Probabilistic Tools in Algorithms

Curs 2019

Review of basic mathematics

- Arithmetic Series: $\sum_{i=1}^{n} i = \frac{n(n+1)}{2} = \Theta(n^2)$.
- Geometric Series: for $x \neq 1$, $\sum_{i=0}^{n} x^i = \frac{x^{n+1}-1}{x-1}$.
- Geometric Series: for $|x| < 1$, $\sum_{i=0}^{n} x^i = \frac{1}{1-x}$.
- Harmonic Series: for n finite,

$$H_n = \sum_{i=0}^{n} \frac{1}{i} = \ln n + O(1).$$

Note that if $n \to \infty$ then $\sum_{i=0}^{n} \frac{1}{i}$ diverges.

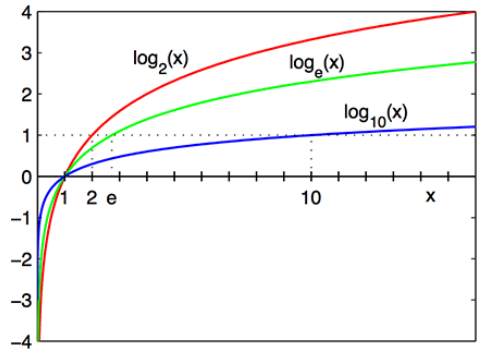
Review of basic mathematics: Log and Exponential

$\log_b n = x$ means $n = b^x$,

$\log(x_1 + x_2 + \cdots + x_n) = \sum_{i=1}^{n} \log x_i$

$\log(x^{f(x)}) = f(x) \log x \Rightarrow 2^{\lg n} = n.$

$\log_a x = \frac{\log_b x}{\log_a b}$



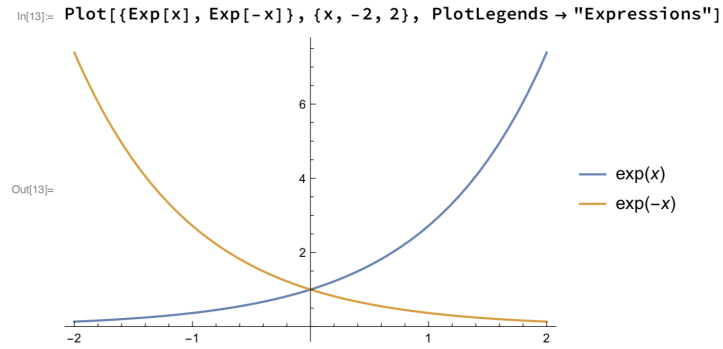
Recall:

$$\frac{d}{dx} \ln(f(x)) = \frac{\frac{d(f(x))}{dx}}{f(x)}.$$

$$\frac{d}{dx} \ln x = \frac{1}{x}.$$

Review of basic mathematics: Exponential

$\ln n = \log_e n = x$ means $n = e^x$,
where $e = \lim_{n\to\infty}(1 + \frac{1}{n})^n \sim 2.71...$
$e^x = \lim_{n\to\infty}(1 + \frac{x}{n})^n$.
$e^{-x} = \lim_{n\to\infty}(1 - \frac{x}{n})^n$. $\frac{d}{dx}e^x = e^x$.



Recall Taylor: $f(x)$ differentiable at a: $f(x) = \sum_{i=0}^{\infty} \frac{f^i(a)}{i!}(x-a)^i$

Therefore $e^x = 1 + x + \frac{x^2}{2!} + \cdots$, $e^{-x} = 1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \cdots$

