EC381/MN308
Probability and Some Statistics

Yannis Paschalidis

yannisp@bu.edu, http://ionia.bu.edu/

Dept. of Manufacturing Engineering

Dept. of Electrical and Computer Engineering

Center for Information and Systems Engineering

Lecture 2 - Outline

1. Set theory.
2. Three key concepts:

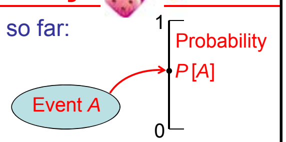
   (Ω, $\mathcal{F}$, P) = (*S*, *E*, *P*)

   = (Sample space, Event space, Probability measure).
3. Probability axioms and some basic properties.

Chapter 1 **Probability**

Summary of what you know so far:

An event *A* is characterized by a number *P* [*A*] between 0 and 1, called the probability of the event.



If *P* [*A*] = 0, the event A never occurs (is impossible).
If *P* [*A*] = 1, the event A always occurs (is certain or sure).
Otherwise, the event A occurs sometimes;
Greater *P* [*A*] means greater likelihood for the event to occur.

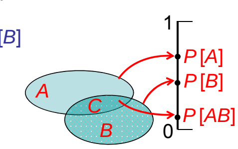
P [*A*] may be determined experimentally by measuring the relative frequency of occurrence, or by adopting some basic principle(s), such as non-prejudice.

Two random events A and B are characterized by the probabilities of the individual events, *P* [*A*] and *P* [*B*], as well as by the probability *P* [*C*] of the joint event *C* = *AB* (i.e., *A* and *B*).

Independence

The event *A* and the event *B* are said to be independent if, and only if, the joint probability is the product of the individual probabilities:

P [*AB*] = *P* [*A*] *P* [*B*]



A formal *axiomatic theory* is necessary to deal with more complex issues such as chaining of events and derived events.

APPROACH:

Definitions: Allow a formulation

Axioms: Accepted without proof

Theorems (and Propositions, Lemmas, etc.): Follow from definitions and axioms

UNDERLYING MATHEMATICAL BASIS:

The mathematical basis of probability theory is *set theory*, which we examine next in simplified form.

1.1 Very Brief Review of Set Theory

A. Definitions

Set = a collection of *elements*

Collection can be: - finite (*B* = {1,2,3}),
- countably infinite (*B* = {1, 2, …})
- or uncountable ($B = \{x : x \in [0,1]\}$)

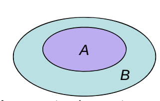
$x \in B$ means that "*x* is an element of the set *B*"

Subset *A* of set *B* = a collection of some of the elements in *B*

$A \subset B$ (symbol $\subset$ indicates "*A* is a subset of *B*")

Equality:

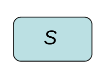
$A = B$ iff $A \subset B$ and $B \subset A$



Venn Diagram

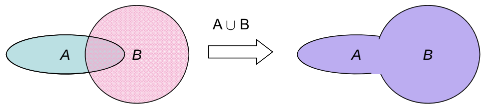
Null Set $\varnothing$ = *A* set with no elements (empty set)

Universal set *S* = the set of all possible elements

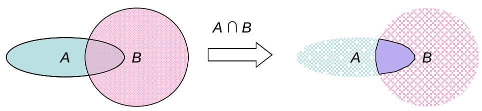


B. Set Operations

Set *union* A ∪ B = {x: x in *A* or x in *B*}

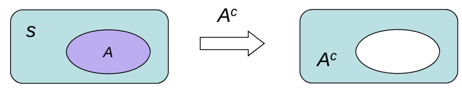


Set *intersection* *A* ∩ *B* = {x: x in *A* and x in *B*}



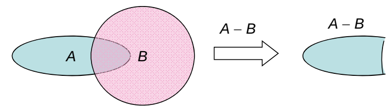
Notations: Sometimes A ∪ B is written as *A*+*B* and *A*∩*B* is written as *AB*

