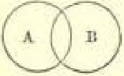
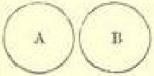
Introduction to data science & artificial intelligence (INF7100)

Arthur Charpentier

#401 Mathematics

été 2020

Mathematics

(i) Diagrammatic	(ii) Common Logic	(iii) Quantified	(iv) Symbolic
	All A is B, All B is A	All A is all B	$A\bar{B}=0$, $\bar{A}B=0$
	All A is B, Some B is not A	All A is some B	$A\bar{B}=0$, $\bar{A}B=v$
	All B is A, Some A is not B	Some A is all B	$\bar{A}B=0$, $A\bar{B}=v$
	Some A is B, Some A is not B, Some B is not A	Some A is some B	$AB=v$, $A\bar{B}=v$, $\bar{A}B=v$
	No A is B	No A is any B	$AB=0$

John Venn, Symbolic Logic, 1881.

Maths? ($\neq$ calculus)

In order to better understand

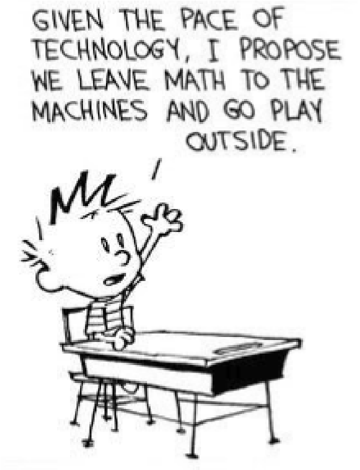
- logarithm
- derivatives, integrals
- optimisation
- vectors, matrices
- projections
- probabilities

log : multiplicative $\rightarrow$ additive

exp : additive $\rightarrow$ multiplicative

$\mathbb{E}$: average value

$\int$: sum vs. $\dfrac{\partial}{\partial x}$: difference



Maths vs. Computational (Language Issues)

"*average*" refers to the arithmetic "*mean*", the sum of the numbers divided by how many numbers are being averaged.

$$\overline{x} = \frac{x_1 + \cdots + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

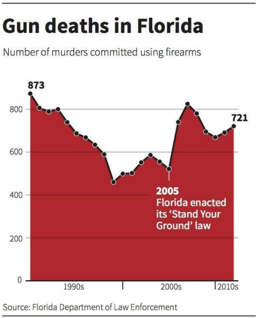
Algorithm 1: Mean (average)

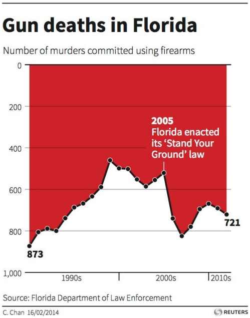
initialization : observations $x_1, \cdots, x_n$ and $s \leftarrow 0$;
for $i = 1, 2, ..., n$ **do**
  $s \leftarrow s + x_i$;
return: s/n

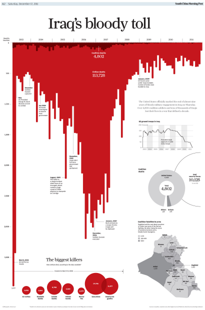
Functions, x and y (and visualization)

Graph $x \mapsto f(x)$ (plot $\{x, y = f(x)\}$)

(possibly upside down, see businessinsider)







Logarithm

For $x > 0$, $b > 0$ and $b \neq 1$,

$$\log_b(x) = y \quad \text{if} \quad b^y = x$$

In python

```
> import math
> math.log(2)
0.693147180559945З
```

and in R

```
> log(x = 2,base = exp(1))
[1] 0.6931472
> log(2)
[1] 0.6931472
```

Logarithm

For $x > 0$, $b > 0$ and $b \neq 1$,

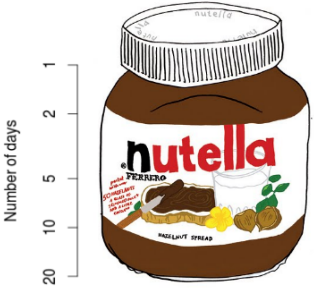
$$\log_b(x) = y \text{ if } b^y = x$$

$$\log_b(xy) = \log_b x + \log_b y,$$

$\log_b$ for $b > 1$ is the only increasing function f satisfying $f(b) = 1$ and $f(xy) = f(x) + f(y)$.

