Lecture 1

The Mathematical Theory of Probability

1. Introduction

Today we will do §2.1 and 2.2. We will skip Chapter 1.

We all have an intuitive notion of probability. Let's see.

What is the probability of tossing two heads in a row with a fair coin?

Method 1

List all possible outcomes

$\{HH, HT, TH, TT\}$

so $p = ?$

Question

What did we just assume to arrive at that answer?

Another way

$$\underset{\text{1st toss}}{\underline{H}} \quad \underset{\text{2nd toss}}{\underline{H}}$$

However it is important to put probolity into a formal mathematic framework for many reasons.

1. Even "elementary" problems become too hard unless we can break them down into simpler problems using the rules of <u>set theory</u>.

<u>Examples</u>

Let's see how you can deal with these now and later

(there is another reason which we will run into later — we often have infinite sets and need calculus e.g. financial math.)

Problems

1. What is the probability of getting <u>one</u> head in one hundred tosses of a fair coin?

2. What is the probability of getting 27 heads in one hundred tosses of a fair coin?

Prediction – nobody will get this one now.
In two weeks everybody will.

2. Transition from the naive theory to the formal mathematical theory

To make the transition we introduce the word "experiment" which will be taken to mean "any action or process whose outcome is subject to uncertainty"

text - pg 47.

Examples

- Tossing a fair coin 100 times.
- Dealing 5 cards from a 52 card deck — a poker hand.

Dealing 13 cards from a 52 card deck — a bridge hand

Definition

The set of all possible outcomes of on experiment will be called the sample space of that experiment and denote S.

Experiment 3 tosses of a fair coin

$$S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

Definition

A subset A of $\mathcal{S}$ is called an event.

Problem

Find $P\left(\begin{array}{l}\text{at least one head in} \\ \text{3 tosses of a fair coin}\end{array}\right)$

We are looking for $P(A)$ where A is a subset of the previous $\mathcal{S}$.

$$S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

We will call ~~this~~ "our favorite sample space" from now on.

3. The Formal Mathematical Theory

Let $\mathcal{S}$ be a set (the sample space). A probability measure P on $\mathcal{S}$ is a rule (function) which assigns a real number $P(A)$ to any subset A of $\mathcal{S}$ (ie to any event) such that the following axioms are satisfied

1. For any event $A \subset \mathcal{S}$ we have $P(A) \geq 0$

2. $P(\mathcal{S}) = 1$

3. If $A_1, A_2, \ldots, A_n, \ldots$ is a possibly infinite collection of pairwise disjoint (mutually exclusive) events then

$$P(A_1 \cup A_2 \cup \cdots \cup A_n \cup \cdots) = \underbrace{\sum_{n=1}^{\infty} P(A_n)}_{}$$

sum of an infinite series not just ordinary sum

mutually exclusive means $A_i \cap A_j = \phi$ for any pair i, j with $i \neq j$.

Special cases

1. Two mutually-exclusive events A_1 and A_2 (so $A_1 \cap A_2 = \phi$)

$$P(A_1 \cup A_2) = P(A_1) + P(A_2)$$

2. n mutually-exclusive events $A_1, A_2, \ldots, A_n$

$$P(A_1 \cup A_2 \cup \cdots \cup A_n)$$

$$= P(A_1) + P(A_2) + \cdots + P(A_n)$$

A Class of Examples

Let $\mathcal{S}$ be a set with n elements. Let $A \subset \mathcal{S}$ be any subset. Define

$$P(A) = \frac{\#(A)}{\#(\mathcal{S})} = \frac{\#(A)}{n}$$

Then P satisfies the axioms 1., 2. and 3.

Here $\#(A)$ means the number elements in A.
This is called the "equally likely probability measure".