Set *complement* A^c = {x in *S* and x not in *A*}



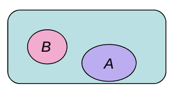
Set *difference* *A* – *B* = {x in *A* and x not in *B*}

($A - B = A \cap B^c$)

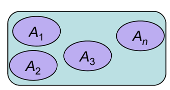


C. Other Set Concepts

A and *B* are disjoint (mutually exclusive) sets if and only if $A \cap B = \varnothing$

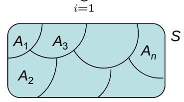


Finite collection $A_1, \ldots, A_n$ is mutually exclusive if and only if $A_i \cap A_j = \varnothing$ for any $i \neq j$ in {1, …, *n*}



Finite collection $A_1, \ldots, A_n$ is collectively exhaustive in *S* if and only if (⇔) $A_1 \cup A_2 \cup \ldots \cup A_n \equiv \bigcup_{i=1}^{n} A_i = S$

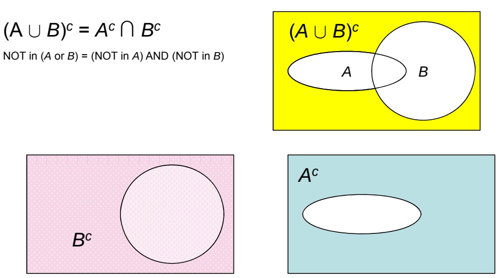
Extensible to *countable* collections



D. De Morgan's Theorems

$(A \cup B)^c = A^c \cap B^c$

NOT in (*A* or *B*) = (NOT in *A*) AND (NOT in *B*)



- Also $(A \cap B)^c = A^c \cup B^c$

end of brief review of set theory

1.2 Axiomatic Theory of Probability

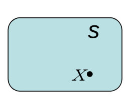
A. Concepts of Experiment, Outcome, Sample Space & Event

Experiment = a procedure that generates observable outcomes. Each procedure generates a single outcome.

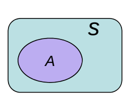
Outcome = any possible observation of an experiment.

Sample Space S = the finest-grain, mutually exclusive, collectively exhaustive set of all possible outcomes.

Each outcome x is a sample point (element or atom) in *S*.



Event = set of all outcomes (sample points) satisfying some property that characterizes that event. The event *A* is a subset of *S*.



Each event is a subset of *S*. A single element (outcome) can be an event.

Example 1

Experiment: Roll a normal six-sided die once

Outcomes: Each outcome is a number *i* = 1, … , 6

Sample Space: 6 distinct numbers: *S* = {1,2,3,4,5,6}

Examples of Events: E_1 = set of all odd outcomes

E_2 = set of all outcomes greater than 2

E_3 = set of all outcomes that are the square of an integer

NOTES: Experiment is sometimes called Procedure

Outcomes are sometimes called Observations

Sample Space is based on a Model (in this example, the model is that each of the six die faces is equally likely to come up and that the result of each experiment is unrelated to that of previous experiments.)

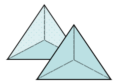
Example 2

Experiment: 2 rolls of a quadrilateral (four-sided) die, record both numbers.

Outcomes: Pairs of numbers {1,2,3,4} x {1,2,3,4}

Sample space: 16 distinct pairs if order matters;

or 10 distinct pairs if order does not matter



Examples of events: E_1 Set of all outcomes with a sum equal to 4

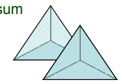
E_2 Set of all outcomes with an odd sum

Example 3

Experiment: 2 rolls of a quadrilateral (four-sided) die; record the sum

Outcomes: Sum of the two numbers, a number between 2 and 8

Sample space: {2,3,4,5,6,7,8}



Examples of events: E_1 Set of all even numbers {2, 4, 6, 8}

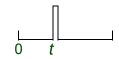
E_2 Set of numbers > 5 {6, 7, 8}

Example 4

Experiment: Time of arrival of the edge of a pulse within a time window $[0,T]$.

