Probabilistic Analysis and Randomized Algorithms

Alexandre David

B2-206

Today

- Counting.
- Basic probability.
- Introduction to randomized algorithms.

Appendix C

Chapter 5

Counting

- Rule of sum.
  - Number of ways of choosing from one of two sets.

$$|A \cup B| = |A| + |B| - |A \cap B|$$

- Rule of product.
  - Number of ways of choosing an ordered pairs from two sets.

$$|A \times B| = |A||B|$$

Counting

- Permutations (on ordered sequences):
  - (A,B,C,D) ≠ (B,A,C,D) ≠ (C,D,A,B)…
  - How many possible permutations?
  - Choose 1st element, choose 2nd … ⇒ $\boxed{n!}$?
- k-permutations:
  - Choose k elements among n: Choose 1st, 2nd, … kth. How many? ⇒ $\boxed{\dfrac{n!}{(n-k)!}}$?

Counting

- k-combination – k subsets:
  - Set {A,B,C,D,E} (n=5), choose 3 among it (k=3): {A,B,C}, {A,B,D}, {B,C,D}…
  - How many combinations?
  - We count a k-permutation and we keep only one representant for every set of equivalent combination, e.g., {A,B,C}={B,A,C}={C,A,B}…

?

$$\boxed{\binom{n}{k} = \frac{n!}{(n-k)!k!}}$$

Counting

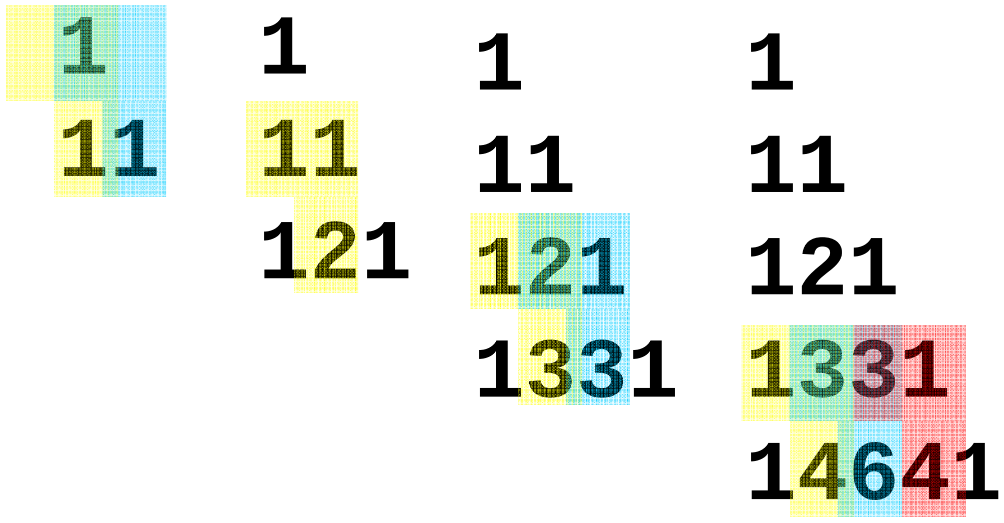
- Binomial coefficients (*n choose k*)

$$\binom{n}{k} = \binom{n}{n-k} = \binom{n-1}{k} + \binom{n-1}{k-1}$$

symmetry

recursion
useful for Pascal's triangle

Pascal’s Triangle



```
1       1       1       1
11      11      11      11
        121     121     121
                1331    1331
                        14641
```