The natural logarithm of x is defined as

$$\log(x) = \int_1^x \frac{1}{t}\, dt$$



! $\log(x)$ exists only when $x > 0$ (and $\log(0)$ does not exist)

Inverse of a Function

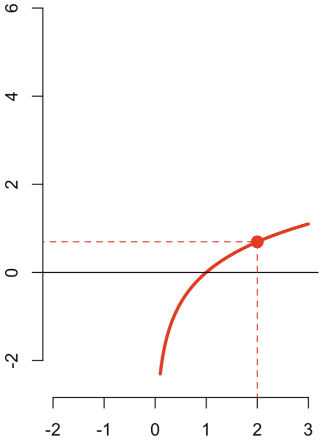
For $x > 0$, $b > 0$ and $b \neq 1$,

$$\log_b(x) = y \quad \text{if} \quad b^y = x$$

Let $f(x) = \log_b(x)$ $(= y)$.
The inverse f^{-1} is such that $x = f^{-1}(y)$, so

$$f^{-1}(y) = b^y \ (= x)$$

Intuition :
$f\big(f^{-1}(y)\big) = y$ and $f^{-1}\big(f(x)\big) = x$
$f(x) = \log(x)$, $f^{-1}(y) = \exp(y)$



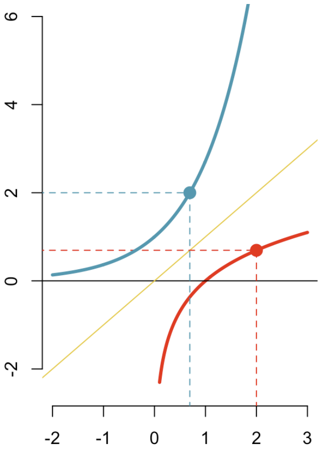
Inverse of a Function

Visually, the inverse is the symmetric with respect to the first diagonal ($y = x$).

$$\log_b(xy) = \log_b x + \log_b y$$

while

$$b^x \cdot b^y = b^{x+y}$$



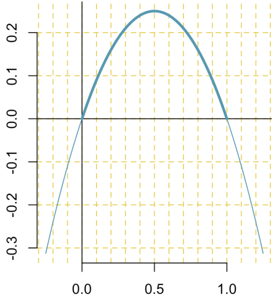
Polynomials

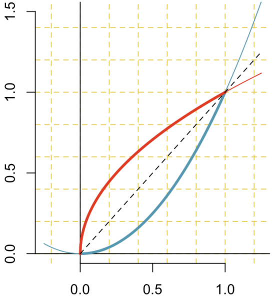
Example: $P(x) = 5x^4 + x^2 - 7x + 3$
is a polynomial of degree 4

Example: $P(x) = -x^2 + x = x \cdot (1 - x)$
is a polynomial of degree 2 (quadratic)
the graph of P is a parabola
Note: $\text{argmax}\{P(x)\} = 1/2$

Example: $x \mapsto= x^2$
Inverse of P (on $[0, \infty)$) is $x \mapsto \sqrt{x}$
i.e. if $y = x^2$ (with $x \geq 0$), $x = \sqrt{y}$

Note: $x^2 \leq x \leq \sqrt{x}$ for any $x \in [0, 1]$.
$\sqrt{x} \leq x \leq x^2$ for any $x \in [1, \infty)$.

