Basics on Probability

RA-MIRI QT Curs 2020-2021

Review of basic mathematics

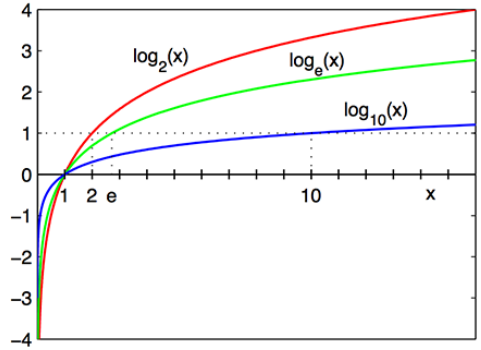
- Arithmetic Series: $\sum_{i=1}^{n} i = \frac{n(n+1)}{2} = \Theta(n^2)$.
- Geometric Series: for $x \neq 1$, $\sum_{i=0}^{n} x^i = \frac{x^{n+1}-1}{x-1}$.
- Geometric Series: for $|x| < 1$, $\sum_{i=0}^{n} x^i = \frac{1}{1-x}$.
- Harmonic Series: for n finite,

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

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

Review of basic mathematics: Log and Exponential

$\log_b n = x$ means $n = b^x$,
$\log(xy) = \log x + \log y$
$\log(x^{f(x)}) = f(x) \log x \Rightarrow 2^{\lg n} = n.$
$\log_a x = \frac{\log_b x}{\log_a b}$



Recall:

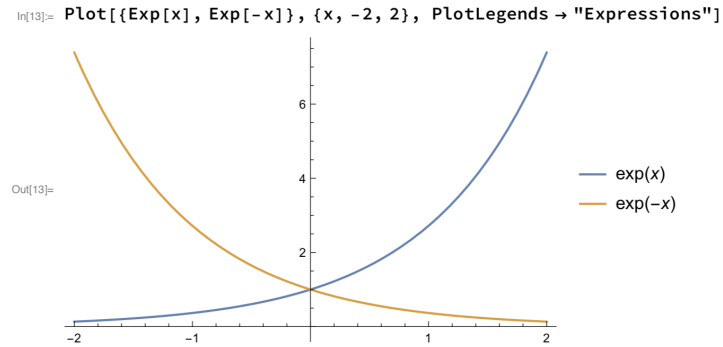
$$\frac{d}{dx} \ln(f(x)) = \frac{\frac{d(f(x))}{dx}}{f(x)}.$$