$\boxed{e^x < 1 + x}$ and $\boxed{e^{-x} > 1 - x}$

In fact, for x is vey small $0 < x << 1$: $\boxed{e^{-x} \sim 1 - x}$

$e^{-0.4} = 0.67032, e^{-0.1} = 0.904837, e^{-0.01} = 0.99005$

Binomial

- Stirling: $n! = \sqrt{2\pi n}(\frac{n}{e})^n + \gamma + O(1/n)$,

$$\boxed{n! \sim n^n}.$$

- Binomial: $\binom{n}{k} = \frac{n!}{(n-k)!k!}$
- Binomial Thm.: $(a+b)^n = \sum_{i=0}^{n} \binom{n}{i} a^{n-i} b^i$.
  $\therefore (1+x)^n = \sum_{i=0}^{n} \binom{n}{i} x^i = 1 + nx + \frac{n(n-1)}{2}x^2 + \cdots + x^n$
- Important: $\boxed{\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{ne}{k}\right)^k}$
- Also useful If $k = o(\sqrt{n})$ then $\binom{n}{k} \sim \frac{n^k}{k!}$

Why using asymptotic notation?

Considering that an instance with size $n = 1$ takes 1 μ second:

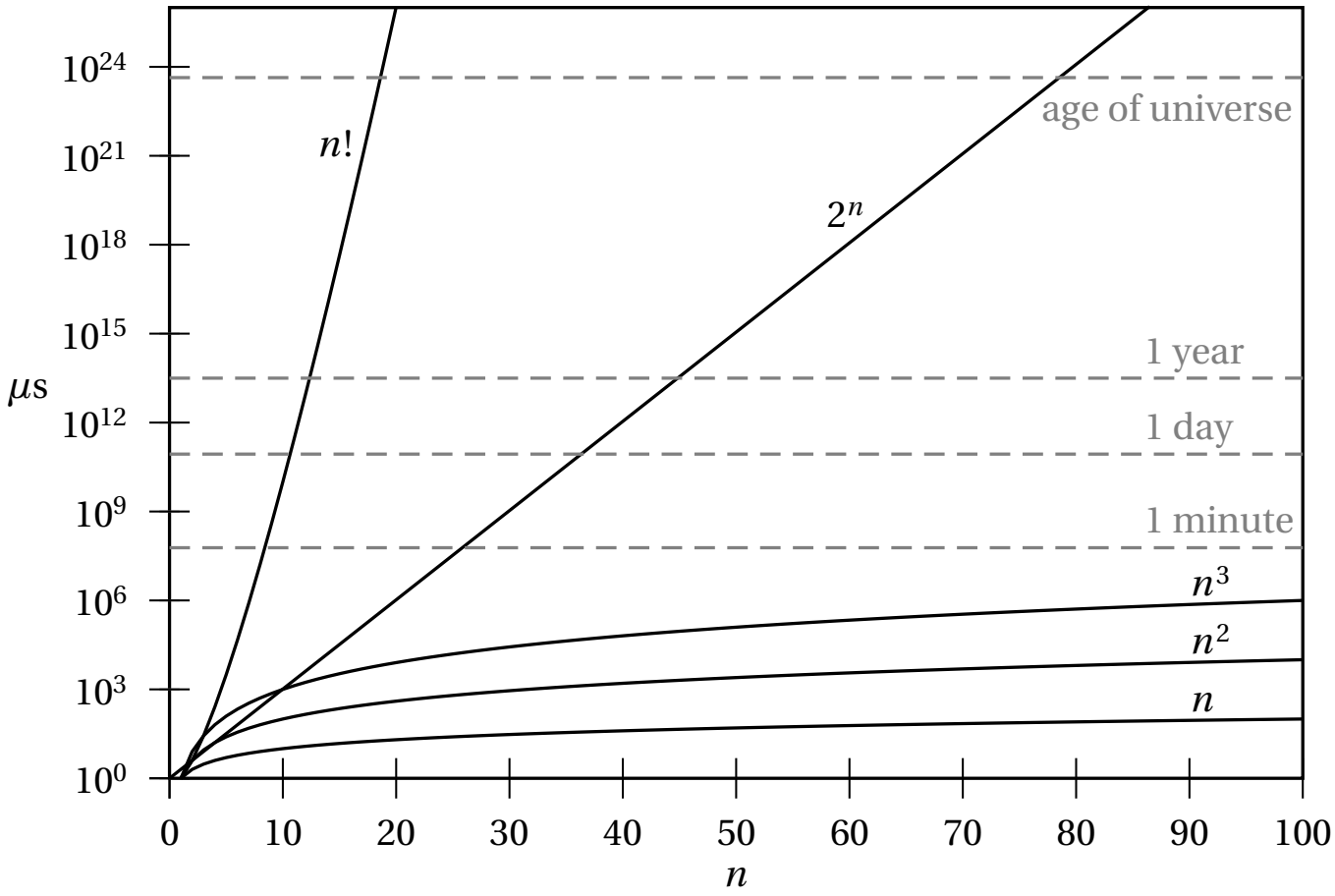


Table of computing times according to the size of an instance.