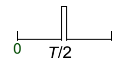
Outcomes: A continuous-valued number $\{t\} \in [0,T]$

Sample Space: An infinite number of all possible values within $[0,T]$

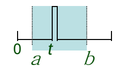


Examples of Events:

E_1: $\{(T/2)\}$



E_2: $\{(t): a<t<b\}$



Example 5

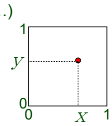
Experiment: Two people choose numbers in [0,1].

Or, pick a point in a unit square [0,1] x [0,1] (toss a dart at it…)

Or, location of an electron or a photon within a square area

Outcomes: Ordered pairs of continuous-valued numbers

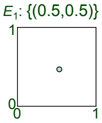
$\{(x, y)\} \in [0,1] \times [0,1]$



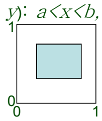
Sample space: An infinite number of pairs of continuous-valued numbers $\{(x, y)\} \in [0,1] \times [0,1]$

Examples of events:

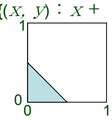
E_1: {(0.5,0.5)}



E_2: $\{(x, y): a<x<b,\ c<y<d\}$



E_3: $\{(x, y) : x + y < 0.5\}$



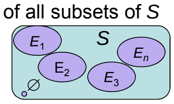
B. Set Theory Description of Probability

Probability Theory		Set Theory
Sample Space	⟺	Universal Set
Outcome	⟺	Element
Event	⟺	Set

Table 1.1 (p. 9). Terminology of set theory and probability.

Event Space E = set of events = set of all subsets of S

Members of a sample space are outcomes but members of an event space are events.



Example: Flip 2 coins, a penny and a dime. $S = \{hh, ht, th, tt\}$, get 4 outcomes. Let E_i = {outcomes with i heads}. Each E_i is an event containing one or more outcomes, e.g., $E_0 = \{tt\}$ contains 1 outcome; $E_1 = \{ht, th\}$ contains 2. Event space $E = \{E_0, E_1, E_2\}$. The event space is not a sample space since it lacks a fine-grain property: learning that an experiment produces event E_1 tells us that one coin came up heads, but doesn't tell us which one.

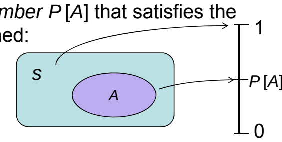
An event space must satisfy several conditions. At a minimum:

- If E_1, E_2 are in E, then $E_1 \cap E_2$ and $E_1 \cup E_2$ are also in E (wish to be able to discuss probability of E_1 and E_2, E_1 or E_2)
- If E is an event, then E^c is an event (probability of not E)
- The sample space is an event: $S \in E$
- The empty set is an event: $\varnothing \in E$
- *Countable* unions, intersections of events are also events:

If $E_1, E_2, \ldots \in E$ then $\bigcup_{i=1}^{\infty} E_i \in E$ and $\bigcap_{i=1}^{\infty} E_i \in E$

C. Probability Axioms

To every event $A \in E$, *a real number* $P[A]$ that satisfies the following three axioms is assigned:



1) $1 \geq P[A] \geq 0$

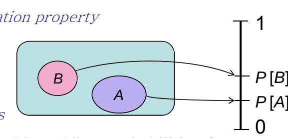
Probability is a nonnegative number

2) $P[S] = 1$

S is the sure event – Normalization property

3) If $A \cap B = \varnothing$, then $P[A \cup B] = P[A] + P[B]$

Additivity property for disjoint events



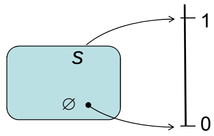
Many probability problems are solved by adding probabilities for mutually exclusive events.

