ACM 116: Lecture 1

Agenda

- Philosophy of the Course

- Definition of probabilities
- Equally likely outcomes
- Elements of combinatorics
- Conditional probabilities

Philosophy of the Course

Probability is the language of uncertainty.

There are many ways of dealing with the subject of probability; e.g. emphasis on

- Measure theory
- Statistics
- Intellectual puzzles

Here, emphasis on the study of probability models. Stochastic models are useful for describing many real phenomena.

This Lecture

- Definition of probabilities
- First examples and calculations

Definition of Probability

First way of defining probabilities is due to Kolmogorov.

- Sample space
- Events
- Probability measures

Sample Space

Sample Space: Set of all possible outcomes of an experiment. The sample space is commonly denoted by Ω; a generic element of Ω is denoted by ω.

- Example A: TCP packets, i.e. finite sequence of bits.
  Each bit is either 0 or 1.
  - e.g. 3 bits.

$$\Omega = \{000, 001, 010, 011, 100, 101, 110, 111\}$$

  - outcome 010
    - first bit is 0
    - second bit is 1
    - third bit is 0

- Example B: Number of jobs in print queue of a mainframe computer.

$$\Omega = \{0, 1, 2, 3, \ldots\}$$

  In practice, upper limit on how large the print queue can be.

$$\Omega = \{0, 1, 2, \ldots, N\}$$

- Example C: Length of time between specific earthquakes in a specific region that are greater in magnitude than a given threshold.

$$\Omega = \mathbf{R}^{+} = \{t,\, t \geq 0\}$$

Events

Event: Subset of the sample space (of possible outcomes).

- Example A: A is the event that the last bit is a 1.

$$A = \{001, 011, 101, 111\}$$

- Example B: A is the event that there are fewer than 5 jobs in the printer queue

$$A = \{0, 1, 2, 3, 4\}$$

- Example C: A is the event that the time interval is greater than one year and less than two years

$$A = \{t,\ 1 < t < 2\}$$

Language of Sets

- Union: $A \cup B$. E.g.
  - A: first bit is a one
  - B: third bit is a one

$$A = \{100, 101, 110, 111\}, \quad B = \{001, 011, 101, 111\}$$

$$A \cup B = \{100, 001, 011, 101, 110, 111\}$$

- Complement: denoted with A^c. E.g., first bit is a zero

$$A^c = \{000, 001, 010, 011\}$$

- Intersection

Probability Measures

A **probability measure** on Ω is a function P from the sets of Ω to the real numbers satisfying the following axioms.

1. $P(\Omega) = 1$.
2. $A \subset \Omega$, $P(A) \geq 0$.
3. A_1 and A_2 disjoint events

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

More generally, $A_1, A_2, \ldots$ mutually disjoint

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

Properties

1. $P(A^c) = 1 - P(A)$.