Recall: Asymptotic notation

Símbol	$L = \lim_{n\to\infty} \frac{f(n)}{g(n)}$	intuïció ..
$f(n) = O(g(n))$	$L < \infty$	$f \leq g$
$f(n) = \Omega(g(n))$	$L > 0$	$f \geq g$
$f(n) = \Theta(g(n))$	$0 < L < \infty$	$f = g$
$f(n) = o(g(n))$	$L = 0$	$f < g$
$f(n) = \omega(g(n))$	$L = \infty$	$f > g$
$f(n) \sim g(n)$	$L \to 1$	

For ex. $\log_a x = \Theta(\log_b x)$, for any $a, b > 0$.

Remember: Basic Combinatorics

- Permutation of a set S with n elements are all ordered sequences of length n without repetition.
  Ex.: $S = \{a, b, c\}$ then $abc, acb, bac, bca, cab, cba$.
  There are $n!$ permutations of S.
- k-Permutation of a S with $|S| = n$ $(k \leq n)$ are all ordered sequences of length k without repetition.
  Ex.: $S = \{a, b, c\}$, then 2-permutation ab, ac, ba, bc, ca, cb.
  There are $P(n, k) + \frac{n!}{(n-k)!}$ k-permutations of S.
  $P(n, k) = n(n-1)(n-2)\cdots(n-k+1)$.
- Given S, $|S| = n$ and $m > n$ the number of ordered m-sequences with repetitions we can form with elements in S is S^m.
  Ex. The number of binary sequences with length 5 is $\{0, 1\}^5 = 2^5$.

k-Combination: Binomial

A k-Combination of a S with $|S| = n$ ($k \leq n$) are all non-ordered sequences of length k without repetition. Ex.: $S = \{a, b, c\}, k = 2$ then ab, ac, bc

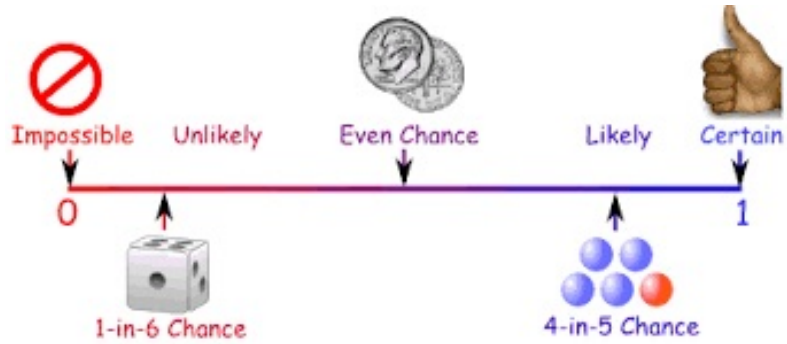
Let $\binom{n}{k}$ the number of different (non-ordered) k-subsets from a set S with k elements.

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!} = \frac{n!}{k!(n-k)!}.$$

Notice $\binom{n}{0} = \binom{n}{n} = 1$ and $\binom{n}{k} = \binom{n}{n-k}$.

What is probability?

Probability: useful technique to simulate and explain real world. Any english speaking person understands the words *likely* and *unlikely*.



But in everyday life, do we consciously think in terms of probability?

What is probability?

As far as we know, many phenomena in *nature* seem to be generated by random choices, but it is difficult to simulate truly unpredictable random experiments:

Flipping a coin or tossing a dice are *deterministic* experiments; Given the initial angle of the coin, the spin, humidity, etc. we can predict the outcome of flipping a coin.