An example in the above class

Take our favorite sample space

$$\mathcal{S} = \left\{ \begin{array}{l} HHH, HHT, HTH, HTT \\ THH, THT, TTH, TTT \end{array} \right\}$$

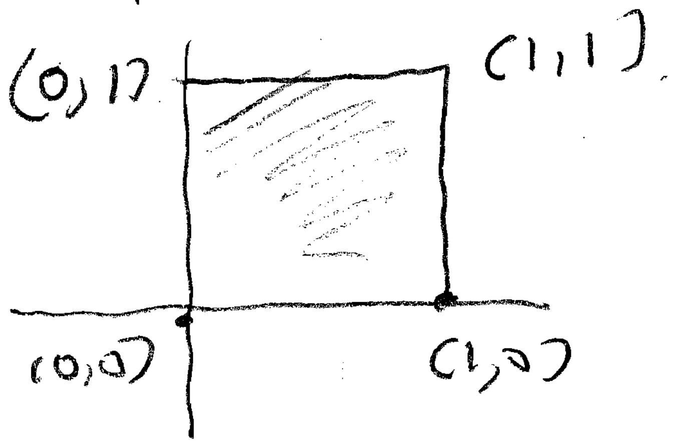
Let A be the subset (event) of outcomes with at least one head and one tail.

All the outcomes are equally likely (because the coin is fair) so

$$P(A) = \frac{\#(A)}{\#(\mathcal{S})} = \frac{?}{8}$$

A Continuous Example

Consider the unit square $\mathcal{S}$ in the plane



Let $A \subset \mathcal{S}$ be any subset.

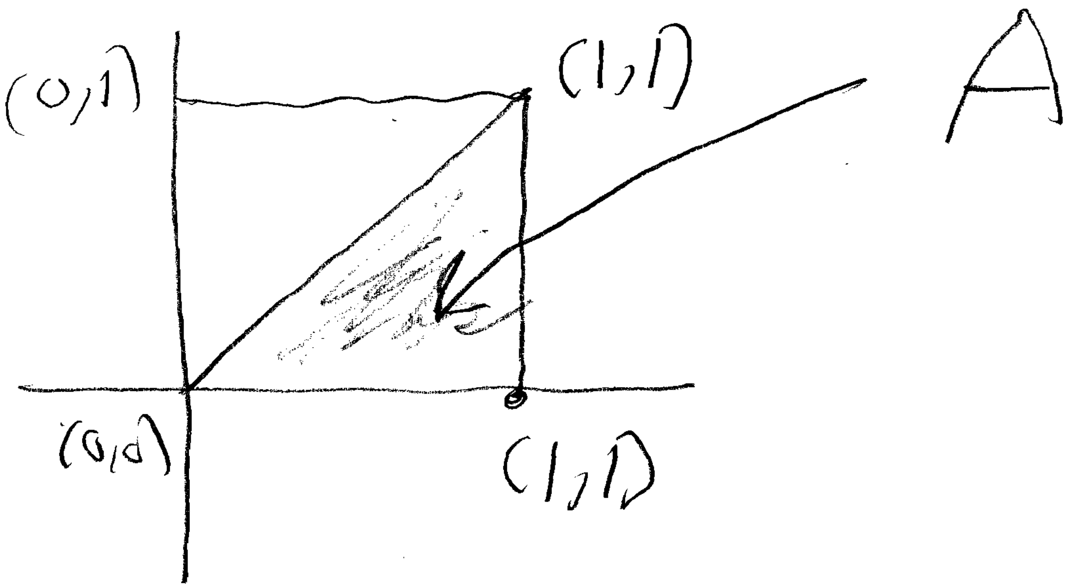
Define

$$P(A) = \text{Area of } A$$

Then P satisfies the axioms 1., 2. and 3.

Let A be the subset of points in the square below the diagonal.

What is $p(A)$?