$$P(A \cup A^c) = P(\Omega) = P(A) + P(A^c)$$

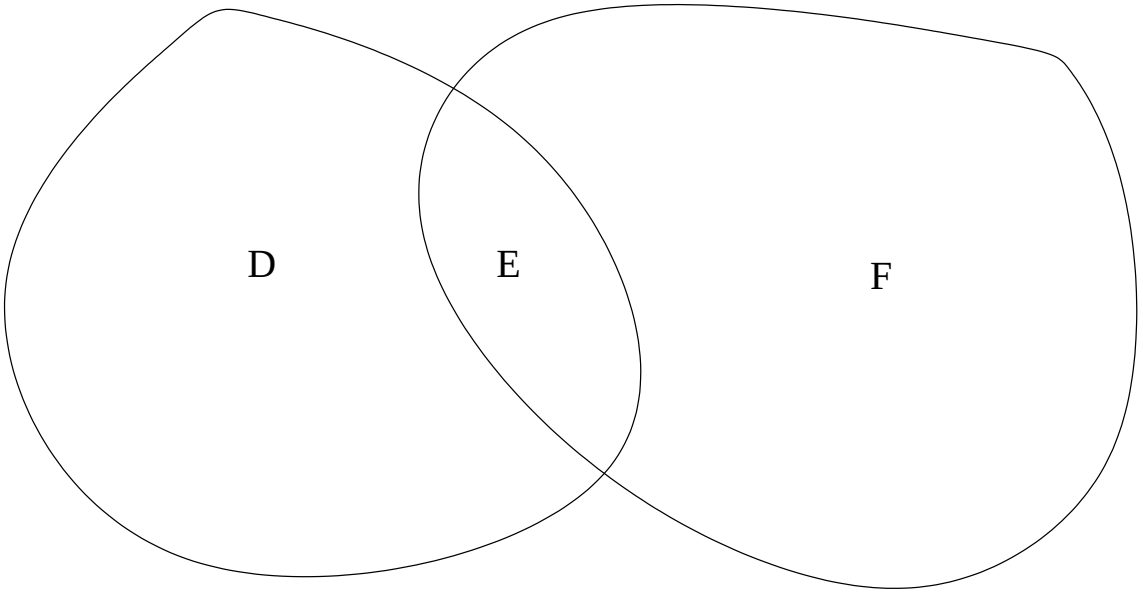
2. $P(\emptyset) = 0$.

3. $A \subset B \quad \Rightarrow \quad P(A) \leq P(B)$.

$$\begin{aligned} B &= A \cup (B \cap A^c) \\ P(B) &= P(A) + P(B \cap A^c) \end{aligned}$$

4. Addition rule

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



$$
\begin{aligned}
P(A \cup B) &= P(D) + P(E) + P(F) \\
P(A) &= P(D) + P(E) \\
P(B) &= P(E) + P(F)
\end{aligned}
$$

which gives

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

Early Models

Setup

- Finite sample space Ω
- All the elements/outcomes of Ω are equally likely

$$P(\{\omega\}) = \frac{1}{\#\Omega}, \quad \omega \in \Omega$$

- Event $A \subset \Omega$

$$P(A) = \frac{\#A}{\#\Omega}$$

- Consequence: probability is related to combinatorics (science of counting)

Things to keep in mind

- Those models are very useful/fruitful
- Developed in connection with the games of chance
- Probability is very different from combinatorics

Interpretation of Probabilities

- Very difficult to try defining probability
- 17th Century:

  Probability of an event = # favorable outcomes/ # possible outcomes

- Interpretation of probabilities is a mess
- The concept of randomness is delicate
  - Coin flipping
  - Computer generated random numbers
  - Sex of an offspring

Which type of mathematics?

- We use probability to do the mathematics of random phenomena.
- What is random? A coin toss?
- Randomness is often a good way to describe a phenomena that is too complex to describe exactly. What we think as random/deterministic has often more to do with the nature of our knowledge than with the essence of the phenomena
- The models that have been used for biological phenomena are mainly stochastic models (differently from physics that got to stochastic modeling only very late)

Pragmatic Approach

Model

- reality
- uncertainty

We want to explore the consequences of those mathematical tools

Examples

- Price of a stock option
- Insurance premium
- How many DSL lines (ISP)?

What is the role of models?

What we see is the solution to a computational problem, our brains compute the most likely causes from the photon absorptions within our eyes. H. Helmholtz

- All models are wrong. Some are useful. (Cox)
- We use models to learn about nature and to mimic nature abilities (vision is one example)