In the same way, in todays computers, the *random generator* functions are *deterministic* programs, which simulates randomness. What is denoted pseudorandom generators.

Probability and computers

The most basic method is the linear congruential generator: from a seed integer $x_0 \in \mathbb{Z}^+$, produce a sequence of pseudo-random values

$$x_{n+1} = (ax_0 + b) \mod m,$$

for a, b constants and m a large integer.
For ex, C/C++ (rand( )) m a 32-bit integer, $a = 22695477, b = 1$
So a computer deterministically generates pseudorandom numbers.

How would generate a sequence of pseudorandom bits?

```
for (i = 0; i < n; i++) {
values[i]=rand() % 2;
printf("%d", values[i]);
}
```

Some applications of probability in CS

- Algorithm design: Making algorithms run faster by introducing probability choices, agains "bad" inputs.
- Data structure: when implementing most of the used data structures, e.g. dictionaries, the use of probability helps to speed up search and reduce space.
- Learning theory: in learning theory one assumes the data is generated according to specific probability distributions.
- Studying and design mechanisms for large complex networks: The design of algorithms for Internet, WWW, Facebook, etc, is based in the design realistic probabilistic models for those huge networks.

Some applications of probability in CS

- Data science: To design efficient algorithm for huge data set, usually we do sampling of a small portion of the data and compute the statistic, rather than read all the data.
- Cryptography: Randomness and number theory, are essential for cryptography and crypto-hashing.
- Data compression: improving data compression algorithms passes through analysing and modelling the underlying probability distribution of the data, and evaluating its information-theoretic contets.
- Modelling and analysing the spread of particular infections: Probabilistic ad-hoc graph models and techniques, have play an important role in helping to stop or mitigated massive infections, including e-infections.

Experiments and Events

Probability space (Ω): the set of outcomes associated with an experiment.

Basic events: the elements in Ω.
Event: $E \subseteq \Omega$, i.e. an event is any collection of outcomes.

Example: Flip two coins:

- Basic events $\Omega = \{HH, HT, TH, TT\}$. $|\Omega| = 4$.
- Non-basic event: Let A be the event of having at least one H, then $A = \{HH, HT, TH\}$.

Given Ω, define $\mathcal{F}$ as the set of all events in the power set of Ω, $\mathcal{P}(\Omega)$.

For any event $E \in \mathcal{F}$, let $\bar{E}$ the set of events $\mathcal{F} \backslash E$

Probability

Given $\mathcal{F}$ on Ω, define the probability function $\mathbf{Pr} : \mathcal{F} \to [0, 1]$ such that:

1. For any event $A \in \mathcal{F}$: $0 \leq \mathbf{Pr}[A] \leq 1$, $\mathbf{Pr}[\Omega] = 1, \mathbf{Pr}[\emptyset] = 0$.
2. Given all basic events $\{E_i\}_{i=1}^n, \sum_{i=1}^n \mathbf{Pr}[E_i] = 1$,
3. If $\{A_j\}_{j=1}^k$ are mutually exclusive in $\mathcal{F}$ then

$$\mathbf{Pr}\left[\cup_{j=1}^k A_j\right] = \sum_{j=1}^k \mathbf{Pr}[A_j].$$

In a Probability Space $(\Omega, \mathcal{F}, \mathbf{Pr}[])$, the set of basic events $\{E_i\}_{i=1}^n$ form a partition of Ω, i.e. they are disjoints, therefore $\sum_{i=1}^n \mathbf{Pr}[E_i] = 1$.

Probability distribution

If $\Omega = \{E_i\}_{i=1}^{n}$ define the Probability distribution $\mathcal{D}$ for $\mathbf{Pr}[]$ as $\mathcal{D} = \{\mathbf{Pr}[E_1], \ldots, \mathbf{Pr}[E_n]\}$.

In discrete probability, if $|\Omega| = n$, then $\mathcal{D}$ is said to be the uniform distribution if for any basic event E_i we have $\mathbf{Pr}[E_i] = \frac{1}{n}$.