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

$\lg = \log_2$, $\ln = \log_e$, $\log = \log_{10}$

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

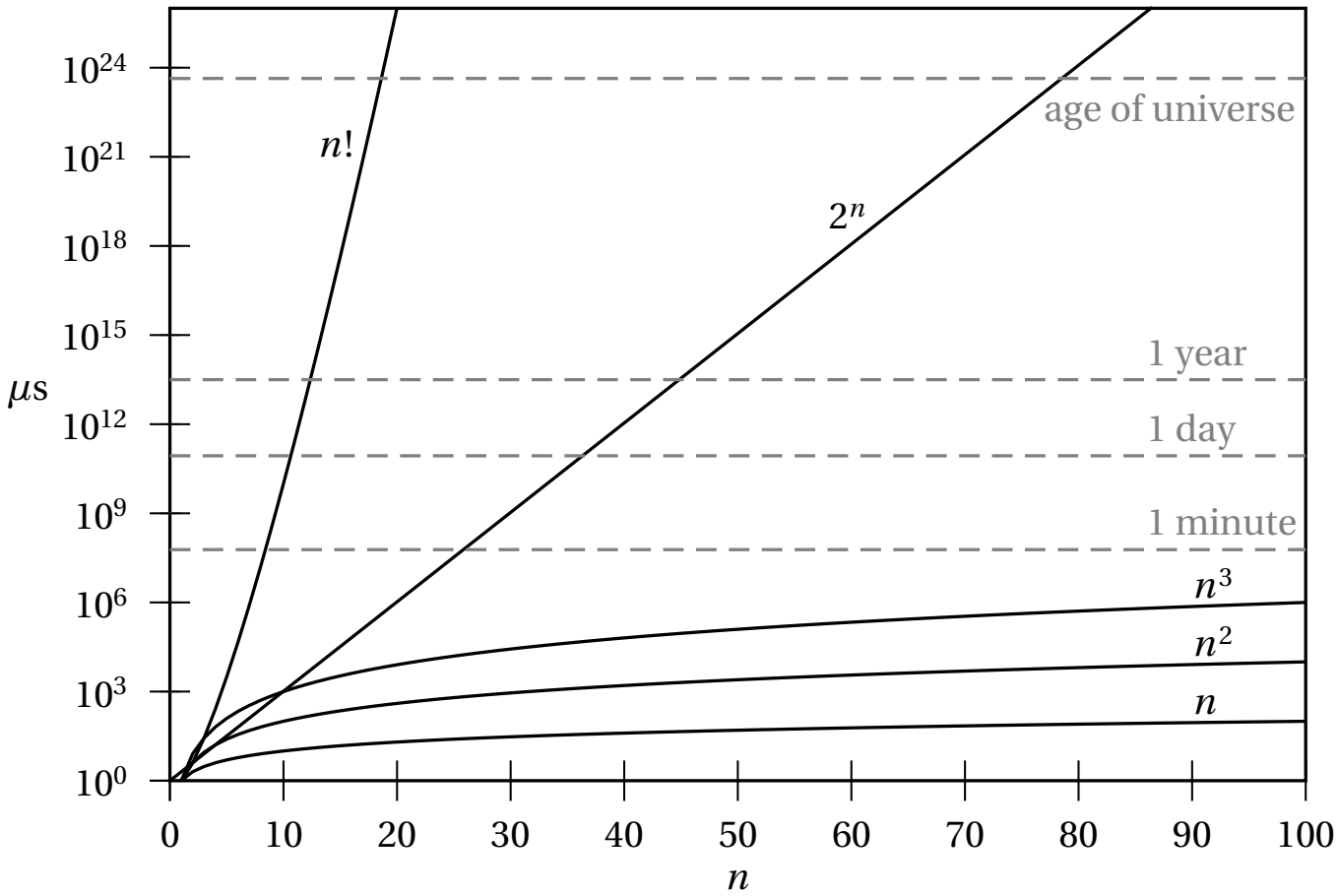


Table of computing times according to the size of an instance.

Recall: Asymptotic notation

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

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

Remember: Basic Combinatorics

For a set S with n elements

- The Permutations of S are all the 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.
- The k-Permutation of S ($k \leq n$) are all the ordered sequences of length k without repetition.
  The 2-permutations of $\{a, b, c\}$ are ab, ac, ba, bc, ca, cb.
  There are $P(n, k) + \frac{n!}{(n-k)!}$ k-permutations of S.
  $P(n, k) = n(n-1)(n-2)\cdots(n-k+1)$.
- For $m > n$ the number of ordered m-sequences with repetitions that we can form with elements in S is n^m.
  Ex. The number of binary sequences with length 5 is 2^5.

k-Combination: Binomial

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

This is the same as the number of different k-subsets, i.e.,

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

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

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 (distribution) $\mathbf{Pr} : \mathcal{F} \to [0,1]$ 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 events 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 mutually disjoint, therefore $\sum_{i=1}^n \mathbf{Pr}[E_i] = 1$ follows from 1 and 3.

Uniform distribution

In a discrete probability space, $|\Omega| = n$, the uniform distribution assigns to any basic event E_i, $\mathbf{Pr}[E_i] = \frac{1}{n}$.

Given a probability space we select uniformly at random (u.a.r.) an element in Ω if we choose wiith 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 $\mathcal{S}$ with side 1, and inside a circle C centered at the central point of S 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 red 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 $A \subseteq B$ then $\mathbf{Pr}\left[B\right] = \mathbf{Pr}\left[A\right] + \mathbf{Pr}\left[B \backslash A\right] \geq \mathbf{Pr}\left[A\right]$.
- $\mathbf{Pr}\left[A \cup B\right] = \mathbf{Pr}\left[A\right] + \mathbf{Pr}\left[B\right] - \mathbf{Pr}\left[A \cap B\right]$.
  Pf. Events $(A \backslash B)$, $(A \backslash B)$ and $(A \cap B)$ are disjoint.
- **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 to K, so that no two integers in S are repeated.

We watn to compute:

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

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

assignments 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 and consider the events: A, there is at least 1 head, and B, there is at least one tail.

$\Omega = \{HH, TT, TH, HT\} \Rightarrow \mathbf{Pr}[A] = \frac{3}{4} = \mathbf{Pr}[B] = \frac{3}{4}$

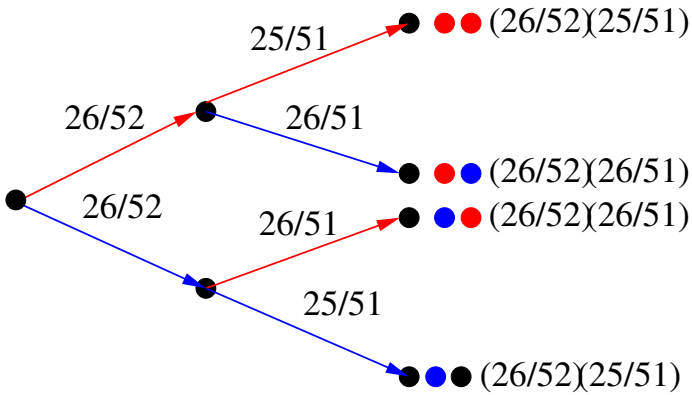
but $\mathbf{Pr}[A \cap B] = \frac{2}{4} \neq \frac{3}{4}\frac{3}{4} = \frac{9}{16}$