Binomial Expansion

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

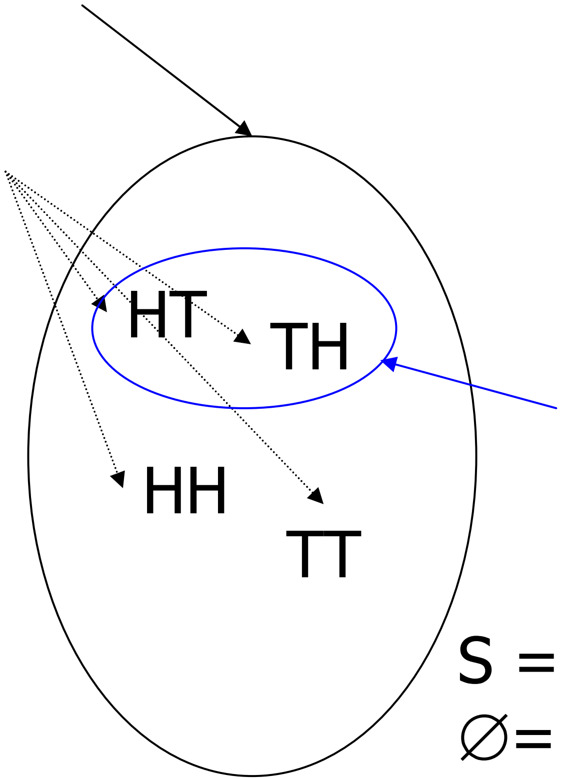
1	$(x+y)^0 = 1$
11	$(x+y)^1 = x+y$
121	$(x+y)^2 = x^2 + 2xy + y^2$
1331	$(x+y)^3 = x^3 + 3x^2y + 3xy^2 + y^3$
14641	$(x+y)^4 = x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4$

Don't panic!

Probability

Sample space S

Elementary events



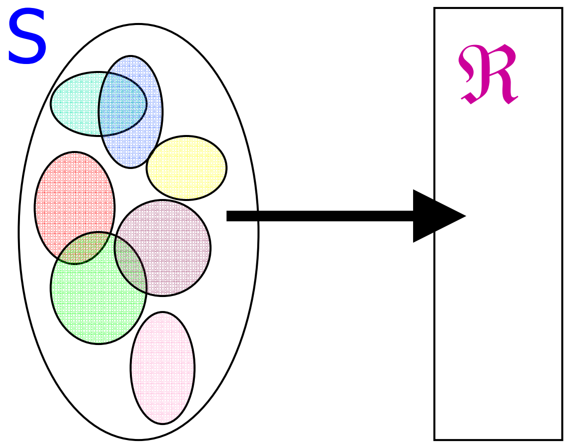
Probability defined in terms of a *sample space*.
Elementary events = outcomes of an experiment, e.g. head/tail.
An *event* is a subset of S: obtaining one tail and one head = {HT,TH}.

S = certain event.
$\emptyset$= null event.
Exclusive events=disjoints subsets.
Elementary events are exclusive.

Don't panic!

Axioms of Probability

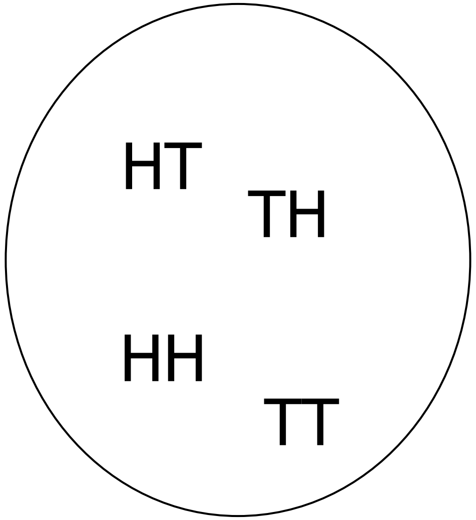
Probability distribution Pr{} =mapping from events of S to real numbers s.t.