Given a probability space with a distribution, we select uniformly at random (u.a.r.) an element in Ω if we choose equal probability among all basic events.

Examples:
Flip 3 coins $|\Omega| = 2^3 = 8$, so probability of choosing u.a.r :
$\mathbf{Pr}[000] = \mathbf{Pr}[011] = 1/8$.
If A is the event that we choose an element with two 1's,
$\mathbf{Pr}[A] = \mathbf{Pr}[011] + \mathbf{Pr}[101] + \mathbf{Pr}[110] = 3/8$

More on events

In general, an event A is a collection of outcomes, i.e. $A \subseteq \Omega$
Given an event $A \subseteq \Omega$ we define its probability:

$$\mathbf{Pr}[A] = \sum_{\omega \in \Omega} \mathbf{Pr}[\omega],$$

where indistinctly we can denote the basic events as E or ω.

Example-1: Flip a fair coin. If it comes up heads, roll a 3-sided die; if it comes up tails, roll a 4-sided die. What is the probability that the die roll is at least 3?

$\Omega = \{(H,1),(H,2),(H,3),(T,1),(T,2),(T,3),(T,4)\}$, $|\Omega| = 7$

As $A = \{(H,3),(T,3),(T,4)\}$
$\Rightarrow \mathbf{Pr}[A] = \mathbf{Pr}[(H,3)] + \mathbf{Pr}[(T,3)] + \mathbf{Pr}[(T,4)] = 5/12.$

Examples

Example-2: We have a unit square $\mathcal{S}$, and inside a circle C centered at the point $(0.5, 0.5)$ and of radius $r = 1/4$. If we throw u.a.r. a point to $\mathcal{S}$, which is the probability it hits inside C?

The probability is $= \frac{\text{Area } C}{\text{Area } \mathcal{S}} = \pi(1/4)^2 = 0.1965$

Example-3: A bag contains 100 balls, 50 reds and 50 blue. We select 5 balls independently and u.a.r. What is the probability that 3 are blue and 2 are red?

The total number of outcomes $|\Omega| = \binom{100}{5}$. Therefore the probability is:

$$\frac{\binom{50}{3}\binom{50}{2}}{\binom{100}{5}}.$$

Some consequences of the probability properties

Given $A, B, C \in \mathcal{F}$:

- $\mathbf{Pr}\left[\bar{A}\right] = 1 - \mathbf{Pr}\left[A\right]$.
- If $A \subset B$ then $\mathbf{Pr}\left[B\right] = \mathbf{Pr}\left[A\right] + \mathbf{Pr}\left[B \backslash A\right] \geq \mathbf{Pr}\left[B\right]$.
- $\mathbf{Pr}\left[A \cup B\right] = \mathbf{Pr}\left[A\right] + \mathbf{Pr}\left[B\right] - \mathbf{Pr}\left[A \cap B\right]$.
  Pf. Events $(A \backslash B)$, $(A \backslash B)$ and $(A \cap B)$ are independent.
- **Inclusion-Exclusion 3 events**

$$\begin{aligned}\mathbf{Pr}\left[A \cup B \cup C\right] = & \mathbf{Pr}\left[A\right] + \mathbf{Pr}\left[B\right] + \mathbf{Pr}\left[C\right] \\ & - \mathbf{Pr}\left[A \cap B\right] - \mathbf{Pr}\left[B \cap C\right] - \mathbf{Pr}\left[A \cap C\right] \\ & + \mathbf{Pr}\left[A \cap B \cap C\right].\end{aligned}$$

Inclusion-Exclusion and Union-Bound

Inclusion-Exclusion Given n events $\{A_1, \ldots, A_n\}$,

$$\mathbf{Pr}\left[\cup_{i=1}^{n} A_i\right] = \sum_{i=1}^{n} \mathbf{Pr}\left[A_i\right] - \sum_{i<j} \mathbf{Pr}\left[A_i \cap A_j\right] + \sum_{i<j<k} \mathbf{Pr}\left[A_i \cap A_j \cap A_k\right] - \cdots (-1)^{l+1} \sum_{i_1<\cdots<i_l} \mathbf{Pr}\left[\cap_{r=1}^{l} A_{ir}\right].$$