Can you find A so that

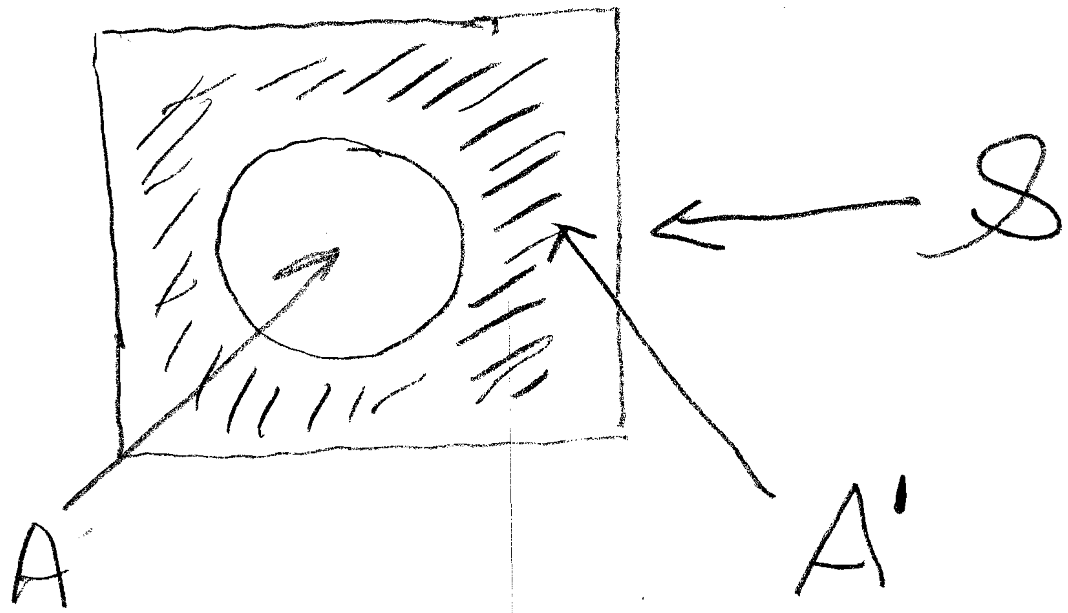
$$p(A) = \frac{1}{\pi} ?$$

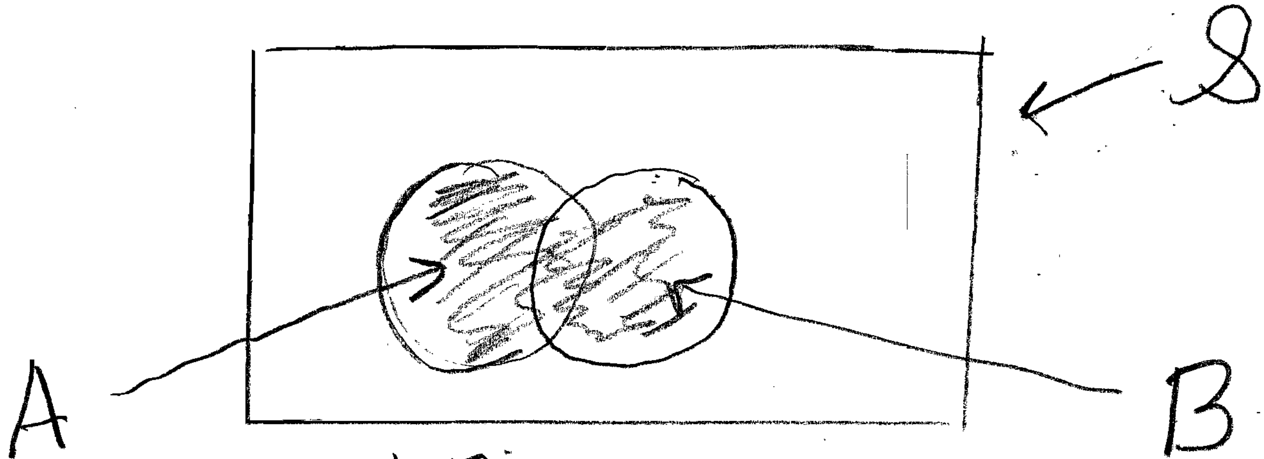
4 A Quick Trip Through Set-Theory (pg 49-50)

Let S be a set and A and B be subsets. Then we have $A \cup B$ (union), $A \cap B$ (intersection) and A' (complement).