- $\Pr\{A\} \geq 0$ for any A.
- $\Pr\{S\} = 1$.
- $\Pr\{A \cup B\} = \Pr\{A\} + \Pr\{B\}$ for exclusive events.
- Pr{A} is called the probability of the event A.
    - $\Pr\{\emptyset\} = 0$
    - $A \subseteq B \Rightarrow \Pr\{A\} \leq \Pr\{B\}$
    - $\Pr\{\underline{A}\} = 1 - \Pr\{A\}$
    - $\Pr\{A \cup B\} = \Pr\{A\} + \Pr\{B\} - \Pr\{A \cap B\}$

Example

Experiment: Toss 2 coins.
Elementary events: outcomes.
Each elementary event has probability ¼.



Probability of getting at least one head?

$$\Pr\{HT,TH,HH\}=\Pr\{HT\}+\Pr\{TH\}+\Pr\{HH\}$$

$$=3/4.$$

$$=\Pr\{\overline{TT}\}=1-\Pr\{TT\}$$

Don't panic!

Discrete Probability Distributions

- A probability distribution is discrete if it is defined over a **countable** sample space.

$$\Pr\{A\} = \sum_{s \in A} \Pr\{s\}$$

(s elementary event)

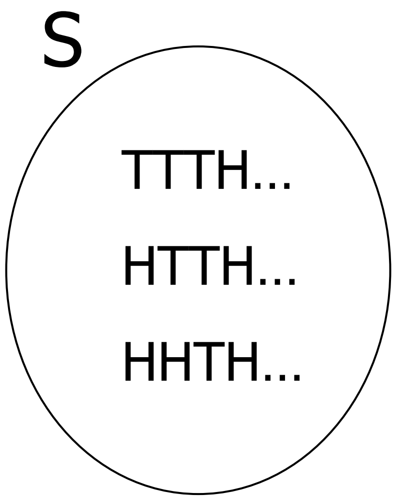
- Uniform probability distribution if

$$\Pr\{s\} = 1/|S|$$

(Pick an element at random, fair coin…)

Example

Experiment: Flip a *fair* coin *n* times.
Sample space = {H,T}^{n} of size 2^n.
Elementary events = sequences of length *n*, of probability 1/2^n.



Pr{A} with ?
A={exactly k heads and exactly n-k tails}

$$|A| = \binom{n}{k}$$

$$\Pr\{A\} = \binom{n}{k} / 2^n$$

Conditional Probability

- We have partial knowledge of the outcome.
  - Flip 2 coins.
  - Tell your friend one is a head.
  - It’s not possible that both are tails! Pr{TT} knowing there is one H=0.
  - The remaining events are equally equal {HT,TH,HH}.
  - Pr{HH} knowing there is one H=1/3.
  - Pr{HH} not knowing there is one H=1/4.

Don't panic!

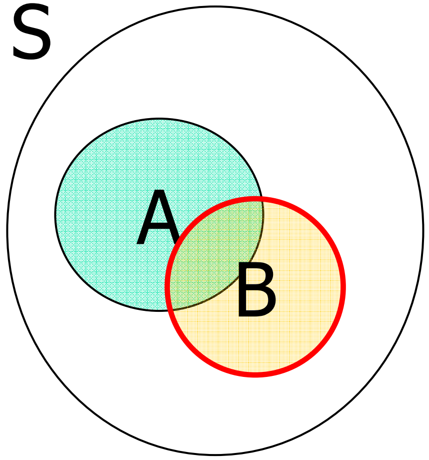
Conditional Probability

- "Probability of A given B" (with Pr{B}≠0):

$$\Pr\{A|B\} = \frac{\Pr\{A \cap B\}}{\Pr\{B\}}$$

← Both A and B occur.

← Outcome is in B, we normalize.



Previous example:
Pr{HH|there is a H}=(1/4)/(3/4).

Independent Events

- Events A_i are pair-wise independent if $\Pr\{A_i \cap A_j\} = \Pr\{A_i\}\Pr\{A_j\}$.
- Events are mutually independent if

$$\Pr\left\{\bigcap_{i \in K} A_i\right\} = \prod_{i \in K} \Pr\{A_i\}$$ for a subset K of indices.

