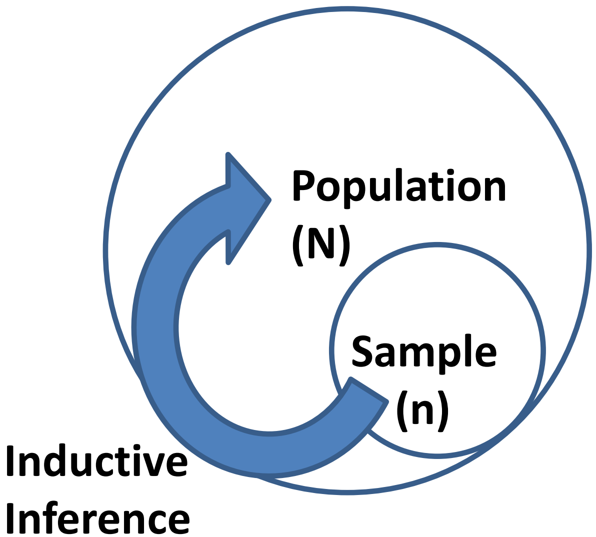
- What do statistics deal with?

  Collection and analysis of data

  Variability - finding a pattern

CHAPTER 1– PROBABILITY

- **Main goal: make statements about the population based in the observation of a sample that have some validity.**



For economic and logistical reasons, to avoid errors of various kinds, only in exceptional cases, the whole population is analysed.

Inductive Inference ⟺ uncertainty

Accuracy ⟺ measure of uncertainty

Population: complete set of items in which the investigator is interested.

Sample: subset of population (randomness)

CHAPTER 1– PROBABILITY

- **Probability theory ⟺ uncertainty measurement.**

  How much likely or probable is it?

- **Random experiment:**
  - The outcome cannot be predicted with certainty.
  - The collection of every possible outcome can be described and, in some cases, listed.

CHAPTER 1 – PROBABILITY

- **Examples of random experiments:**

a) Flip of a coin and observation of head or tails.

b) Toss of a dice and observation of number of spots.

c) A coin is flipped successively at random until the first head is observed and the number of flips counted .

d) Hourly number of phone calls in a call centre.

e) Measure of body fat based on person's weight (kgs) under water.

f) Observation of the daily change (Δ) in an index of stock market prices.

CHAPTER 1– PROBABILITY

Definition 2. 1. Sample Space. Event.

Element or sample point (*s*):

Each of the possible outcomes of a random experiment.

Sample Space (S)

Set of all possible outcomes from a random experiment.

CHAPTER 1 – PROBABILITY

- Sample Spaces can be:

  $a_{1.1}$ **discrete - countable** or **(finite)**

  examples: a) S = {H,T}. $\# S = 2^1$ =2 elements.

  b) S = {1, 2, 3, 4, 5, 6}. $\# S = 6$

  $a_{1.2}$ **discrete - infinite though countable nº of elements**

  examples: c) S=$\{\mathrm{H}, \mathrm{TH}, \mathrm{TTH}, \dots, \mathrm{TT} \dots \mathrm{TH}\}$

  d) S = {0, 1, 2, 3, 4, …}.

  $\boldsymbol{b}$) **continuous** - **uncountable**

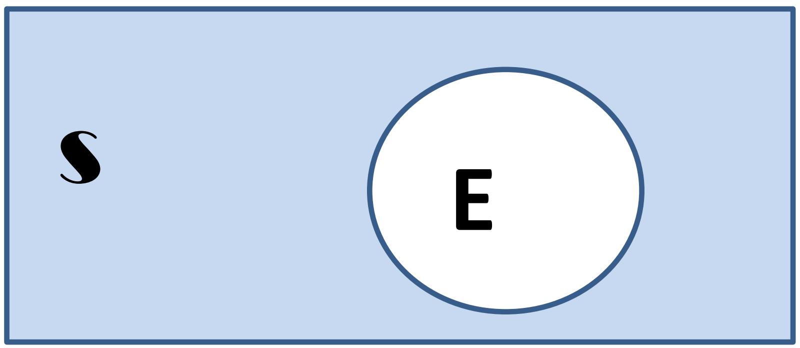
  examples: e) S = $\{s: 0 < s < 7\}$.

  f) S = $\{\Delta: 0 < \Delta < 1\}$

CHAPTER 1– PROBABILITY

Definition 2. 2. - EVENT.

Any subset of the sample space **S**.



$$\boldsymbol{E} = \{s_1, s_2, \cdots, s_p\} \subset \boldsymbol{S}$$

If E = $\{s\}$ *it is an element.*

$\mathcal{S}$ *is an event too.*

An event $\boldsymbol{E} \subset \mathcal{S}$ occurs if the random experiment results is one of its elements

$$\boldsymbol{s_i \in E \quad i = 1, 2, \cdots, p}$$

CHAPTER 1– PROBABILITY

Example: Random Experiment - A coin is flipped twice

Sample space: $\mathbb{S}$ = {HH, HT, TH, TT},

{HH},{HT},{TH},{TT} are elements.

Events (**E** $\subset \mathbb{S}$):

E_1 = {HT,TH} «occurence of exactly one head/tail»;

E_2 = {HH, HT} «occurence of a head in first toss»;

E_3 = {HT,TH,HH} =«occurence of at least one head».

CHAPTER 1– PROBABILITY

Example

random experiment: cast of two dices and observation of number of spots in first (i), and second (j) dices.

Sample space: $\mathbb{S}$ = {(i, j): i, j = 1,2,3,4,5,6}.

$\mathbb{S}$ = 6^2 = 36 elements.

Events:

E_1 = {(4,4),(4,5),(5,4),(5,5)}

«occurence of just 4 or 5 spots in both dices»;

E_2 = {(1,1),(1,2),(1,3),(2,1),(2,2),(3,1)}

«total number of spots less than 5».

CHAPTER 1– PROBABILITY

Example

random experiment: observation of a light bulb life span (hrs)

Sample space: $\mathbb{S} = \{x : x > 0\} \subset \mathbb{R}$

Events:

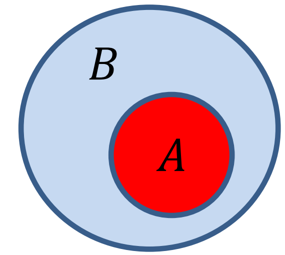
$E_1 = \{x : 17000 < x < 43000\}$;
«life span greater than 17000 and lower than 43000 hours »;

$E_2 = \{x : x \leq 50000\}$ «life span less than or equal to 50000 hrs»;

$E_3 = \{x : x > 60000\}$ «life span greater than 60000 hrs».