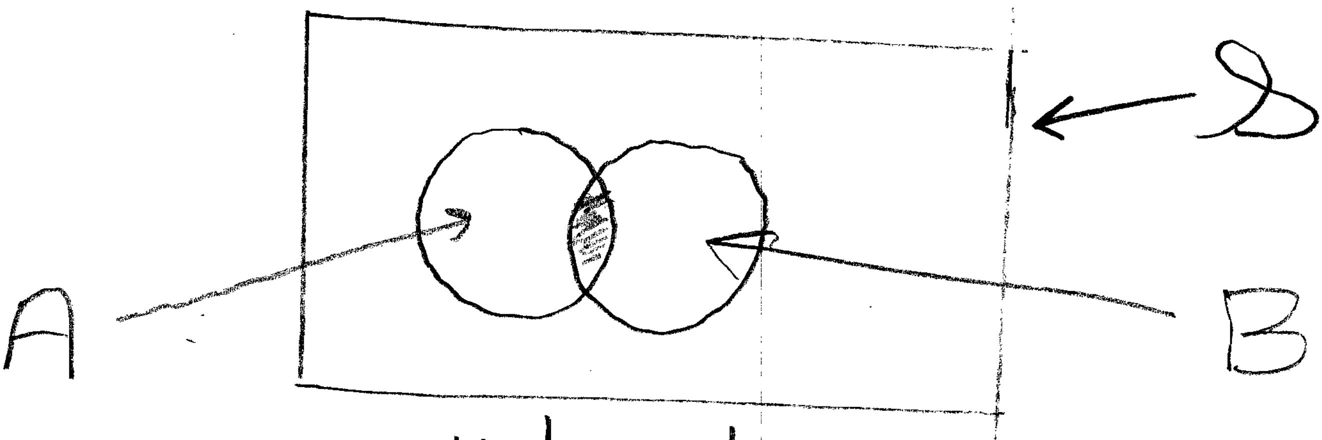
Venn diagrams

complement





union



intersection

$A \cup B$ = "everything in S that is in either A <u>or</u> B"

$A \cap B$ = "everything in S that is in A <u>and</u> B"

The formulas linking $\cup$, $\cap$ and $'$

To help you remember the formulas that follow use the analogy

$$\mathcal{S} \longleftrightarrow \text{set of numbers}$$
$$\cup \longleftrightarrow +$$
$$\cap \longleftrightarrow \cdot$$

The commutative laws

$$A \cup B = B \cup A \quad \left(\begin{array}{c}\text{analogue}\\ a+b=b+a\end{array}\right)$$

$$A \cap B = B \cap A \quad \left(\begin{array}{c}\text{analogue}\\ a\cdot b=b\cdot a\end{array}\right)$$

The associative laws

$(A \cup B) \cup C = A \cup (B \cup C)$ $\left(\begin{array}{c}\text{analogue}\\ (a+b)+c = a+(b+c)\end{array}\right)$

$(A \cap B) \cap C = A \cap (B \cap C)$ $\left(\begin{array}{c}\text{analogue}\\ (a \cdot b) \cdot c = a \cdot (b \cdot c)\end{array}\right)$

Now we have laws that relate two or more of $\cup$, $\cap$ and $'$.

The distributive laws

$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ $\left(\begin{array}{c}\text{analogue}\\ a \cdot (b+c) = (a \cdot b) + (a \cdot c)\end{array}\right)$

$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ no analogue

Problem What would the analogue of the second distributive law say. It isn't true.

De Morgan's Laws

(no analogy with $+, \cdot$)