Genetics

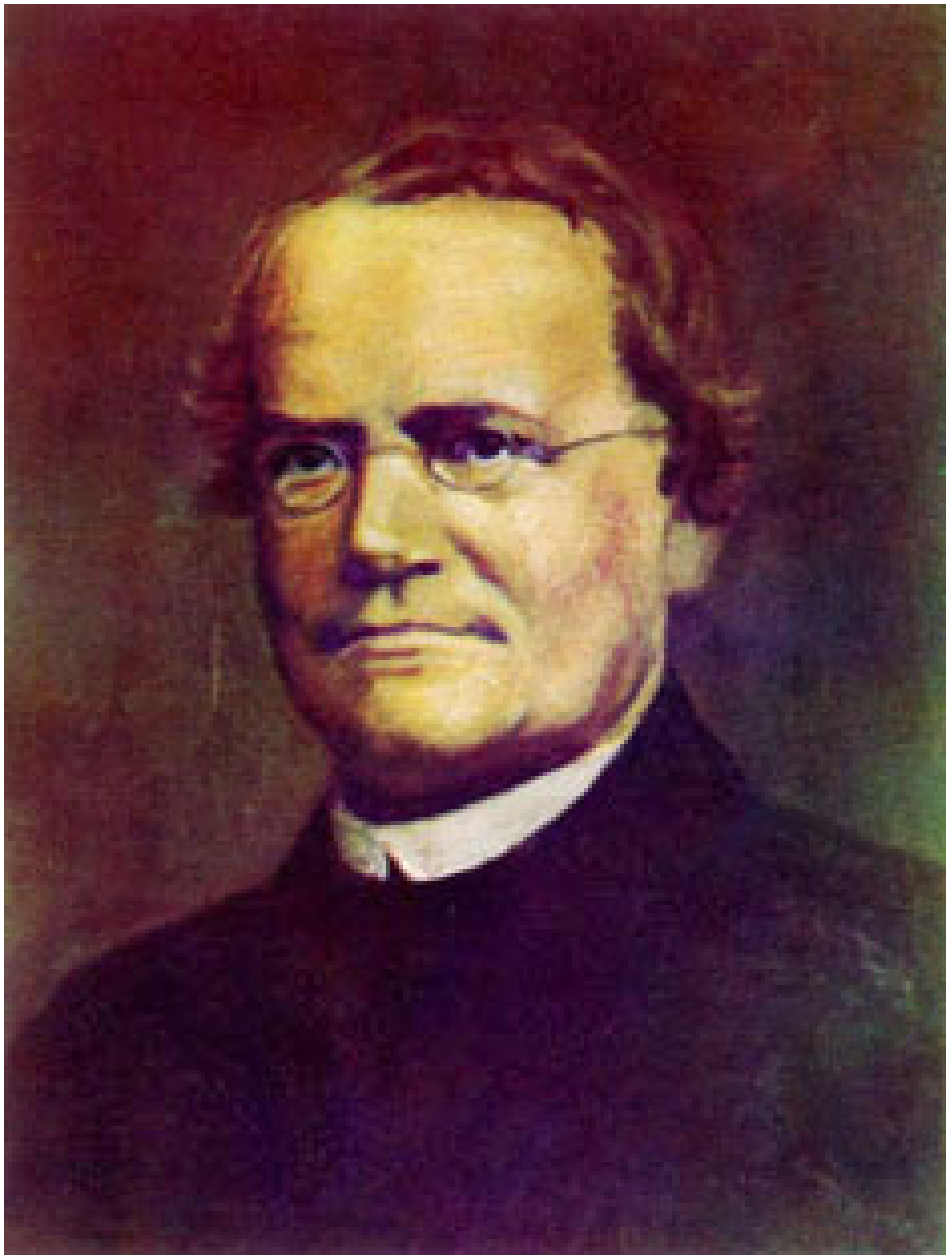
I shall never believe that God plays dice with the world

Albert Einstein

- Randomness: how we can describe reality
- Mendel theory of genetics
  Laws of inheritance of physical traits
  - Mechanism of heredity is based on gene pairs
  - Gene pairs control biological characteristics in several ways. One way is dominance.
- Illustrates the power of simple probability models
- Today's applications: finding causing-disease genes