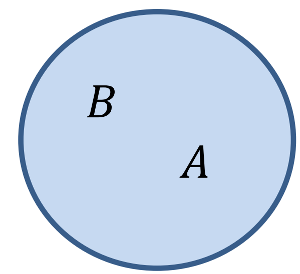
CHAPTER 1– PROBABILITY

- In the study of probability, sets and events are interchangeable.
- Algebra of events = Algebra of sets



Let A, B be events in sample space $\mathbb{S}$

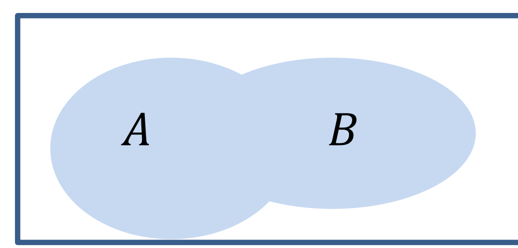
Event A occurs if and only if B occurs. A $\subset$ B.



Event A=Event B if and only if $\forall s \epsilon A \Rightarrow$ s ϵB :

$$A = B \Rightarrow A \subset B \; e \; B \subset A$$

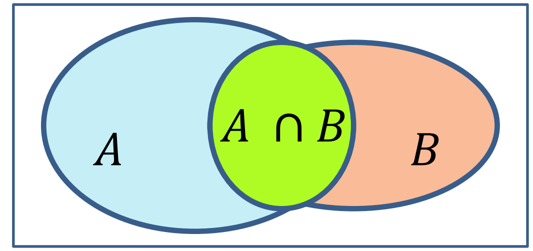
Union of events $\boldsymbol{A} \cup \boldsymbol{B}$
occurs if and only if either A or B or both occur



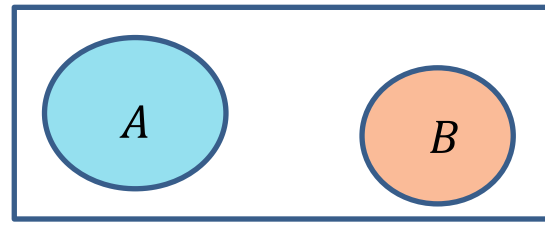
Null event (∅): event with no elements.

CHAPTER 1– PROBABILITY

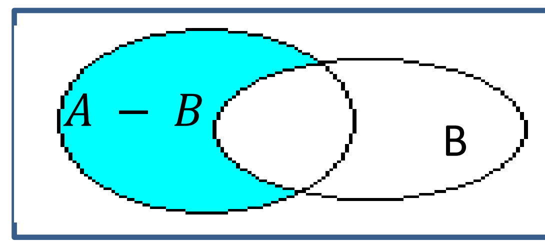
Intersection of events $A \cap B$: is the set of all elements in $\mathbb{S}$ that belong to both A and B



If events A and B have no common elements they are called **Mutually exclusive events.** $A \cap B = \emptyset$

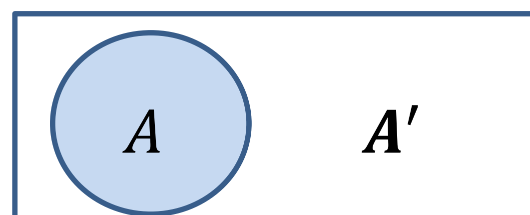


Diference of events $A - B$: is the event that occurs if and only if A occurs but not B.



Complement $\boldsymbol{A'} = \mathbb{S} - A$: the set of elements belonging to $\mathbb{S}$ but not to A. It is obvious that

$$A \cap \boldsymbol{A'} = \emptyset \text{ and } A \cup \boldsymbol{A'} = \mathbb{S}$$



CHAPTER 1– PROBABILITY

Some properties of operations over events:

1. Associativity: $A \cup (B \cup C) = (A \cup B) \cup C$; $A \cap (B \cap C) = (A \cap B) \cap C$

2. Comutativity: $A \cup B = B \cup A$; $A \cap B = B \cap A$

3. Distributivity: $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$;
   $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

4. De Morgan Laws: $\overline{A \cup B} = \bar{A} \cap \bar{B}$; $\overline{A \cap B} = \bar{A} \cup \bar{B}$

5. *Union of k events* $A_{1,} A_{2,} \cdots, A_k = \cup_{i=1}^{k} A_i$

6. *Intersection of k events* $A_{1,} A_{2,} \cdots, A_k \cdots = \cap_{i=1}^{k} A_i$

CHAPTER 1– PROBABILITY

- **Measure of probability. Probability Axioms.**
- **Probable, likely, credible, possible, etc.,**
- **Probability measure the uncertainty degree associated to the occurence of random events .**
- The **measure of the probability** can be formalized through a small number **of postulates.**
- These postulates are due to Bernstein and Kolmogorov decisive contribution. **Kolmogorov Probability postulates .**

CHAPTER 1– PROBABILITY

Definition: Measure of probability

Probability of an event is a real valued set function *P that assigns to each event* A *in the sample space* $\mathbb{S}$ *(*$A \subset \mathbb{S}$*) a real number* $P(A)$*, called probability of event* A*, such that the following postulates are satisfied*:

P1 – $P(A) \geq 0.$

P2 – $P(\mathbb{S}) = 1.$

P3 – If $A_1, A_2, \cdots, A_k$ are events satisfying $A_i \cap A_j = \emptyset$, i $\neq j$, then
$$P(A_1 \cup A_2 \cup A_3 \cup \cdots \cup A_k) = P(A_1) + P(A_1) + \cdots + P(A_k)$$
for a finite or an infinite but countable number of events.

CHAPTER 1– PROBABILITY

Postulate P3 can be generalized saying that:

- **Theorem 2. 1**– If there is *an infinite but countable number of events*

$$A_{1,} A_{2,} \cdots , A_{n,} \cdots \Rightarrow \cup_{i=1}^{\infty} A_i$$

that are mutually exclusive events,

$$A_i \cap A_j = \emptyset \quad (i \neq j)$$

$$then\ P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$$

CHAPTER 1– PROBABILITY

- **Properties of the measure of probability in a sample space S:**

Theorem 2.3. P($\boldsymbol{A'}$) = 1 − P(A).