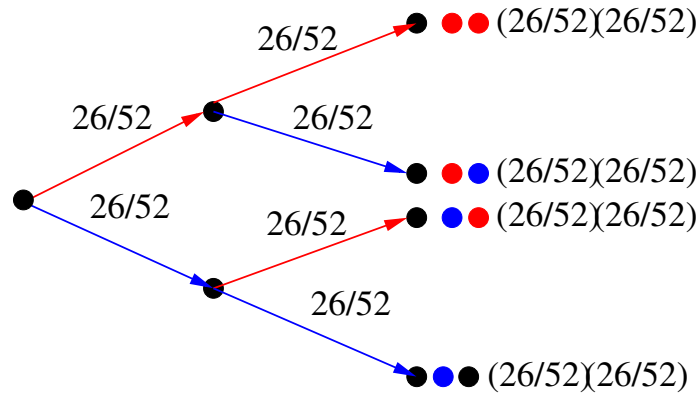
Therefore A and B, are dependent (correlated).

Sampling with replacement simplifies life

Important Example: We draw sequentially 2 cards from a deck with 52 cards, where 26 of the cards are red and the other half blue. 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 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.





Without replacement: $\mathbf{Pr}[R_1 \cap R_2] = \frac{26}{52} \cdot \frac{25}{51} \neq \mathbf{Pr}[R_1] \cdot \mathbf{Pr}[R_2]$

With replacement: $\mathbf{Pr}[R_1 \cap R_2] = \frac{26}{52} \cdot \frac{26}{52} = \mathbf{Pr}[R_1] \cdot \mathbf{Pr}[R_2]$

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

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

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

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 a revolver with six chambers, all empty. The players put two bullets into adjacent chambers of the barrel. The first player takes the gun and spins the barrel, then he puts the gun in his head and pulls the trigger and no bullet!
He gives the gun to the second player. Which would be better for the second player, to spin the barrel first, or just pull the trigger?

The Russian roulette

If player 2 spins the barrel, the probability of getting a bullet is $2/6 = 1/3$ so the probability of survival is $1 - 1/3 = 2/3$, i.e 66.66%

If he does not spin the barrel, we are conditioning to the fact that we are positioned right after one of the 4 empty chambers. Only one of the empty chambers leads to one with 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

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

Principle of deferred decisions

Not to assume that the entire set of random choices is made in advance. Rather, at each step of the process concentrate only on the random choices that are relevant to the algorithm outcome

When applicable it provides a simplified probability space to perform the probabilistic analysis.

Analyzing the Clock Solitaire game

The Clock Solitaire game: randomly shuffle a standard pack of 52 cards. Then, split the cards into 13 piles of 4 cards each; label piles as A, 2, . . . , 10, J, Q, K; take the first card from the "K" pile; take the next card from the pile "X", where X is the value of the previous card taken; repeat until:

- either all cards removed ("win")
- or you get stuck ("lose")

We want to evaluate the probability of "win".

From MR 3.5

Game termination?

The last card we take before the game ends (either winning or loosing) is a "K".

Let us assume that at iteration j we draw card X but the pile X is empty (thus the game terminates).

Let $X \neq K$ (i.e. we lose). Because pile X is empty and $X \neq K$, we must have already drawn (prior to draw j) 4 X cards. But then, we can not draw an X card at the jth iteration, a contradiction.

There is no contradiction if the last card is a "K" and all other cards have been already removed (in that case the game terminates with win).

Game win?

We win if the fourth “K” card is drawn at the 52 iteration.

Whenever we draw for the 1st, 2nd or 3rd time a“K” card, the game does not terminate because the K pile is not empty so we can continue.

When the fourth K is drawn at the 52nd iteration then all cards are removed and the game’s result is “win”

The probability of win?

According to the previous observations

$$\begin{aligned} Pr\{win\} &= Pr\{\text{4th ``K'' at the 52nd iteration}\} \\ &= \frac{\#\text{game evolutions: 52nd card} = \text{4th ``K''}}{\#\text{all game evolutions}} \end{aligned}$$

Considering all possible game evolutions is a rather naive approach since we have to count all ways to partition the 52 cards into 13 distinct piles, with an ordering on the 4 cards in each pile. This complicates the probability evaluation because of the dependence introduced by each random draw of a card.

We define another probability space that better captures the random dynamics of the game evolution.

The principle of deferred decisions

Basic idea: rather than fix (and enumerate) the entire set of potential random choices in advance, instead let the random choices unfold with the progress of the random experiment.