- If A and B are independent, $\Pr\{A|B\} = \Pr\{A\}$.

Don't panic!

Bayes' Theorem

- For two events A and B with non-zero probabilities:

$$\Pr\{A \cap B\} = \Pr\{B\}\Pr\{A \mid B\}$$

$$= \Pr\{A\}\Pr\{B \mid A\}$$

$$\downarrow$$

$$\boxed{\Pr\{A \mid B\} = \frac{\Pr\{A\}\Pr\{B \mid A\}}{\Pr\{B\}}}$$

Bayes’ Theorem

- Alternative form using

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

↓

$$\boxed{\Pr\{A \mid B\} = \frac{\Pr\{A\}\Pr\{B \mid A\}}{\Pr\{A\}\Pr\{B \mid A\} + \Pr\{\overline{A}\}\Pr\{B \mid \overline{A}\}}}$$

$$\Pr\{A \mid B\} = \frac{\Pr\{A\}\Pr\{B \mid A\}}{\Pr\{A\}\Pr\{B \mid A\} + \Pr\{\overline{A}\}\Pr\{B \mid \overline{A}\}}$$

Example

We have a fair coin and a biased coin (always head).
Experiment: Choose a coin, flip it twice.
Suppose it comes up head twice.
What is the probability that it is biased?

?

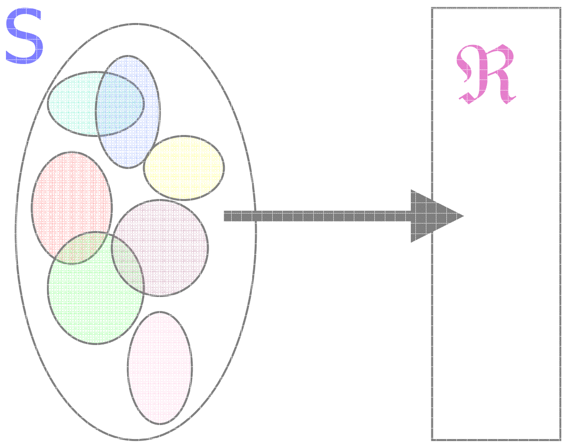
A=Event that the biased coin is chosen.
B=Event that the coin comes up head twice.
Pr{A|B}?

$\Pr\{A\}=1/2$, $\Pr\{B|A\}=1$, $\Pr\{\overline{A}\}=1/2$, $\Pr\{B|\overline{A}\}=1/4$

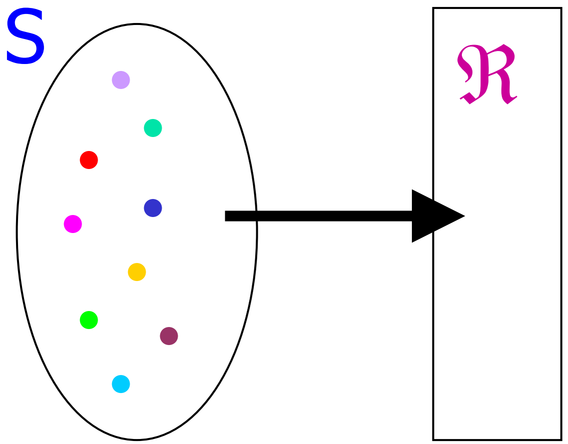
$$\Pr\{A \mid B\} = \frac{(1/2)\cdot 1}{(1/2)\cdot 1 + (1/2)\cdot(1/4)} = 4/5$$

Discrete Random Variables

Probability distribution Pr{}
=mapping from events of S
to real numbers s.t. … (axioms)



Discrete random variable X=
function from *outcomes of S*
to real numbers.



Don't panic!