Theorem 2.4. $P(\emptyset) = 0\ (*)$.

Theorem 2.5. If $A \subset B \Rightarrow P(A) \leq P(B)$.

Theorem 2.6. $0 \leq P(A) \leq 1$.

Theorem 2.7. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Theorem 2.8.

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

These properties are easily proved.

(*) Null probability of an event does not mean that it can't occur.

CHAPTER 1– PROBABILITY

- **Exercise:** A corporation takes delivery of some new machinery that must be installed and checked before it becomes available. The corporation is sure that it will take no more than 7 days for the instalation and check-up to take place.

  Let: $A = \{It\ will\ be\ more\ than\ 4\ days\ before\ the\ machinery\ becomes\ available\}$

  $B = \{It\ will\ be\ less\ than\ 6\ days\ before\ the\ machinery\ becomes\ available\}$

  (a) Describe the event that is the complement of A.

  (b) Are events A and B mutually exclusive?

- **Exercise:** Draw one card from a deck of cards. The sample space is the set of 52 cards.

  Let $A = \{x: x\ is\ a\ jack, queen\ or\ king\}$; $B = \{x: x\ is\ a\ 9\ or\ 10\ and\ red\}$

  $C = \{x: x\ is\ a\ club\}$; $D = \{x: x\ is\ a\ diamond, heart\ or\ spade\}$

  Find: (a) $P(A)$, (b) $P(A \cap B)$, (c) $P(A \cup B)$, (d) $P(C \cup D)$, (e) $P(C \cap D)$.

- **Exercise:** If $P(A) = 0.4$; $P(B) = 0.5$; $P(A \cup B) = 0.7$

  Find $P(A \cap B)$, $P(A - B)$, $P(A')$, $P(A' \cup B')$, $P(A' \cap B')$

CHAPTER 1– PROBABILITY

- **Exercise:** Let x equal a number randomly selected from the closed interval $[0, 1]$. Use your intuition to assign values to:

  (a) $P(x\colon 0 \leq x \leq {}^{1}/_{3})$, (b) $P(x\colon {}^{1}/_{3} \leq x \leq 1)$, (c) $P(x\colon x = {}^{1}/_{3})$

  (d) $P(x\colon {}^{1}/_{2} < x < 5)$

- **Exercise**: A point is picked “at random” from a unit square centered in (0,0)

  (a) Define the sample space. $S = \{(x, y)\colon -{}^{1}/_{2} \leq x \leq {}^{1}/_{2}, -{}^{1}/_{2} \leq y \leq {}^{1}/_{2}\}$

  (b) For the probability assignment if $A \in S$, we assign $P(A) = \acute{a}rea_{\,A}$.

  Find $P(A)$:

  $b_1)\ A = \{(x, y)\colon 0 \leq x \leq {}^{1}/_{2}, 0 \leq y \leq {}^{1}/_{2}\}$,

  $b_2)\ A = \{(x, y)\colon (x - {}^{1}/_{2})^2 + (y - {}^{1}/_{2})^2 = 1\}$

  $$b_3)\ A = \begin{Bmatrix} set\ of\ all\ points\ that\ are\ at\ the\ most \\ a\ unit\ distance\ from\ origin \end{Bmatrix}$$

CHAPTER 1– PROBABILITY

- Except for the probability of the *null* event (*Ø)* and the *sample space (S)*, probability theory doesn't say how to assign probabilities to events.

- It only says how to find probabilities of certain events from the probabilities of other, usually simpler, events.

- The assignment of probabilities to an event depends on the knowledge about the random experiment and the interpretation that is given to the concept of probability.

- 3 main interpretations of the concept of probability : **classical, relative frequency and subjective.**

CHAPTER 1– PROBABILITY

Definitions of the concept of probability .

1. **Classical (De Moivre 1718, Laplace** beggining of XIX century – gambling problems).

   Probability of an even is the proportion of times that the event occurs assuming that all elements in the sample space are equally likelly to occur.

main features:

Sample Space $\mathbb{S}$ is **finite**, $\mathbb{S}$ = $\{s_1 , s_2 , \cdots , s_N\}$

Principle of symmetry

(all outcomes\elements in the sample space **are equally likelly** to occur**).**

$$\boldsymbol{P(s_i) = \frac{1}{N} \quad (i = 1, 2, \cdots , N)}$$

Let event A = $\{s_3 , s_5 , \cdots , s_k\}$ $(k < N)$

$$\boldsymbol{P(A) = \frac{N_A}{N} = \frac{\#A}{\#\mathbb{S}} = \frac{number\ of\ outcomes\ satisfying\ A}{total\ number\ of\ outcomes\ in\ the\ sample\ space}}$$

CHAPTER 1– PROBABILITY

- **Example: toss of two "fair" coins, $\mathbb{S}= \{HH, HT, TH, TT\}$.**

  Let, $A = \{occurence\ of\ exactly\ one\ Head\ or\ Tail\} = \{HT, TH\}$;

  $B = \{occurence\ of\ one\ Head\ in\ the\ 1^{st} coin\} = \{HH, HT\}$;

  $C = \{occurence\ of\ at\ least\ one\ tail\} = \{HT, TH, TT\}$;

  Find $P(A)$, $P(B)$, $P(C)$.

- **Example: A manager has a group of eight employees who could be assigned to a project-monitoring task. Three are women and five are men. Two of the men are brothers.**

  Let, $A = \{chosen\ employee\ is\ a\ man\}$;

  $B = \{chosen\ employee\ is\ one\ of\ the\ brothers\}$;

  Find $P(A)$, $P(B)$, $P(A \cap B)$, $P(A \cup B)$.

- **Consider the course of the Dow-Jones average over two days.**

  Let $\mathbb{S} = \{O_1 , O_2, O_3 . O_4\}$; O_1=avg rises on both days; O_2=avg rises on 1st day but not in 2nd; O_3=avg rises 2nd day but not in 1st; O_4=avg doesn't rise on either day. Let, $A = \{market\ will\ rise\ on\ at\ least\ one\ of\ the\ days\}$.

CHAPTER 1– PROBABILITY

- To deal with more complex sample spaces a powerful tool are the Methods of Enumeration : permutations and combinations.

- The classical definition was in rule until the beggining of XX century.

- Two main drawbacks of the classical definition :

  – The principle of symmetry is seldom verified ;
  – Many sample spaces are not finite or countable.