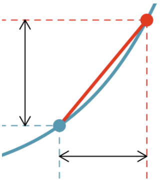




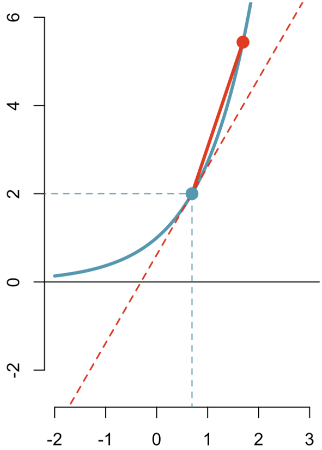
Derivative of a Function

Definition:

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$



corresponds to the limit of the slope



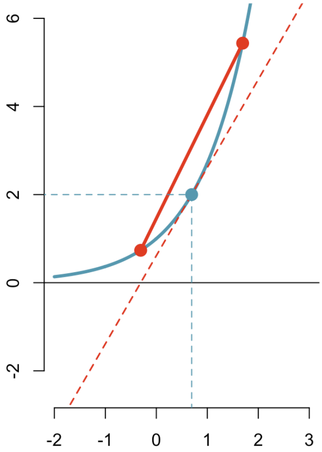
Derivative of a Function

An alternative expression is

$$f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x-h)}{2h}$$

also denoted $\dfrac{\partial f(x)}{\partial x}$

(better numerical properties,
error in h^2, against h previously)



Derivative of a Function

Standard properties

$$(f+g)' = f' + g' \text{ and } (fg)' = f'g + fg'$$

$$(\exp[g])' = g'\exp(g) \text{ and } (\log[g])' = \frac{g'}{g}$$

Chain rule $z = f(y)$ and $y = g(x)$,

$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial y} \cdot \frac{\partial y}{\partial x} = \frac{\partial f(y)}{\partial y} \cdot \frac{\partial g(x)}{\partial x} = f'(g(x)) \cdot g'(x).$$

Integral of a Function

In python

```
> import scipy.integrate as integrate
> integrate.quad(lambda x: 1/x,1,2)
(0.6931471805599454, 7.695479593116622e-15)
```

and in R

```
> integrate(function(x) 1/x,1,2)
0.6931472 with absolute error < 7.7e-15
```

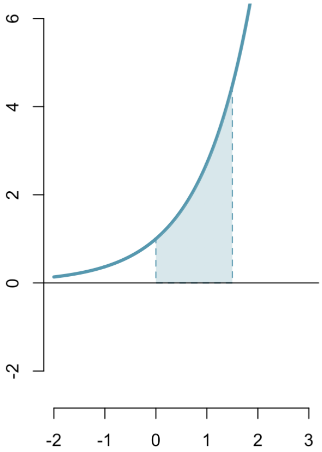
Integral of a Function

If $g(x) = f'(x)$,

$f(x) = \int_a^x g(t)dt$ for some a

$$\int_x^y f'(t)dt = f(y) - f(x)$$

Example: $\displaystyle\int_1^x \frac{1}{t}\, dt = \log(x).$

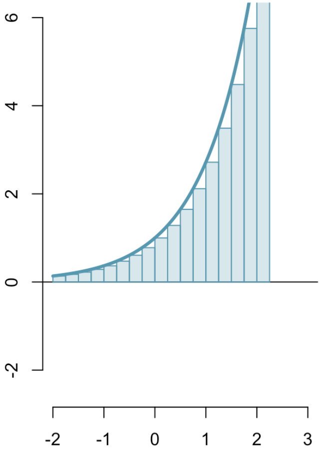


Integral of a Function

Integral $\simeq$ sum
Example
(up to Euler's constant)

$$\sum_{k=1}^{n} \frac{1}{k} \simeq \int_1^n \frac{1}{x} dx = \log(n)$$

With a log scale on x a logarithmic function is linear
With a log scale on y an exponential function is linear