Very useful upper-bound to the probability of non-exclusive events:
Trick 1: Union-Bound. Given non-independent events $\{A_i\}_{i=1}^{n}$,

$$\mathbf{Pr}\left[\cup_{i=1}^{n} A_i\right] \leq \sum_{i=1}^{n} \mathbf{Pr}\left[A_i\right].$$

Basic Example

Given a k-dimensional vector $K[1, \ldots, k]$ and a set $S = \{1, 2, \ldots, n\}$, where $n >> k$, we want to compute the probability of having a random assignment $S \to K$, such that no two integers in S are repeated.

We want:

(# assignments $S \to K$ without repeated integers)/(total # of assignments)

Total # of assignments $S \to K$: n^k

assignments $S \to K$ without repeated integers:

$n(n-1)(n-2)\cdots(n-k+1)$

Therefore,

$$\frac{n(n-1)\cdots(n-k+1)}{n^k} = \frac{n}{n}\frac{n-1}{n}\cdots\frac{n-k+1}{n} = 1\cdot(1-\frac{1}{n})\cdot(1-\frac{2}{n})\cdots(1-\frac{k-1}{n})$$

where $(1 - \frac{j}{n})$ is the probability of no-choosing the same integer in $K[j]$ that in any of the previous $K[i]$ for $1 \leq i < j$.

Independent and correlated events

Given events A, B on Ω, the are said to be independent (mutually independent) if if $\mathbf{Pr}[A \cap B] = \mathbf{Pr}[A] \times \mathbf{Pr}[B]$, otherwise they are said to be correlated or dependent.
Events $A_1, A_2, \cdots A_n$ are independent if

$$\mathbf{Pr}[A_1 \cap A_2 \cap \cdots \cap A_n] = \prod_{i=1}^{n} \mathbf{Pr}[A_i].$$

Notice the basic events in Ω are not independent, although they are disjoint.
For example, if we flip a coin, and E_1 is the event of (H), and E_2 is the event of (T), then $\mathbf{Pr}[E_1]\,\mathbf{Pr}[E_2] = \frac{1}{4} \neq 0 = \mathbf{Pr}[E_1 \cap E_2]$

But if the experiment is flipping twice a coin and E_1 is the event of (H) in the 1st flip E_2 = event of (H) in the 2nd. flip, then E_1 and E_2 are independent.

Independent and correlated events

Toss 2 fair coins ($\mathcal{D} = \{1/4, 1/4, 1/4, 1/4\}$) and consider the A: there is at least 1 head, B: there is at least one tail.
$\Omega = \{HH, TT, TH, HT\} \Rightarrow \mathbf{Pr}[A] = \frac{3}{4} = \mathbf{Pr}[B] = \frac{3}{4}$
but $\mathbf{Pr}[A \cap B] = \frac{2}{4} \neq \frac{3}{4}\frac{3}{4} = \frac{9}{16}$
Therefore A and B, are dependent (correlated).

We draw sequentially 2 cards from a 52 card deck.
Let R_1 be the event of drawing a red card on the first trial and R_2 the event of drawing a red card on the second trial. (26 are red and 26 are black)
If the draws are with replacement R_1 and R_2 are independent, if it is without replacement R_1 and R_2 are not independent. Why?

Formal proof sampling without replacement are not independent events

Draw sequentially 2 cards from a 52 deck. Let R_1 be the event of drawing a red card on the first trial and R_2 the event of drawing a red card on the second trial. If we draw without replacement, R_1 and R_2 are not independent.

Let B_1 event of drawing a black card 1st. trial.

Recall: $\mathbf{Pr}[R_1] = \frac{26}{52}$ and $\mathbf{Pr}[B_1] = \frac{26}{52}$.