- The **relative frequency definition** tried to overcome the criticisms above. It was adopted since early XX century and is the dominant definition till now.

CHAPTER 1– PROBABILITY

Relative frequency definition of probability

- Consider repeating a random experiment $\boldsymbol{n}$ times.

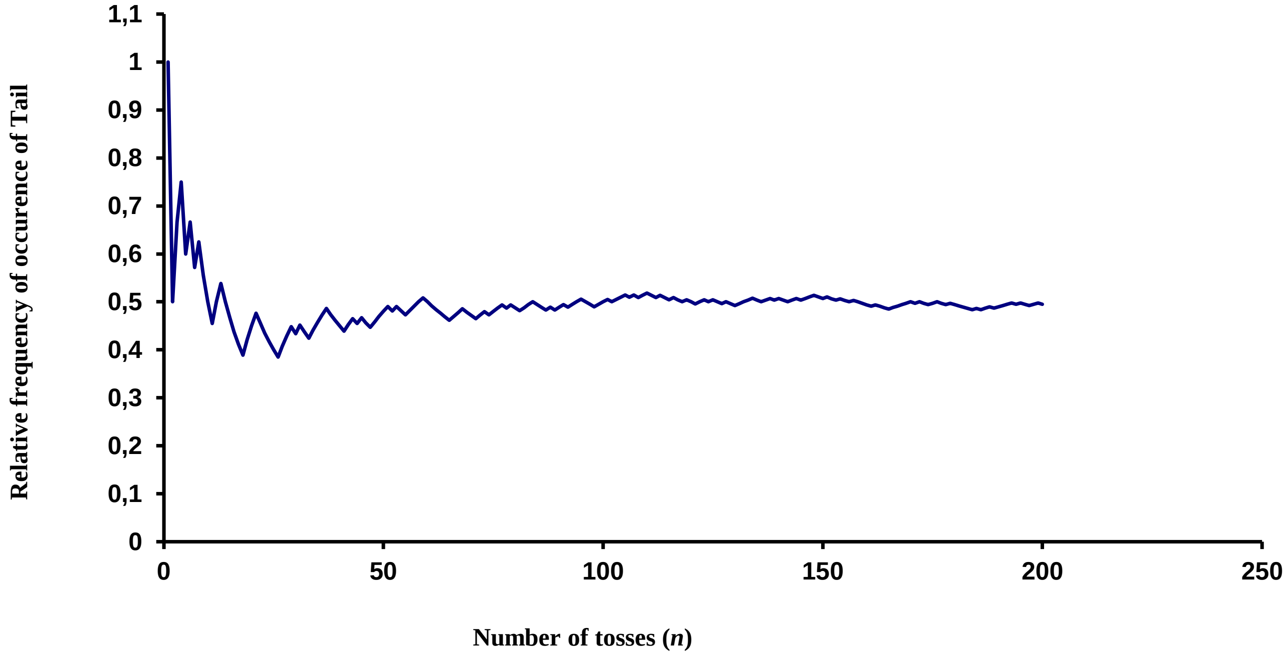
- Count the number of times that event $\boldsymbol{A}$ occured throughout these $\boldsymbol{n}$ repetitions. This number n_A is called the frequency of event $\boldsymbol{A}$.

- The ratio $\frac{n_A}{n}$ is called relative frequency of event $\boldsymbol{A}$.

- The relative frequency is usually very unstable for small values of $\boldsymbol{n}$ but it tends to stabilize for large $\boldsymbol{n}$.

CHAPTER 1– PROBABILITY

- **Example –** A regular coin is tossed 200 times. The number of times that a Tail occured troughout these $\boldsymbol{n}$ repetitions was counted. The relative frequency of occurence of a Tail calculated for $\boldsymbol{n} = 1, 2, \cdots, n$.



- The probability of event $\boldsymbol{A}$ is the limit of the relative frequency of event $\boldsymbol{A}$ in a large number of trials.

$$P(A) = \lim_{n \to \infty} n_A / n.$$

CHAPTER 1– PROBABILITY

- In a more formalized way we can say that:

  For every $\varepsilon > 0$, there is a number n of experiments so that

$$P(A) - \varepsilon < f(A) < P(A) + \varepsilon,$$

is virtually certain

or

$$\lim_{n\to\infty} P[|f(A) - P(A)| < \epsilon] = 1$$

This remarkable result is called the **law of large numbers**.

When ***n*** is set, the relative frequency of an event verify the Kolmogorov postulates.

CHAPTER 1– PROBABILITY

- The **subjective definition** of probability deal with the problem of calculating probabilities when the experiment is not symmetric and can't be successively repeated.

- It expresses an individual's degree of belief about the chance that an event will occur

- These subjective probabilities are used in certain management decisions.

  Examples:

  - profitability of a certain project exceed $x\%$.
  - A certain marketing plan will provide in two years a turnover of approximately y thousands euros.
  - The inflation rate next year will be 1%.
  - The exports growth rate in the next quarter will be 2%.

CHAPTER 1– PROBABILITY

Subjective probabilities may be determined by finding the odds at which people will consider it fair to bet on the outcome.

Example: A high school principal feels that the odds are 7 to 5 against her getting 100€ raise and 11 to 1 against her getting 200€ raise. Furthermore, she feels that it is an even money bet that she will get one of these raises or the other.

$$P\Big(\underbrace{\text{her getting 100€ raise}}_{A}\Big) = 5/12$$

$$P\Big(\underbrace{\text{her getting 200€ raise}}_{B}\Big) = 1/12$$

$$P\Big(\underbrace{\text{getting one of these raises or the other}}_{A \cup B}\Big) = 6/12 = 5/12 + 1/12$$

So the probabilities are consistent

CHAPTER 1– PROBABILITY

- Example: A party official states that the odds are 2 to 1 that he will run for the House of Representatives, and 4 to 1 that he will not run for the Senate. Furthermore, he feels that the odds are 7 to 5 that he will run for one or the other.

$$P\Big(\underbrace{\text{run for the House of Representatives}}_{A}\Big) = 2/3$$

$$P\Big(\underbrace{\text{run for the Senate}}_{B}\Big) = 1/5$$

$$P\Big(\underbrace{\text{run for one or the other}}_{A \cup B}\Big) = 7/12$$

$P(A \cup B) \neq P(A) + P(B)$, so the probabilities are not consistent

CHAPTER 1– PROBABILITY

- **Methods of Enumeration**

Theorem 1.1 Multiplication Principle
If an operation consists of two steps, of which the 1st can be done in n_1 ways and for each of these the 2nd one can be done in n_2 ways, then the whole operation can be done in $n_1 \times n_2$ ways.