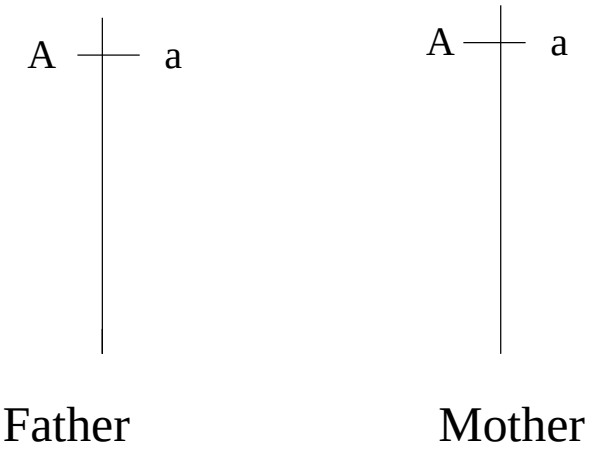
J. G. Mendel

(Austria, 1823-1884)



Laws of Genetic Inheritance

Mendel, second half of the 19th century.



Genotype $\longrightarrow$ Physical trait

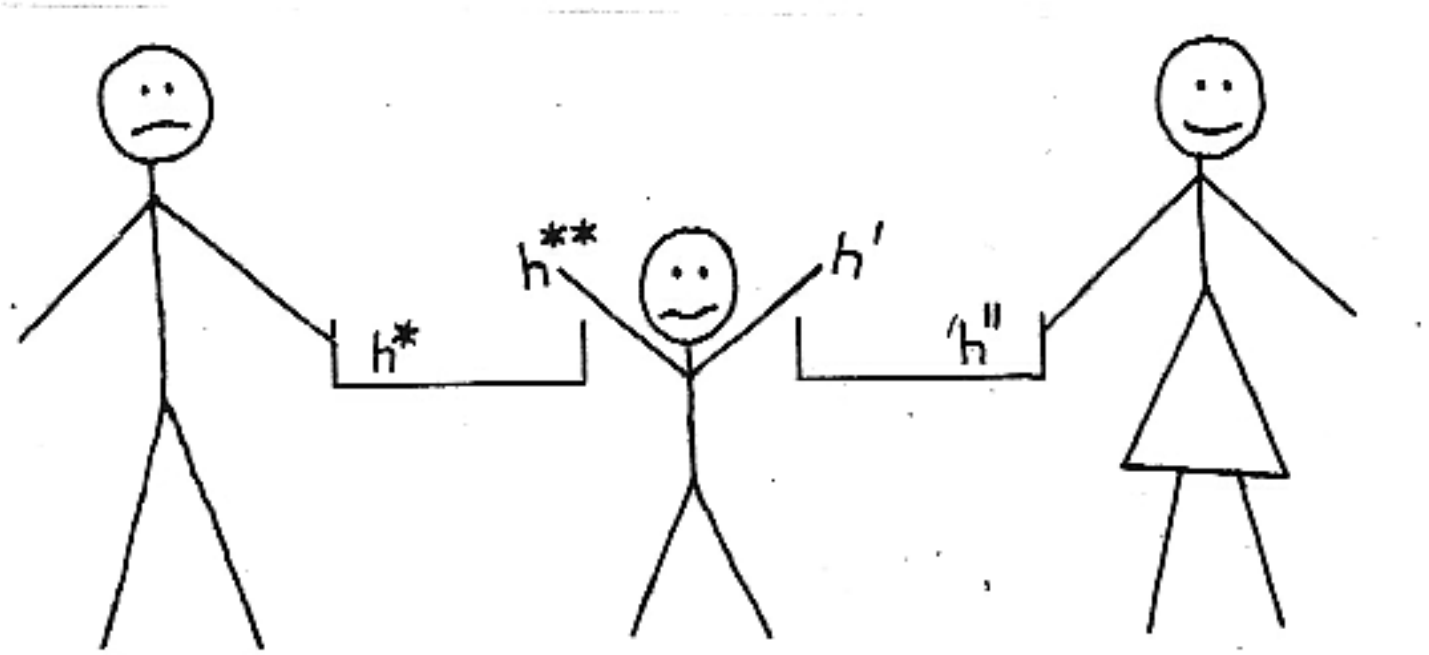
Here Allele with 2 possible values $\boldsymbol{A}$, $\boldsymbol{a}$.