$$(A \cup B)' = A' \cap B'$$

$$(A \cap B)' = A' \cup B'$$

$$C \subset D \iff C' \supset D'$$

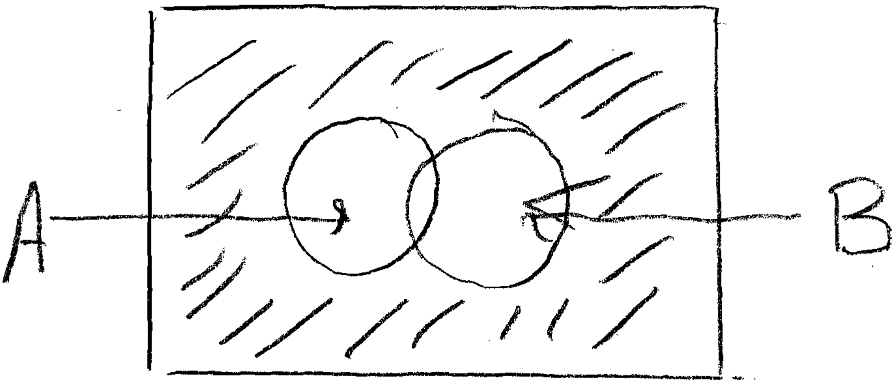
if and only if

(so complement reverses $\cup$, $\cap$ and $\subset$)

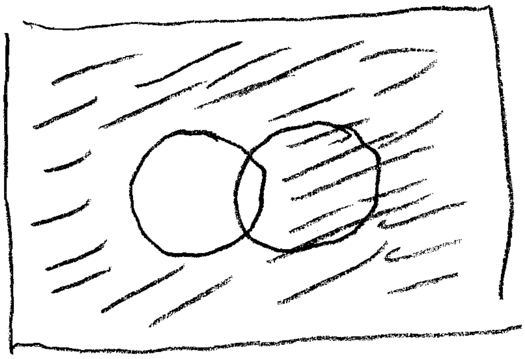
One way to think of the first formula.

not in A or B. = not in A and not in B

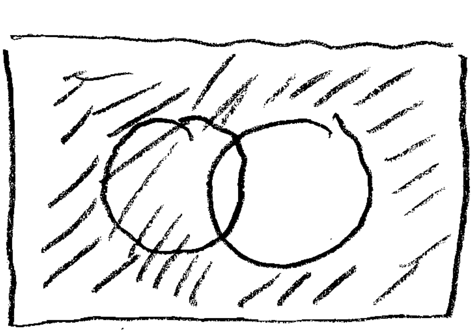
The best way to see it is by a Venn diagram



$(A \cup B)'$ shaded



A' shaded



B' shaded

Top square = intersection of bottom two squares

Consequences of the axioms of probability theory

pg 54 - 56

We will prove two propositions which will be extremely useful to you

Proposition 1 (Complement Law)

$$P(A') = 1 - P(A)$$

Proof

$A \cup A' = S$ so

$P(A \cup A') = P(S) = 1$ (axiom 2) (#)

But $A \cap A' = \phi$ so by

axiom 3, special case 1

$$P(A \cup A') = P(A) + P(A') \quad (\#\#)$$

Putting (#) and (##) together we get

$$1 = P(A) + P(A')$$

qed in TeX $\to \square$

Corollary 1

$$P(\phi) = 0$$

Proof

$$\phi = S'$$

so $P(\phi) = 1 - P(S) = 1 - 1 = 0$

$\square$

Remark

ϕ is not the Greek letter phi
It is a Norwegian letter.
The symbol was chosen by
André Weil.

Corollary 2

$$P(A) \leq 1$$

Proof

$$P(A) = 1 - P(A') \leq 1$$

because $P(A') \geq 0$. $\square$

Bottom line (literally)

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

To illustrate the use of Proposition 1 let us go back to computing

P(at least one head in three tosses)

Put $\mathcal{S}$ = our favorite sample space

A = at least one head SO

A' = no heads = all tails = TTT

so

$$P(A) = 1 - P(TTT) = 1 - \frac{1}{8} = \frac{7}{8}$$

Now we can do 100 tosses

$$P(\text{at least one head}) = 1 - \frac{1}{100} = \frac{99}{100}$$

Recall that two events A and B are mutually exclusive if $A \cap B = \phi$ and axiom 3 says in this case

$$P(A \cup B) = P(A) + P(B) \quad (\#)$$

The following proposition is absolutely critical for computations

<u>Proposition 2</u> (Additive Law)