- Extending to a sequence of more than 2 experiments suppose that a random experiment E_i has $m_i, i = 1,2, \cdots, k$ possible elements.

Theorem 1.2 If an operation consists of k steps of which the 1st can be done in n_1 ways, for each of these the 2nd step can be done in n_2 ways, and so forth, then the whole operation can be done in

$$n_1 \times n_2 \times \cdots \times n_k \text{ elements.}$$

CHAPTER 1– PROBABILITY

- **3 important questions** in the choice of the right method of enumeration**:**

1. The order of the selection is noted or not?
2. The sampling is with replacement?
   - replacement occurs when an object is selected and then replaced before the next object is selected.
3. The sampling is without replacement?
   - sampling without replacement occurs when an object is not replaced after being selected.

CHAPTER 1– PROBABILITY

Methods of enumeration:

Definition 1.1 Permutations.

A permutation is a distinct arrangement of n different elements of a set

Suppose that $\boldsymbol{n}$ positions are to be filled with $\boldsymbol{n}$ distinct objects.

There are $\boldsymbol{n}$ choices for the 1st position, $\boldsymbol{n}$-1 choices for the 2nd position, ⋯, 1 choice for the last position. So by the multiplication principle, the number of possible arrangements are

$$n(n-1)(n-2)\cdots(2)(1)=n!$$

By definition we take $0!=1$

Theorem 1.3

The number of permutations of n distinct objects is $n!$

CHAPTER 1– PROBABILITY

- **Permutation** of n objects (**ordered sampling is without replacement**)

  Example: The number of permutations of the four-letters a,b,c,d $=4! = 24$

- **Permutation** of n objects (**ordered sampling is with replacement**)

  The number of possible arrangements are $\boldsymbol{n^n}$

  Example: The number of possible four-letters code words using a,b,c,d $=4^4 = 256$

CHAPTER 1 – PROBABILITY

Ordered sampling without replacement

- **Theorem 1.4 Permutation** of $\boldsymbol{n}$ objects taken **r** at a time

If only $\boldsymbol{r}$ positions are to be filled with $\boldsymbol{n}$ distinct objects $\boldsymbol{r} \leq \boldsymbol{n}$, the number of possible **ordered** arrangements is

$$_{n}P_{r} = n(n-1)(n-2)\cdots(n-r+1)! = \frac{n!}{(n-r)!}$$

Example 1: With pieces of cloth of 4 different colours how many distinct three band vertical coloured- flags can one make if the colours can’t be repeated?

$$_{4}P_{3} = \frac{4!}{(4-3)!} = 24$$

Example 2: The number of ways of selecting a president, a vice president, a secretary and a treasurer in a club of 10 persons is

$$_{10}P_{4} = \frac{10!}{(10-4)!} = 5040$$

CHAPTER 1 – PROBABILITY

Non ordered sampling without replacement

- **Theorem 1.7**

  The number of combinations of $\boldsymbol{n}$ objects taken $\boldsymbol{r}$ at a time is

$$C_k^n = \binom{n}{k} = \frac{n!}{k!\,(n-k)!}$$

Example 1: In a restaurant there are 10 different menus. How many subsets of 5 different menus can we make from the 10 menus available?

$$C_5^{10} = \frac{10!}{5!(10-5)!} = 36$$

Example 2: The number of possible 5-card hands (in five card poker) drawn from a deck of 52 playing cards is

$$\binom{52}{5} = \frac{52!}{5!\,(52-5)!} = 2598960$$

CHAPTER 1 – PROBABILITY

- **Theorem 1.8 Distinguishable Permutation**

  Suppose a set of $\boldsymbol{n}$ objects of $\boldsymbol{r}$ distinguishable types .

  From these $\boldsymbol{n}$ objects, k_1 are similar, k_2 are similar, $\cdots$, k_r are similar, so that $k_1 + k_2 + \cdots + k_r = n$,

  The number of distinguishable permutations of these $\boldsymbol{n}$ objects is:

$$\frac{n!}{k_1! \times k_2! \times \cdots \times k_r!}$$

Multinomial Coeficient

- When r =2 , it is called **Binomial Coeficient** **=**$C_k^n = \frac{n!}{k_1!(n-k_1)!}$.

CHAPTER 1 – PROBABILITY

- **Distinguishable Permutation**

Example 1: With 9 balls of 3 different colours, 3 black, 4 green and 2 yelow, how many distinguishable groups can one make?

$$\frac{9!}{3! \times 4! \times 2!}$$

Example 2: Suppose that you have a collection of 12 objects, 3 of type A, 4 of type B and 5 of type C. The number of distinguishable permutations of these 12 objects is

$$\frac{12!}{3! \times 4! \times 5!}$$

CHAPTER 1 – PROBABILITY

- A particularly important application of the methods of enumeration is the resolution of the following problem:
- Consider a population of $\boldsymbol{N}$ elements, from which $\boldsymbol{M}$ has a certain attribute. Suppose that a sample with $\boldsymbol{n}$ elements $(n > 0)$ was selected, what is the probability that there are $\boldsymbol{x}$ $(0 \leq x \leq n)$ elements in the sample with this particular attribute?
- The answer depends on the type of sampling :

Without replacement $\rightarrow$ **hipergeometric model**

or

With replacement $\rightarrow$ **binomial model**

CHAPTER 1 – PROBABILITY

- **Sampling without replacement $\rightarrow$ hipergeometric model**

number of possible arrangements: Combination of $\boldsymbol{n}$ objects taken $\boldsymbol{n}$ at a time

the population has $\boldsymbol{M}$ elements with a certain attribute and, $N - M$

without it; the sample has $\boldsymbol{x}$ elements with such attribute and

$\boldsymbol{n} - \boldsymbol{x}$ without it. The number of **favorable arrangements** is,

$$\binom{M}{x}\binom{N-M}{n-x} \quad x \geq max\{0, n-(N-M)\} \text{ e } x \leq min\{n, M\}$$

Then the probability that there are $\boldsymbol{x}$ elements in the sample with this particular attribute is:

$$P(x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

CHAPTER 1 – PROBABILITY

- **Sampling with replacement $\rightarrow$ Binomial model**

  The probability that there are $\boldsymbol{x}$ elements in the sample with this particular attribute is:

$$P_r = \frac{M^x (N-M)^{n-x} \binom{n}{x}}{N^n} = \binom{n}{x} \frac{M^x}{N^x} \frac{(N-M)^{n-x}}{N^{(n-x)}} = \binom{n}{x} p^x q^{(n-x)}$$

  where $p = M \;/\; N$ is the proportion of elements in the population with the attribute and $q = 1 - p$ is the proportion of elements in the population without the attribute.