Discrete Random Variables

- For a random variable X and a real number x (we choose),
  the event "X=x" is defined as
  {s∈S : X(s)=x}.

$$\Pr\{X = x\} = \sum_{\{s \in S : X(s) = x\}} \Pr\{s\}$$

(called the probability density function of X)

Example

Experiment: Roll two 6 sided dices.
36 elementary events in the sample space.
Uniform probability distribution: $\Pr\{s\}=1/36$.
X=maximum of the two values on the dice.
$\Pr\{X=3\}$?

Corresponding set of elementary events:
(1,3), (2,3), (3,3), (3,2), (3,1).
$\Pr\{X=3\}=5/36$.

Expected Value

- *Average* of the values of a random variable.

$$E[X] = \sum_{x} x \Pr\{X = x\}$$

$$E[X + Y] = E[X] + E[Y]$$

Example

Game: You flip two fair coins.

You earn \$3 for heads, but lose \$2 for tails.

X=your earning.

$E[X]$? ?

$$E[X]=6*\Pr\{HH\}+1*\Pr\{HT\}+1*\Pr\{TH\}-4*\Pr\{TT\}$$

$$=6(1/4)+1/4+1/4-4(1/4)=1$$

Geometric & Binomial Distributions

- Appendix C.4.
- Interesting read.

Probabilistic Analysis and Randomized Algorithms

- What is it about?
  - Worst case analysis = worst cost or running time of an algorithm.
  - Probabilistic = average in terms of cost or running time.
  - Randomized algorithms = algorithms that have a randomized decision but the result is not random!

Hiring Problem Example

- Hire new office assistant and fire the old (worse) one.
    - Cost of interviewing and hiring $O(nc_i+mc_h)$.
    - Idea: Cheap to interview, expensive to hire.
    - Result is independent of the ordering.

```
best=0
for i = 1 to n do
        interview candidate i
        if candidate i is better than candidate best then
                best=i
                hire candidate i
```

Hiring Problem Example

- Worst case = hire everyone (focus on hiring): $O(nc_h)$.
- More interesting: probabilistic analysis.
  - Saying that applicants arrive in a random order is the same as choosing randomly any possible permutation of applicants (n! permutations).
  - Expected cost of hiring?
  - How do we do?

Indicator Variables

- I{A}=1 if A occurs, I{A}=0 if it does not.
- Example: Expected number of heads if we flip one fair coin.
  - S={H,T}
  - X=I{Y=H}
  - E[X]=1*Pr{Y=H}+0*Pr{Y=T}=1/2.
- Lemma: Given an event A in S, let X_A=I{A}. Then E{X_A}=Pr{A}.

Hiring Problem cont.

- X_i=I{hire candidate i}.
- X=Number of hired candidates=$X_1+X_2…X_n$.
  - $E[X]=E[X_1]+E[X_2]+…E[X_n]$
- $E[X_i]$=Pr{hire candidate i}=1/i.
- E[X]=sum(1/i) harmonic serie, see A.7 $E[X]=\ln n+O(1)$.
- Expected cost = $O(c_h \ln n)$

Randomized Algorithms

- What we did: Knowing the distribution (uniform) of an input, we analyzed an algorithm.
- When we don’t know, we can impose a distribution beforehand ⇒ we can analyze such a *randomized* algorithm for *any input*.
- The point: Even your worst enemy cannot produce a bad input since the execution time depends on the (randomized) algorithm.
  - Every run is different.
  - The end result is the same.

Permuting Arrays

- A way to randomize inputs
- Permute by sorting
  - Assign random priorities and sort $\Theta(n \lg n)$.

```
n=length[A]
for i = 1 to n do P[i]=Random(1,$n^3$)
sort A using P as keys
```

  - It is a uniform random permutation, i.e., every output has probability 1/n!.

Permuting Arrays

- Randomize in-place.
    - Swap elements randomly.

```
n=length[A]
for i = 1 to n do swap(A[i],A[Random(i,n)])
```

    - It is also a uniform random permutation.
    - Proof technique: Based on a loop invariant.
        - Initialization.
        - Maintenance.
        - Termination.