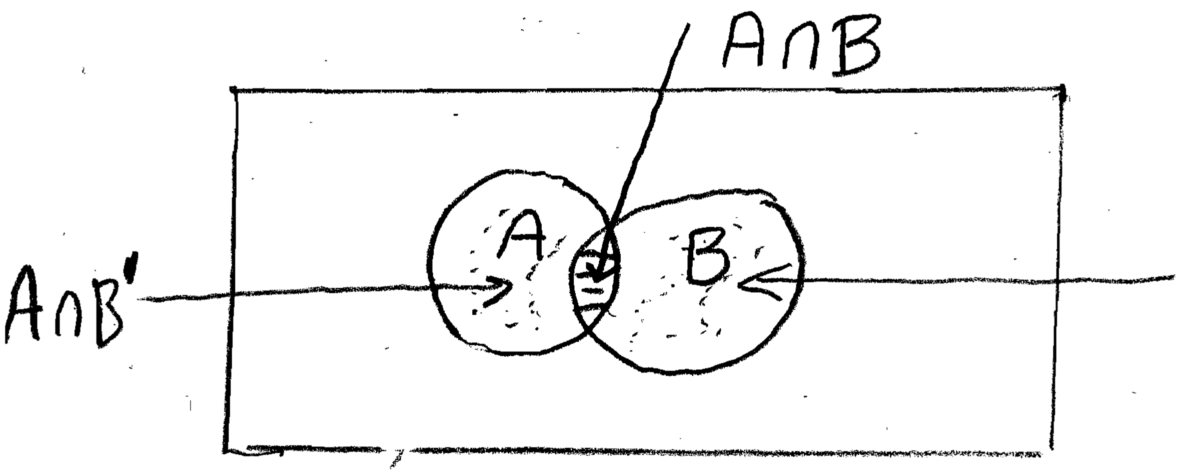
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Note that this is consistent with (#) above because if $A \cap B = \phi$ then

$$P(A \cap B) = P(\phi) = 0$$

<u>Proof</u> The proof is hard.

It depends on the following Venn diagram.



We see that $A \cup B$ is the union of three <u>mutually</u> <u>exclusive</u> sets

$$A \cup B = (A \cap B') \cup (A \cap B) \cup (B \cap A')$$

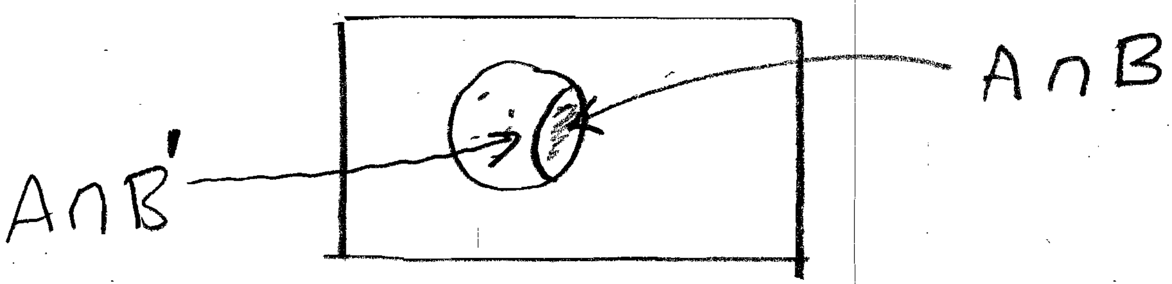
so by axiom 3 with $n = 3$

$$P(A \cup B) = P(A \cap B') + P(A \cap B) + P(B \cap A') \quad (\#\#)$$

How do we compute the first and third terms?

We have a disjoint union (ie union of mutually exclusive sets)

$$A = (A \cap B) \cup (A \cap B')$$



so by axiom 3

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

whence

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

Similarly

$$P(B \cap A') = P(B) - P(A \cap B) \qquad (3)$$

Plug (1) and (3) into (##).

$\square$

What about the intersection of three terms?

Proposition 3

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

This is (more or less) "the principle of exclusion and inclusion"

1. include the singletons A, B, C
2. exclude the pairs $A \cap B, A \cap C, B \cap C$
3 include the triple $A \cap B \cap C$