- In the binomial model only the value of $\boldsymbol{p}$ matters. The values of $\boldsymbol{N}$ and M are not important.

CHAPTER 1 – PROBABILITY

- **Hipergeometric model**

  example: In a lake there are 100 fishes from which 10 are red. If a fisherman take 5 fishes out of the lake, randomly and **without replacement**, what is the probability that none of them are red?

$$P_{red} = \frac{\binom{10}{0}\binom{100-10}{5-0}}{\binom{100}{5}}$$

- **Binomial**

  example : Consider the same lake. If a fisherman take 5 fishes out of the lake, randomly and **with replacement**, what is the probability that two of them are red?

$$P_{red} = \binom{5}{2} 0.1^2 \times 0.9^3$$

CHAPTER 1 – PROBABILITY

- **Condicional Probability of event A given the sample space $\mathbb{S}$**
- Two dice are cast. We are interested in the sum of spots of both dice. Let event $\boldsymbol{A}$ = « get a sum of 5 spots»

$$P(A) = P(Sum\ of\ spots = 5) = 4/36 = 1/9$$

Let event $\boldsymbol{B}$ be the number of spots in one of the dice is 4.

Once it occurred, probability of event A should be reassessed because the **sample space narrowed** since it is now composed only by all $(4,*)$ or $(*, 4)$. The occurrence of event $\boldsymbol{A}$ will now happen only if the other die has one spot.

$$P(Sum\ of\ spots = 5|\ one\ of\ the\ dice\ got\ 4) = 2/12 = 1/6$$

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

CHAPTER 1 – PROBABILITY

- **Definition 2.4. - Condicional Probability**

Let $\boldsymbol{A}$ and $\boldsymbol{B}$, be any two events in a sample space $\boldsymbol{\mathcal{S}}$, $P(\boldsymbol{A}) > 0$. The **conditional probability** of event $\boldsymbol{B}$ given that event $\boldsymbol{A}$ has occurred is denoted by the simbol $P(A|B)$ and is defined by :

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

The **conditional probability is a measure of probability** and verifies the Kolmogorov postulates.

The **conditional probability** can be seen as a reassessement of the probability of an event when there is information about the occurrence of another event. Once one got the new information that event $\boldsymbol{B}$ has occurred, the sample space is no longer $\boldsymbol{\mathcal{S}}$ but a subset of $\boldsymbol{\mathcal{S}}$, associated to the occurrence of event $\boldsymbol{A}$.

CHAPTER 1 – PROBABILITY

- Previous example:

$$P(B) = \frac{1}{6} + \frac{1}{6} \ \ casos\ (4,*)\ e\ (*,4)$$

$$P(A \cap B) = 2/36\ casos\ (4,1)\ e\ (1,4)$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{2/36}{1/3} = \frac{1}{6}$$

Example – Totoloto has 49 numbers. Playing with 6 numbers the probability of getting a 1st prize is

$$\frac{1}{\binom{49}{6}} = 6!/(49 \times 48 \times \cdots \times 44) = 7{,}15 * 10^{-8}$$

Hitting the first number, the conditional probability of winning the first prize changes to

$$5!/(48 \times 47 \times \cdots \times 44) = 5{,}84 * 10^{-7}$$

CHAPTER 1 – PROBABILITY

Example : Suppose that we are given 20 tulip bulbs very similar in appearance and told that 8 will bloom early and 13 will be red in accordance with the following combinations:

	Bloom		**Total**
Colour	Early	Late	
Red	5	8	13
Yellow	3	4	7
Total	8	12	20

If one bulb is selected at random, what is the probability that it will produce a red tulip? #Sample space($\mathbb{S}$) =20, number of favorable events=13, so $P(Red) = 13/20$.

Suppose that a close examination of the bulb reveals that it will bloom early. After this information is known, the new sample space (E) is the set of all bulbs that bloom early. $\#E = 8$, number of favorable events=5, so $P(Red|Early) = 5/8$.

CHAPTER 1 – PROBABILITY

- **Theorem 2. 9. - Multiplication rule:**

  The probability that two events, A and B in sample space S, occur simultaneously is

$$P(A \cap B) = P(A|B) \times P(B) = P(B|A) \times P(A)$$

with $P(B) \neq 0$ and\or $P(A) \neq 0$

This rule can easily be generalized to three or more events.

- **Theorem 2. 10.**

  If A, B and C are any three events in a sample space S such that $P(A \cap B) \neq 0$, then

$$\begin{aligned} P(A \cap B \cap C) &= P[(A \cap B) \cap C] \\ &= P(A \cap B) \times P(C|A \cap B) \\ &= P(A) \times P(B|A) \times P(C|A \cap B) \end{aligned}$$

CHAPTER 1 – PROBABILITY

- Four cards are to be dealt successively at random and without replacement from a deck of playing cards. The probability of receiving in order a spade (S), a heart (H), a diamond (D) and a club(C) is

$$P(S \cap H \cap D \cap C) =$$
$$= P(S) \times P(H|S) \times P(D|S \cap H) \times P(C|S \cap H \cap D)$$
$$= \frac{13}{52} \times \frac{13}{51} \times \frac{13}{50} \times \frac{13}{49}$$

CHAPTER 1 – PROBABILITY

- **Definition 2.5. Independence**

Let A and B be any two events. These events are said to be **statistically independent** if and only if

$$\boldsymbol{P(A \cap B) = P(A) \times P(B)}$$

From the multiplication rule it follows that

$$P(A|B) = P(A) \quad if\ P(B) > 0$$

$$P(B|A) = P(B) \quad if\ P(A) > 0$$

CHAPTER 1 – PROBABILITY

- Toss a coin twice and observe the sequence of head and tails.

$$\mathbb{S} = \{HH, HT, TH, TT\}$$

$A = \{heads\ on\ the\ first\ toss\} = \{HH, HT\}$

$B = \{tails\ on\ the\ second\ toss\} = \{HT, TT\}$

$C = \{Tails\ on\ both\ tosses\} = \{TT\} \subset B \Rightarrow P(C) = 1/4$

$$P(A) = P(B) = \frac{1}{2}$$

$P(B \cap C) = \frac{1}{4} \neq P(B) \times P(C) = 1/8$ so the events B and C are not statistically independent.