Exponentials & Logarithms

$$y = \underbrace{2 \cdot 2 \cdot 2 \cdots 2}_{n \text{ times}} = 2^n$$

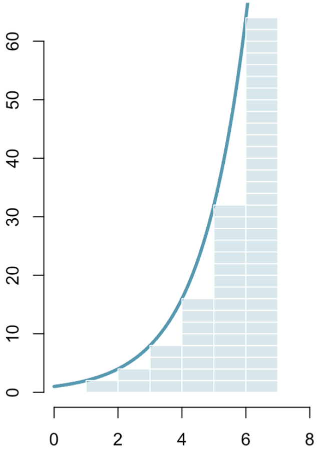
If $x_1, \cdots, x_p \in \{0,1\}$ the number of vectors $(x_1, \cdots, x_p)$ that can be generated is 2^p = number of models that can be considered with p features

$p = 10,\ 2^{10} = 1,024$

$p = 20,\ 2^{20} = 1,048,576$

$p = 40,\ 2^{40} \simeq 1,099,511,627,776$

Note : $2^{2p} = \left(2^p\right)^2$



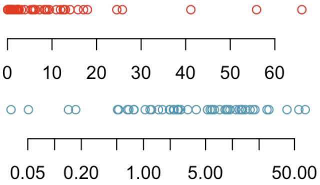
Exponentials & Logarithms

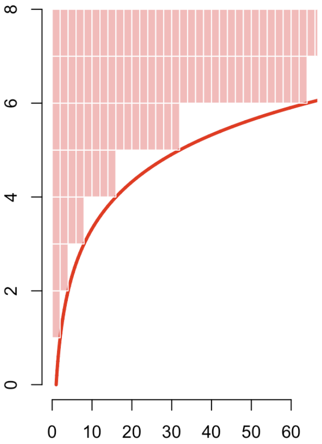
What is n so that

$$y = \underbrace{2 \cdot 2 \cdot 2 \cdots 2}_{n \text{ times}} = 2^n$$

$$n = \log_2(y)$$

see also Log Scales for visualization





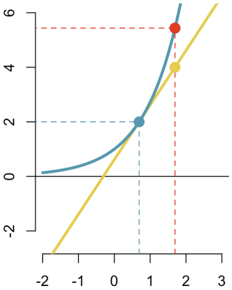
Approximation

We have seen that, if h small,

$$f'(x) \simeq \frac{f(x+h)-f(x)}{h}$$

that we can write

$$f(x+h) \simeq f(x) + \frac{f'(x)}{1} h$$



Taylor approximation (expansion),

$$f(x+h) \simeq f(x) + \frac{f'(x)}{1!} h + \frac{f''(x)}{2!} h^2 + \frac{f'''(x)}{3!} h^3 + \cdots$$

or

$$f(x) \simeq f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + \cdots$$

Counting

- how many ways to order four items $\{A, B, C, D\}$?

$$n! = n(n-1)(n-2)\cdots 3\cdot 2\cdot 1$$

(here $4! = 24$)

- how many combinations of two items out of four $\{A, B, C, D\}$?

(here (A, B), (A, C), (A, D), (B, C), (B, D) and (C, D), i.e. 6)

$$\binom{n}{k} = \frac{n!}{(n-k)!k!} = \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!}$$

for k elements out of n.

See also the birthday paradox (#131), $\displaystyle\binom{n}{2} = \frac{n(n-1)}{2} \simeq \frac{n^2}{2}$

Wrap-up

- if $b^y = x$ then $\log_b(x) = y$ (and conversely)
- $\log(xy) = \log(x) + \log(y)$ while $b^{x+y} = b^x \cdot b^y$
- $f(x) = \int^x f'(t)dt$, while $f'(x) = \left.\dfrac{\partial f(t)}{\partial t}\right|_{t=x}$
- $f(x+h) \simeq f(x) + f'(x)h$
- $\binom{n}{k} = \dfrac{n!}{(n-k)!k!} = \dfrac{n(n-1)(n-2)\cdots(n-k+1)}{k!}$