Need $\mathbf{Pr}[R_1 \cap R_2] =^{?} \mathbf{Pr}[R_1]\,\mathbf{Pr}[R_2]$

After R_1, prob. drawing another $R = \frac{25}{51} \Rightarrow \mathbf{Pr}[R_1 \cap R_2] = \frac{26}{52}\frac{25}{51}$

So $\mathbf{Pr}[R \text{ then } R] = \frac{26}{52}\frac{25}{51}$ and $\mathbf{Pr}[B \text{ then } R] = \frac{26}{52}\frac{26}{51}$

$\Rightarrow \mathbf{Pr}[R_2] = \frac{26}{52}\frac{25}{51} + \frac{26}{52}\frac{26}{51} = \frac{26}{51}$.

$$\therefore \mathbf{Pr}[R_1 \cap R_2] = \frac{26}{52}\frac{25}{51} \neq \frac{26}{52}\frac{26}{51} = \mathbf{Pr}[R_1]\,\mathbf{Pr}[R_2].$$

Pairwise independence

Given a set of events $\{A_i\}$ they are said to be pairwise independent if every pair (A_i, A_j) is independent, i.e.

$$\mathbf{Pr}[A_i \cap A_j] = \mathbf{Pr}[A_i]\,\mathbf{Pr}[A_j].$$

Mutually independence $\Rightarrow$ pairwise independence,
but pairwise independence not necessarily $\Rightarrow$ independence:

Example: Throw 2 dice. Let A_1 be the event sum of the points is 7;
Let A_2 be the event dice 1 is 3; Let A_3 be the event dice 2 is 4.
$\mathbf{Pr}[A_1] = \mathbf{Pr}[A_2] = \mathbf{Pr}[A_3] = 1/6$ and
$\mathbf{Pr}[A_1 \cap A_2] = \mathbf{Pr}[A_1 \cap A_3] = \mathbf{Pr}[A_3 \cap A_2] = 1/36$
but $\mathbf{Pr}[A_1 \cap A_2 \cap A_3] = \frac{1}{36}$, while $\mathbf{Pr}[A_1] \cdot \mathbf{Pr}[A_2] \cdot \mathbf{Pr}[A_3] = \frac{1}{216}$.

Conditional probability

One of the important concepts in probability is conditioning, which means revising probabilities on an event A based on *partial information* that we know, i.e. based in another event B.

Flip 2 fair coins. Given that event B that one of them is H, what is the probability of the even A that both of them are H?
$\mathbf{Pr}[A|B] = 1/3$, *as the information B reduces the probability space to $\{TH, HT, HH\}$, each one with probability $1/3$.*

Formal definition of conditional probability:

$$\mathbf{Pr}[A|B] = \frac{\mathbf{Pr}[A \cap B]}{\mathbf{Pr}[B]} = \frac{\mathbf{Pr}[B \cap A]}{\mathbf{Pr}[B]} = \frac{\mathbf{Pr}[B|A]\,\mathbf{Pr}[A]}{\mathbf{Pr}[B]}.$$

In previous ex.: $\mathbf{Pr}[A|B] = \frac{\mathbf{Pr}[A \cap B]}{\mathbf{Pr}[B]} = \frac{1/4}{3/4}$.

Alternative definition of independence:
A and B are independent iff $\mathbf{Pr}[A|B] = \mathbf{Pr}[A]$.

The Russian roulette

Two people play one round of Russian roulette. The gun is revolver with six chambers, all empty. The players put two bullets into adjacent chambers of the barrel. The first player takes the gun and spin the barrel, I put the gun in his head and pull the trigger and no bullet. Gives the gun to the second player for his turn. Which would be better for the second player, to spin the barrel first, or just pull the trigger?

The Russian roulette

Two people play one round of Russian roulette. The gun is revolver with six chambers, all empty. The players put two bullets into adjacent chambers of the barrel. The first player takes the gun and spin the barrel, I put the gun in his head and pull the trigger and no bullet. Gives the gun to the second player for his turn. Which would be better for the second player, to spin the barrel first, or just pull the trigger?