- $\boldsymbol{A}$: Dominant
- $\boldsymbol{a}$: Recessive

e.g. Eye color

- $\boldsymbol{A}$: Black
- $\boldsymbol{a}$: Blue

Example



- d: dominant allele
- r: recessive allele

$$\left.\begin{array}{l} d/d \\ d/r \\ r/d \end{array}\right\} \rightarrow \text{dominant trait}$$

$$r/r \rightarrow \text{recessive trait}$$

Example

- Eye Color

Emmanuel	Chiara
blue/black	blue/black

- Chance of having an offspring with blue eyes

$$p = 1/2 \times 1/2 = 1/4.$$

Equally likely outcomes

- Ω finite sample space
- All the outcomes are equally likely, i.e.

$$P(\omega) = \frac{1}{\#\Omega} \qquad \text{for every } \omega \in \Omega$$

- Probability of an event :

$$P(A) = \frac{\#A}{\#\Omega} \qquad \text{for every } A \subset \Omega$$

  Follows from the third axiom

$\Rightarrow$ The calculation of probabilities involves some combinatorics

An important abstraction

Random placement of r balls in n cells

- *What is the sample space ?*
  e.g. 2 balls in 3 cells :

$\underline{ab}\ \underline{\ }\ \underline{\ }$	$\underline{\ }\ \underline{ab}\ \underline{\ }$	$\underline{\ }\ \underline{\ }\ \underline{ab}$
$\underline{a}\ \underline{b}\ \underline{\ }$	$\underline{a}\ \underline{\ }\ \underline{b}$	$\underline{b}\ \underline{a}\ \underline{\ }$
$\underline{b}\ \underline{\ }\ \underline{a}$	$\underline{\ }\ \underline{a}\ \underline{b}$	$\underline{\ }\ \underline{b}\ \underline{a}$

- *Cardinality of the sample space ?*

$$N = \#\Omega = n^r$$

The multiplication rule

Suppose there are p experiments :

- the first has n_1 outcomes,
- the second has n_2 outcomes, ...
- and the p^{th} has n_p possible outcomes.

Then there are a total of

$$N = n_1 \times n_2 \times \ldots \times n_p$$

possible outcomes for the p experiments.

This abstraction proves to be very useful. Many experiments are equivalent to this scheme of placing r balls in n cells.

Examples

Birthday problem: configuration of birthdays of r people

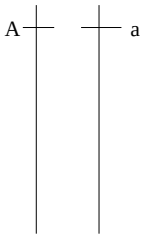
balls	$=$	people
cells	$=$	days of the year

Physics: particles hitting an array of detectors

balls	$=$	particles
cells	$=$	detectors

Examples

Genetics: classification according to the genotype



Genotype (gene pair) = allele pair

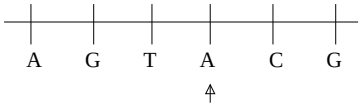
e.g. AA Aa aa

balls = people

cells = genotype

Examples

Bioengineering: DNA mutation



Mutate a letter to another

Screening : record the location of mutations which produce equally or more effective proteins

balls	=	mutations
cells	=	location on the DNA strand

Elements of combinatorics

n objects $\{a_1, a_2, \ldots, a_n\}$

We successively pick r elements and list them in order. In how many ways can we do this ?

- The number of *ordered* samples of r objects from n objects *with replacement* is n^r
- The same, *without replacement* : $n(n-1)\ldots(n-r+1)$

Suppose we are no longer interested in ordered samples but in the constituents of the samples regardless of the order in which they were obtained.

- The number of *unordered* samples of r objects from n objects *without replacement* is $\binom{n}{r} = \frac{n(n-1)\ldots(n-r+1)}{r!} = \frac{n!}{(n-r)!r!}$

Example : quality control

Only a fraction of the output of a manufacturing process is sampled and examined

- n items in a lot
- a sample of size r is taken

Suppose the lot contains k defective items. What is the probability that the sample contains exactly m defective items ?

$$P(A = \text{sample contains } m \text{ defective items}) = \frac{\#A}{\#\Omega}$$

Assumption : each item is equally likely to be sampled

Example : quality control

$$\#\Omega = \binom{n}{r}$$

$$\#A = \binom{k}{m}\binom{n-k}{r-m}$$

$$\Rightarrow P(A) = \frac{\binom{k}{m}\binom{n-k}{r-m}}{\binom{n}{r}}$$

In practice, we observe the number of defective items in the sample and guess the total number of defective items. It's a reverse problem.