Properties derived from the three axioms

i) $P[\varnothing] = 0$

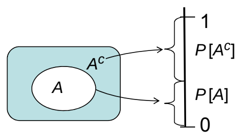
The empty set (null set) is the impossible event

Proof:
Since $\varnothing \cup \varnothing = \varnothing$ and $\varnothing \cap \varnothing = \varnothing$
use axiom # 3 to conclude that
$P(\varnothing) = P(\varnothing) + P(\varnothing)$. Therefore, $P[\varnothing] = 0$.



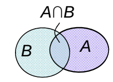
ii) $P[A] + P[A^c] = 1$

Proof:
$A \cup A^c = S$ and $A \cap A^c = \varnothing$
so $P(S) = P(A) + P(A^c)$. Using axiom # 2,
$P(S) = 1$.



EC381/MN308 – 2007/2008

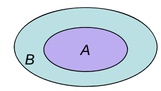
iii) $P[A \cup B] = P[A] + P[B] - P[A \cap B]$



Additivity property

Proof:
CAST A and B AS DISJOINT SETS SO AXIOM #3 CAN BE USED!
$A = (A - B) + A \cap B$; so $P[A] = P[A - B] + P[A \cap B]$
$A \cup B = (A - B) \cup B$; so $P[A \cup B] = P[A - B] + P[B]$
$= P[A] + P[B] - P[A \cap B]$

iv) If $A \subset B$, then $P[B - A] = P[B] - P[A]$
and $P[A] \le P[B]$



Monotonicity property

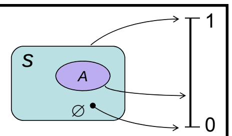
Proof :
CAST A and B AS DISJOINT SETS SO AXIOM #3 CAN BE USED!
$B = (B - A) + A$ and $(B - A) \cap A = \varnothing$
so $P[B] = P[B - A] + P[A]$

EC381/MN308 – 2007/2008

v) $P[A] \le 1$

Proof:

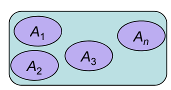
Since $A \subset S$, then $P[A] \le P[S] = 1$



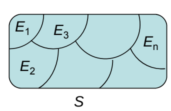
vi) If $A_1, \ldots, A_n$ are mutually exclusive, then

$P[A_1 \cup \ldots \cup A_n] = P[A_1] + \ldots + P[A_n]$

Additivity property
Proof: Simple extension of Axiom #3



vii) If $E_1, \ldots, E_n$ are mutually exclusive and $E_1 \cup E_2 \ldots \cup E_n = S$, then $P[E_1 \cup \ldots \cup E_n] = 1$

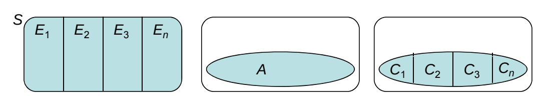


EC381/MN308 – 2007/2008

viii) For any countable set of mutually exclusive events E_i that are collectively exhaustive ($\cup_i E_i = S$), and any other event A,

$$P[A] = \sum_{i=1}^{\infty} P[A \cap E_i]$$

Partition property: Theorem 1.8, p. 15
(Often used when sample space can
Be written in form of a table.)



Proof:
$E = \{E_1, E_2, E_3, E_4\}$ and $C_i = A \cap E_i$ for $i = 1, \ldots, n$.
$A = C_1 \cup C_2 \cup C_3 \cup C_n$,
Since C_i are mutually exclusive, $P[A] = P[C_1] + P[C_2] + P[C_3] + P[C_n]$
The property applies to a countable set $n = \infty$.

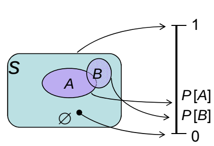
This theorem is important because it allows us to express any event as a union of mutually exclusive events, the cornerstone of probability calculations.