Each time we spin the barrel, the probability of getting a bullet is $2/6 = 1/3$ so the probability of surviving is $1 - 1/3 = 2/3$, i.e 66.66%

In the second time, if we do not spin the barrel, we are conditioning to the fact that we have 4 possible chambers in position, of which one has a bullet. So the probability of having a bullet is $1/4$, therefore the probability of non-having a bullet is $3/4 = 75\%$. So it is better no to spin the barrel.

Total probability law

When dealing with conditional probability, it seems that first we have to compute the probabilities involved in a random experiment, and then we can calculate the conditional probabilities.
In practice we use conditional probabilities to *reduce* the calculation of probabilities for events.

Total Probability Law If a set of events $\{E_i\}_{i=1}^n$ is a **partition** of Ω and $A \in \mathcal{F}$ is a event, then

$$\mathbf{Pr}[A] = \sum_{i=1}^{n} \mathbf{Pr}[A \cap E_i] = \sum_{i=1}^{n} \mathbf{Pr}[A|E_i]\,\mathbf{Pr}[E_i]\,.$$

Randomization and algorithmic: Probabilistic analysis

Given a deterministic algorithm, it happens that a few "instances" may bias the complexity outcome of the algorithm, which for most of the instances seem to work well, for ex. Quicksort.
In this cases, we can perform a probabilistic analysis of the deterministic algorithm as follows:
Define a probability distribution on the set of inputs, parametrized by input size. Often the distribution is the uniform, but not always.
We consider the number of steps as a random variable $T(n)$ and compute its expected value $\mu = \mathbf{E}[T(n)]$.
We also need to prove concentration, i.e. with high probability, for most of the imputs, $T(n)$ is near μ.

Randomization and algorithmic: Randomized algorithms

We can design a randomized algorithms, where the algorithm takes random choices and continues the computation according to the output of the random choices.

In this case, we may have to perform a probabilistic analysis of the complexity.

There are two main types of probabilistic algorithms:

- Monte-Carlo: Always halt in finite time, but may input the wrong answer. If the answer is binary (yes/not) the error can be in one direction, *one-side error*, or the error could be in both answers. In Monte-Carlo algorithms it is important to bound the error probability. *two-side error*.
- Las Vegas: The output is always correct but the running time may be unbounded.

It is easy to convert a Las Vegas algorithm into a Monte-Carlo, how?. The contrary is not always true.

In this course we will be working mainly with Monte-Carlo algorithms.

Generating a permutation uniformly at random

Given an array of n-distinct keys $K[1, \ldots, n]$ recall a permutation Π of K defines a re-ordering of the elements of K, where $\Pi(i) = j$ means $K[i] \rightarrow K[j]$.

If k has size n the number of different permutations is $n!$.

Considering the experiment of generating a random permutation, we get $\Omega = \{\Pi_1, \Pi_2, \ldots, \Pi_{n!}\}$, i.e. $|\Omega| = n!$.

Generating a permutation uniformly at random (u.a.r) means from ordered $K[1, \ldots,]$ generate a permutation $\Pi(K)$ with probability

$$\frac{1}{|\Omega|} = \frac{1}{n!}.$$

Randomized algorithm to generate u.a.r. a permutation

Algorithm Fisher-Yates (also known as Knuth's algorithm)

Given an array $K[1, \ldots, n]$ with different n. keys,

Random-Perm (K, n)
for $i = 1$ to n **do**
  choose $j = \text{Rand}(i, n)$
  swap $K[i] \leftrightarrow K[j]$
end for
return K

Lemma Random-Perm (K, n) produces a uar permutation of K in $\Theta(n)$ steps.

Proof

Time: Random-Perm has of n-iterations, each of cost of $\Theta(1)$.

u.a.r.: Pr. a key $\rightarrow K[1] = \frac{1}{n}$; Pr. a key $\rightarrow K[j] = \frac{1}{n-j}$; Therefore, prob. of a specific $\Pi(K) = \frac{1}{n} \cdot \frac{1}{n-1} \cdot \frac{1}{n-2} \cdots \frac{1}{2} \cdot 1 = \frac{1}{n!}$. □