$P(A \cap C) = P(\emptyset) = 0 \neq P(A) \times P(C) = 1/8$ events A and C are not statistically independent.

$P(A \cap B) = \frac{1}{4} = P(A) \times P(B)$, so events A and B are statistically independent.

CHAPTER 1 – PROBABILITY

- **Statistically independent events and mutually exclusive events mean different relations between events:**

- If $P(A) = 0$, then A is an event independent from any other event. Independent events and mutually exclusive events mean the same only and only if the above situation occurs.

- If $P(A) > 0$ e $P(B) > 0$, and events A e B, are mutually exclusive events then A e B are statistically dependent events, since
$$P(A \cap B) = 0 \neq P(A) \times P(B).$$
- **Independence is a property of the probability measure. It is not a property of events.**
- It is not possible to define independence based on events only.

CHAPTER 1 – PROBABILITY

- **Theorem $\mathbf{2.11}$.** If A and B are statistically independent events, A and B', A' and B, A' and B' are statistically independent events too.

example:

$$P(A \cap B') = P(A - B) = P(A) - P(A \cap B)$$
$$= P(A) - P(A) \times P(B) = P(A)\big(1 - P(B)\big) = P(A) \times P(B')$$

When three events, A, B e C, are considered the following situations may occur :

1 - Events are two by two independent but

$$P(A \cap B \cap C) \neq P(A) \times P(B) \times P(C).$$

2 - $P(A \cap B \cap C) = P(A) \times P(B) \times P(C)$ but A and B or A and C or B and C are statistically dependent events .

3 - Events are two by two independent and

$$P(A \cap B \cap C) = P(A) \times P(B) \times P(C).$$

CHAPTER 1 – PROBABILITY

- Example 1: A box has four balls (1, 2, 3,4). One ball is taken out and the number observed. $\mathbb{S} = \{1, 2, 3, 4\} \Rightarrow \#\mathbb{S} = 4$

Let $A = \{1,2\}, B = \{1,3\}, C = \{1,4\}$ be events from Ω

$P(A) = P(B) = P(C) = 1/2;$

$P(A \cap B) = 1/4 = P(A) \times P(B), P(A \cap C) = 1/4 = P(A) \times P(C),$
$P(B \cap C) = 1/4 = P(B) \times P(C)$ but

$P(A \cap B \cap C) = 1/4 \neq P(A) \times P(B) \times P(C) = 1/8$

Example 2: A regular coin is tossed 3 times. Let A={CCC,FCC,CFF,**FFF**}, B={CCF,CFC, FFC, **FFF**} and C={CCF, CFC, FCF, **FFF**} be events from

$\mathbb{S} = \{CCC, CCF, CFC, FCC, CFF, FCF, FFC, FFF\} \Rightarrow \#\mathbb{S} = 8$

$P(A) = P(B) = P(C) = 1/2\,;$
$P(A \cap B \cap C) = 1/8 = P(A) \times P(B) \times P(C) = 1/8$ but
$P(B \cap C) = 3/8 \neq P(B) \times P(C) = 1/4$

CHAPTER 1 – PROBABILITY

- Example: A regular coin is tossed 3 times.

$$\mathbb{S} = \{CCC, FCC, FCF, FFC, CFC, CCF, FCC, FFF\};\ \#\mathbb{S} = 8$$

Let $A = \{CCC, FCC, CFF, FFF\}$, $B = \{CCF, CFC, FCC, FFF\}$ and $C = \{CCC, FCC, CCF, FFC\}$ be events of the sample space.

$A \cap B = \{FCC, FFF\}$; $A \cap C = \{CCC, FCC\}$; $B \cap C = \{CCF, FCC\}$;

$A \cap B \cap C = \{FCC\}$

Then: $P(A) = P(B) = P(C) = 1/2$

$P(A \cap B) = \frac{1}{4} = P(A) \times P(B)$; $P(A \cap C) = \frac{1}{4} = P(A) \times P(C)$

$$P(B \cap C) = \frac{1}{4} = P(B) \times P(C)$$

$$P(A \cap B \cap C) = \frac{1}{8} = P(A) \times P(B) \times P(C)$$

CHAPTER 1 – PROBABILITY

Definition 2.6. Independence of more than 2 events

Completelly independent or Mutually independent events

Events A, B e C belonging to the same sample space are **Completelly independent events** if and only if:

$$P(A \cap B) = P(A).P(B), P(A \cap C) = P(A).P(C), P(B \cap C) = P(B).P(C)$$

and

$$P(A \cap B \cap C) = P(A) \times P(B) \times P(C)$$

More generally, $E_1, E_2, \cdots, E_k$ are **mutually statistically independent** if and only if

$$P(E_1, E_2, \cdots, E_k) = P(E_1) \times P(E_2) \times \cdots \times P(E_k) \quad \text{and}$$

$$P(E_i) \times P\left(E_j\right) = 0 \quad \mathrm{i} \neq j \quad (i, j = 1, 2, \cdots, k)$$

CHAPTER 1 – PROBABILITY

Sample Space Partition

The class of events $\{B_1, B_2, \cdots, B_k, \cdots\}$ is said to be a sample space partition if and only if

$$\cup_j B_j = \mathbb{S} \text{ and } B_i \cap B_j = \emptyset \ (i \neq j), i, j = 1,2, \cdots, k, \cdots$$

then $P(\cup_j B_j) = \sum_j P(B_j) = 1, j = 1,2, \cdots, k, \cdots$

Theorem 2.12 Total probability theorem

If $\{B_1, B_2, \cdots, B_k \cdots\}$ is a partition of sample space $\mathbb{S}$ and

$P(B_j) > 0 \ (j = 1, 2, \cdots, k, \cdots)$, for any event A,

$$P(A) = \sum_j P(A \cap B_j) = \sum_j P(B_j) \times P(A|B_j)$$

CHAPTER 1 – PROBABILITY

- Bowl 1 contains 2 red and 4 white chips, Bowl 2 contains 1 red and 2 white chips, Bowl 3 contains 5 red and 4 white chips. The probabilities for selecting the bowls are given by $P(B_1) = \frac{1}{3}$, $P(B_2) = \frac{1}{6}$ and $P(B_3) = \frac{1}{2}$.