EC381/MN308 – 2007/2008

Summary of Main Points

The probability measure P is a mapping from the event space E into the space of real numbers [0,1]

$P: E \rightarrow [0,1]$



such that

$P[A \cup B] = P[A] + P[B] - P[A \cap B]$

EC381/MN308 – 2007/2008

Probability Axioms for Countable Events

Probability measure P has the following properties:

1) $P[A] \ge 0$ for $A \in E$
2) $P[S] = 1$
3) If $A_1, \ldots, A_n, \ldots$ is a *countable* list of mutually exclusive events, then

$$P[A_1 \cup A_2 \cup \ldots] = P\left[\bigcup_{i=1}^{\infty} A_1\right] = \sum_{i=1}^{\infty} P[A_i]$$

Countable additivity property

EC381/MN308 – 2007/2008

Probability Space

= Axiomatic abstraction of a random experiment consisting of:

1) A sample space S of all possible outcomes (mutually exclusive, collectively exhaustive, finest grain)
2) An event space E of events, which are subsets of S satisfying the axioms of event spaces
3) A probability measure $P : E \rightarrow [0,1]$ satisfying the axioms of probability measures

Steps for Problem Solving

- Set up sample space from description of experiments (all outcomes)
- Describe probability law on events (atoms if finite)
- Identify event of interest
- Calculate…

Example 1

Experiment: one roll of a six-sided die

Outcomes: each outcome is a number $i = 1, \ldots, 6$

Sample Space: 6 distinct numbers: $S = \{1,2,3,4,5,6\}$

Examples of events: E_1: Set of all odd outcomes

E_2: Set of all outcomes greater than 2

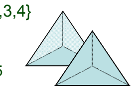
Probabilities: Use principle of nonprejudice to assign to each event that contains a single outcome an equal probability. Since {1,2,3,4,5,6} are mutually exclusive and their union is S, the probability of each must be 1/6. Thus,

$P[E_1] = P[1] + P[3] + P[5] = 3/6 = 1/2$

$P[E_2] = P[3] + P[4] + P[5] + P[6] = 4/6 = 2/3$

Example 2

Experiment: 2 rolls of a quadrilateral (four-sided) die, record both numbers and their order.



Outcomes: Pairs of ordered numbers $(x,y) \in \{1,2,3,4\} \times \{1,2,3,4\}$

Sample space: 16 distinct pairs

Examples of events: E_1: Set of all outcomes with a sum equal to 5

E_2: Set of all outcomes with an odd sum

E_3: $\{x = 1\} = \{(1,1), (1,2), (1,3), (1,4)\}$

E_4: $\{\min(x,y) = 2\} = \{(2,2), (2,3), (2,4), (3,2), (4,2)\}$

Probabilities: Assign to each event that contains a single outcome a probability of 1/16 and use the probability axioms

$P[E_1] = P[\{(1,4)\}] + P[\{(2,3)\}] + P[\{(3,2)\}] + P[\{(4,1)\}] = 4/16 = 1/4$

$P[E_2] = P[\{(1,2)\}] + P[\{(1,4)\}] + P[\{(2,1)\}] + P[\{(2,3)\}] + P[\{(3,2)\}] + P[\{(3,4)\}] + P[\{(4,1)\}] + P[\{(4,3)\}] = 8/16 = 1/2$

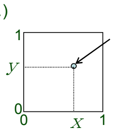
$P[E_3] = P[\{(1,1)\}] + P[\{(1,2)\}] + P[\{(1,3)\}] + P[\{(1,4)\}] = 1/4$

$P[E_4] = P[\{(2,2)\}] + P[\{(2,3)\}] + P[\{(2,4)\}] + P[\{(3,2)\}] + P[\{(4,2)\}] = 5/16$

Example 3 (continuous sample space)

Experiment: Two people choose numbers in [0,1].