Maximum likelihood principle : choose k which maximizes $P(A)$.

$$k = \text{argmax}\frac{\binom{k}{m}\binom{n-k}{r-m}}{\binom{n}{r}} = \text{argmax}\binom{k}{m}\binom{n-k}{r-m}$$

Solution : $k \sim n\frac{m}{r}$

Example : bioengineering - mutation machine

DNA strand : 10 bases and 7 locations

Manufacturer's claim : the locations are equally likely

- Event A : at least 2 mutations occur at the same place

$$P(A) = 1 - P(A^c)$$

$$P(A^c) = \frac{10 \times 9 \times \cdots \times 4}{10^7} = 0.0605$$

- Event B : there is a location with 5 or more mutations

$$(5,1,1) \quad , \quad (5,2) \quad , \quad (6,1) \quad , \quad (7)$$

Example : bioengineering - mutation machine

$$P(5,1,1) = \frac{10\binom{7}{5}\binom{9}{2}2!}{10^7} = \frac{10!}{5!}2!10^{-7} = 0.001512$$

$$P(5,2) = \frac{10\binom{7}{5}9}{10^7} = \frac{10!}{5!}2!\frac{1}{8}10^{-7} = 0.000189$$

$$P(6,1) = \frac{10\binom{7}{6}9}{10^7} = \frac{10!}{6!}\frac{1}{8}10^{-7} = 0.000063$$

$$P(7) = 10 \times 10^{-7} = 10^{-6}$$

$$\Rightarrow P(B) = 0.001765$$

Exercise : you run a little experiment with the machine and discover there is a site with 4 mutations. Do you believe the machine is random ?

Conditional probability

Conditional probability is a very fruitful concept

Situations where we want to compute probabilities when partial information is available

Example :

1. P(location with 5 mutations — there is a location with at least 4)
2. we have a population of individuals

	D^+	D^-	
T^+	$\mathbf{25}$	$\mathbf{14}$	$\mathbf{39}$
T^-	$\mathbf{18}$	$\mathbf{78}$	$\mathbf{96}$
	$\mathbf{43}$	$\mathbf{92}$	$\mathbf{135}$

Pick a person at random

Conditional probability

$$P(D^+) = 43/135$$

$$P(D^+|T^+) = 25/39 = \frac{P(D^+ \cap T^+)}{P(T^+)}$$

Definition : A and B two events with $P(A) \neq 0$

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

conditional probability of B given that A occurred.

Given that B occurred, the relevant sample space becomes B rather than Ω.

$P(A \cap B), P(A)$	Unconditional probabilities
$P(B\|A)$	Conditional probability

Conditional probability

A and B two events

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

Example : A family has 2 children. What is the probability that both are girls given that at least one of them is a girl ?

Sample space $\{b,b\},\{b,g\},\{g,b\},\{g,g\}$

$$P(A = \text{at least one girl}) = 3/4$$

$$P(B = \text{two girls}) = 1/4$$

$$P(B|A) = \frac{1/4}{3/4} = 1/3$$

Law of total probability

$B_1, \ldots, B_n$ mutually exclusive,
$\bigcup_{i=1}^{n} B_i = \Omega$, then for any event A,

$$\underbrace{P(A)}_{\text{marginal prob.}} = \sum_{i=1}^{n} \underbrace{P(A|B_i)}_{\text{conditional prob.}} \underbrace{P(B_i)}_{\text{marginal prob.}}$$

Useful because sometimes

- It is difficult to calculate $P(A)$, while
- It may be relatively easy to compute $P(A|B_i)$, $P(B_i)$