The class of events $\{B_1, B_2, B_3\}$ is a partition of the sample space if and only if

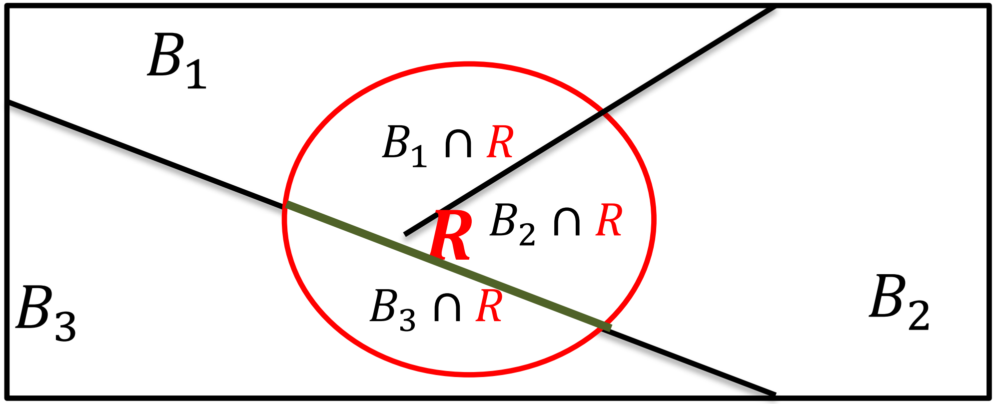
$$B_i \cap B_j = \phi \; i, j = 1, 2, 3$$

and

$$\cup_{j=1}^{3} B_j = \mathcal{S} \Leftrightarrow P(B_1 \cup B_2 \cup B_3) = 1$$

- Consider the experiment selecting a bowl and then drawing a chip at random from the bowl. The probability of drawing a red chip -$P(R)$ is

$$P(R) = P(B_1 \cap R) + P(B_2 \cap R) + P(B_3 \cap R)$$



$$P(R) = P(B_1)P(R|B_1) + P(B_2)P(R|B_2) + P(B_3)P(R|B_3)$$

$$P(R) = \frac{1}{3} \times \frac{2}{6} + \frac{1}{6} \times \frac{1}{3} + \frac{1}{2} \times \frac{5}{9} = \frac{4}{9}$$

Is the **total probability of event R**.

CHAPTER 1 – PROBABILITY

- Now suppose that the outcome of the experiment is a red chip but we don't know from which bowl it was drawn. The conditional probability that the chip was drawn from bowl 1 is

$$P(B_1|R) = \frac{P(B_1 \cap R)}{P(R)} = \frac{P(B_1)P(R|B_1)}{P(B_1)P(R|B_1) + P(B_2)P(R|B_2) + P(B_3)P(R|B_3)}$$

$$P(B_1|R) = \frac{\frac{1}{3} \times \frac{2}{6}}{\frac{1}{3} \times \frac{2}{6} + \frac{1}{6} \times \frac{1}{3} + \frac{1}{2} \times \frac{5}{9}} = \frac{1}{4}$$

CHAPTER 1 – PROBABILITY

- **Theorem 2.13 Bayes Theorem**

If $\{B_1, B_2, \cdots, B_k, \cdots\}$ is a partition of $\mathbb{S}$ and if

$P(B_j) > 0 \ (j = 1, 2, \cdots, k, \cdots)$, for any event A

such that $P(A) > 0$,

$$P(B_j|A) = \frac{P(B_j) \times P(A|B_j)}{\sum_j P(B_j) \times P(A|B_j)}, (j = 1, 2, \cdots, k, \cdots)$$

CHAPTER 1 – PROBABILITY

- **Example** – A stock market analyst examined the prospects of the shares of a large number of corporations. It turned out that 25% performed much better than the market average(E_1), 50% about the same (E_2), 25% much worse (E_3) ? Forty percent of the stocks that turned out to do much better than the market average were rated a "good buy" (A), as were 20% of those that did about as well as the market and 10% of those that did much worse. What is the probability that a stock rated a "good buy" performed better than the average.

CHAPTER 1 – PROBABILITY

$P(E_1) = 0{,}25;\ P(A|E_1) = 0{,}4;\ P(E_2) = 0{,}5; P(A|E_2) = 0{,}2$

$P(E_3) = 0{,}25; P(A|E_3) = 0{,}1$

$P(A) = P(E_1)P(A|E_1)+P(E_2)P(A|E_2)+P(E_3)P(A|E_3)$

$= 0{,}25 * 0{,}4 + 0{,}5 * 0{,}2 + 0{,}25 * 0{,}1 = 0{,}225$

$$P(E_1|A) = \frac{P(E_1 \cap A)}{P(A)} = \frac{P(E_1)P(A|E_1)}{P(A)} = \frac{0{,}25 * 0{,}4}{0{,}225} = 0{,}44(4)$$

CHAPTER 1 – PROBABILITY

$P(E_j)$	$P(A\|E_j)$	$P(E_j)P(A\|E_j)$	$P(E_j\|A)$	
0,25	0,4	0,1	0,444(4)	
0,5	0,2	0,1	0,444(4)	
0,25	0,1	0,025	0,111(1)	
1		0,225	1	

CHAPTER 1 – PROBABILITY

Suppose that you are responsible for detecting the source of error when there is a computer system failure. From past experience you know that there are three sources of error: disk drive, computer memory, operating system.

It is known that 50% of errors are due to the disk drive.

From the pattern of the components performance , it is known that the likelihood that a system failure occurs when there is a disk drive error is 60% and the likelihood that a system failure occurs when there is a memory error is 80%.

It is also known that, a failure having been reported , the likelihood that it was due to a memory error is 40% and that the likelihood of a system failure is 60%.

What is the probability that the system fails when there is an operating system error?

CHAPTER 1 – PROBABILITY

F – system failure

A_j - source of error

A_j	$P(A_j)$	$P(F\|A_j)$	$P(A_j)P(F\|A_j)$	$P(A_j\|F)$
UD	0,5	0,6		
M		0,8		0,4
SO		?		
	1		0,6	1