Or, pick a point in a unit square [0,1] x [0,1] (toss a dart at it…)

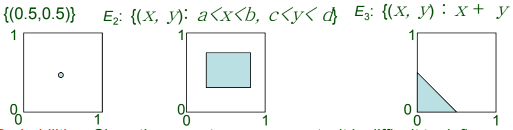


Outcomes: Ordered pairs of continuous-valued numbers $\{(x, y)\} \in [0,1] \times [0,1]$

Sample space: An infinite number of pairs of continuous-valued numbers $\{(x, y)\} \in [0,1] \times [0,1]$

Examples of Events:

E_1: {(0.5,0.5)} E_2: $\{(x, y): a<x<b, c<y<d\}$ E_3: $\{(x, y): x + y < 0.5\}$



Probabilities: Since there are too many events, it is difficult to define a probability measure consistent with axioms. Possible probability measure on interval events is $P[E]$ = area of E

Issues with infinite sample spaces

If S has a finite number of elements, axioms are sufficient:

- Event space E: set of all subsets of S also has a finite number of elements
- Can find mutually exclusive, collectively exhaustive events $E_1, \ldots, E_N$, each set consisting of a single outcome (*atom* events)
- All other events built as unions of these atom events
- Defining a probability measure on each atom event extends to a probability measure on all events consistently

If S has an infinite number of elements:

- Often, probability measure cannot then be defined in terms of atoms
- There may not exist a countable number of atoms
  - Consider $S = [0,1]$. Atoms are $\{x\}$, where $x \in [0,1]$
    - Note: there are an uncountable number of mutually disjoint atoms
    - If all atoms are equally likely, what is $P[\{0.5\}]$?
    - If $P[S] = 1$, how do we define $P[\{x\}]$?

Example: Telephone Calls

Quiz 1.4, p. 16: long (L) or brief (B), voice (V) or data (D) calls
Use Theorem 1.8 (property viii above)

	V	D
L	0.35	0.25
B	0.35	0.05

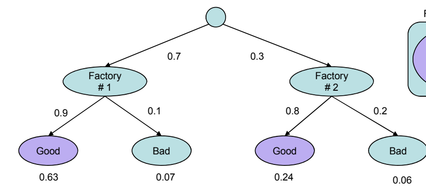
GIVEN P [VL] = 0.35
P[V] = 0.7
P[L] = 0.6

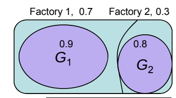
S = {LV, LD, BV, BD}.
Events VL = {LV}, V = {LV, BV}, L = {LV, LD}.
Know P [VL] = 0.35 and P [V] = 0.7 = P [LV] + P [BV]
so P [BV] = 0.35;
Know P [L] = 0.6 and P [L] = P [LV] + P [LD] so P [LD] = 0.25;
Know P [S] = 1 = P [LV] + P [LD] + P [BV] + P [BD]
so P [BD] = 0.05.
We now know P [E] for all atoms, and can thus compute everything:
$P[D \cup L] = P[D] + P[L] - P[D \cap L] = 0.3 + 0.6 – 0.25 = 0.65$
$P[V \cup L] = P[V] + P[L] - P[V \cap L] = 0.7 + 0.6 – 0.35 = 0.95$
$P[V \cup D] = P[V] + P[D] - P[V \cap D] = 0.7 + 0.3 = 1.0$; $P[L \cap B] = 0$

Example: Parts from two factories

Two factories make "identical" parts. However, the
the probability that the part from factory F_1 is okay is 0.9, while
the probability that the part from factory F_2 is okay is 0.8;
Further, factory 1 fabricates 70% of the parts sold at the store, while factory 2 fabricates 30%.

Q: What is the probability that the part you purchase is okay?
Define appropriate event spaces, including one that the part is okay (G):





	G	B
F_1	.63	.07
F_2	.24	.06

One event space is {G B} and another event space is {$F1$ $F2$}. Construct a table as above, to find that $P(G) = P(G F_1) + P(G F_2) = 0.63 + 0.24 = 0.87$. This problem will later be reconsidered from the perspective of conditional probability.