In this particular game at each draw any card not drawn yet is equally likely to be drawn.

A winning game corresponds to a dynamics where the first 51 random draws include 3 "K" cards exactly.

This is equivalent to draw the 4th "K" at the 52nd iteration. So we "forget" how the first 51 draws came out and focus on the 52nd draw, which must be a "K".

The probability of win

We actually have 13x4=52 distinct positions (13 piles, 4 positions each) where 52 distinct cards are placed. This gives a total of 52! different placements.

Each game evolution actually corresponds to an ordered permutation of the 52 cards.

The winning permutations are those where the 52nd card is a "K" (4 ways) and the 51 preceding cards are arbitrarily chosen (51!). Thus:

$$Pr\{win\} = \frac{4\,51!}{52!} = \frac{4}{52} = \frac{1}{13}.$$

The probability of win

We actually have 13x4=52 distinct positions (13 piles, 4 positions each) where 52 distinct cards are placed. This gives a total of 52! different placements.

Each game evolution actually corresponds to an ordered permutation of the 52 cards.

The winning permutations are those where the 52nd card is a "K" (4 ways) and the 51 preceding cards are arbitrarily chosen (51!). Thus:

$$Pr\{win\} = \frac{4\,51!}{52!} = \frac{4}{52} = \frac{1}{13}.$$

A simpler way to get the same: The probability is $\frac{1}{13}$ because of symmetry (e.g. the type of the 52nd card is random uniform among all 13 types).

The idea was to defer, i.e. first consider the last choice and then conditionally the previous ones!

Checking matrix multiplication

Problem: Given 3 square matrices ($n \times n$), A, B and C, we want to see if $A \times B = C$.
Easy solution: compute $A \times B$ and compare with C.

$n \times n$ matrix multiplication:

1. Naive algorithm: $O(n^3)$
2. Strassen (1969): $O(n^{2.81})$
3. Coppersmith-Winograd (1987): $O(n^{2.376})$
4. Vassilevska (2015): $O(n^{2.373})$

Can we check in $O(n^2)$ if $A \times B = C$?

Freivald's algorithm for checking if $A \times B = C$ (1977)

Given $n \times n$ matrices A, B, C

Freivald A, B, C
choose u.a.r. $r \in \{0,1\}^n$
if $A(Br) = Cr$ **then**
  output true
else
  output false

Choosing u.a.r. r can be done choosing independently with probability $1/2$ each of its n bits. This makes the probability of any given r $1/2^n$, and the cost of generate the vector $O(n)$.

The time complexity of Freivald's is $\Theta(n^2)$.
Notice: if $AB = C$ the algorithm yields always the correct answer.
It could be that $AB \neq C$ the algorithms may yield the wrong answer ($AB = C$) with a certain probability
(ex: with prob.=$1/2^n$, $r = (0, 0, \ldots, 0)$)

Error probability

Theorem If $AB \neq C$ then $\mathbf{Pr}\left[A(B(r)) = Cr\right] \leq \frac{1}{2}$

Proof

Neat trick: As $AB \neq C$ taking $D = AB - C$, then $D \neq (0)$.

$\Rightarrow \exists d_{ij} \in D$ s.t. $d_{ij} \neq 0$. W.l.o.g. assume $d_{11} \neq 0$.

If $\exists r$ s.t. $A(Br)) = Cr$ then $Dr = 0$.

$Dr = 0 \Rightarrow \sum_{j=1}^{n} d_{1j} r_j = 0$, but as $d_{11} \neq 0$ then

$$r_1 = \frac{-\sum_{j=2}^{n} d_{1j} r_j}{d_{11}}.$$

Second trick: Choose $r = (r_1, \ldots, r_n)$ from r_n to r_1 and stop at r_2, just before choosing r_1, which could be only 0 or 1.

Then the equality $r_1 = \frac{-\sum_{j=2}^{n} d_{1j} r_j}{d_{11}}$ holds with prob. $= 1/2$

Notice that by considering $r_n, \ldots, r_2$ to be fixed, we reduce the sample space to $r_1 \in \{0, 1\}$

Randomized algorithms and amplification

Notice Freivald's algorithm finish always in finite time ($\Theta(n^2)$) but may output the wrong answer. That type of randomized algorithms are called Monte-Carlo algorithms.
Moreover Freivald's also is a one side error, if $AB = C$ we always get the correct answer, but if $AB \neq C$ we may get the wrong answer with a "small" probability.

One-side Monte-Carlo algorithms have the nice characteristic that can be amplified: Each run of the algorithm can be considered as an independent "experiment", so they can be repeated, at each run we generate a new random choice, and by independence, each run decreases the probability of error.

If we repeat k times Freivald's algo. and each time we generate a new v, the answer keep being true, the probability of error is $\leq 1/2^k$.