Introduction aux fonctions et statistiques

Sommaire

Courbes obtenues avec xcas en ligne et https://www.solumaths.com

I – Fonctions

1 – définitions

☞ Notion de fonction réelle : On considère un sous-ensemble I de $\mathbb{R}$. Définir une fonction f sur I, c’est associer à chaque réel de I au plus un nombre réel. Outre la courbe, le tableau, la façon la plus classique et efficace d’y parvenir est une formule définissant cette association

Exemple 1 : I = $\mathbb{R}$ et $f(x)=33$
Dans ce cas, on associe à chaque réel un unique nombre, la fonction est dite constante

Exemple 2 : I = $\mathbb{R}$ et $f(x)=\frac{1}{x+1}$
Dans ce cas le nombre associé peut être calculé pour tous les nombres réels sauf pour -1 (division par zéro interdite)

☞ Ensemble de définition de f : partie de I sur laquelle, on peut effectivement calculer $f(x)$

Exemple 3 : I = $\mathbb{R}+$ et $f(x)=\frac{1}{x-6}$
Le plus grand ensemble de définition sur I est $[0;6[\cup]6;+\infty[$

☞ Image , antécédent : on pose $f(a)=b$
b est l’image de a par f , a est un antécédent de b par f

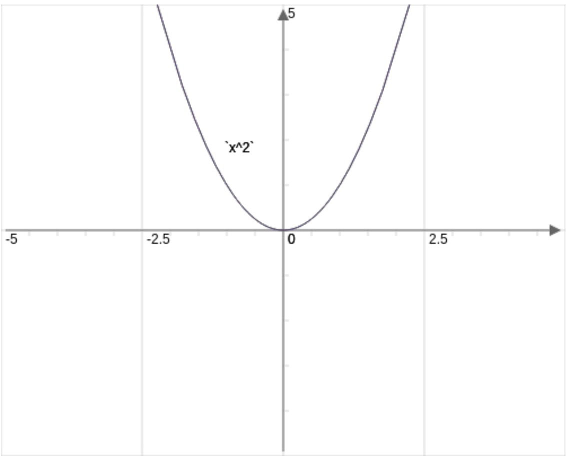
Remarque : l’image est unique mais il peut y avoir 0, 1, plusieurs ou une infinité d’antécédents

Remarque : les antécédents de b sont obtenus en résolvant l’équation $f(x)=b$ où x est l’inconnue recherchée

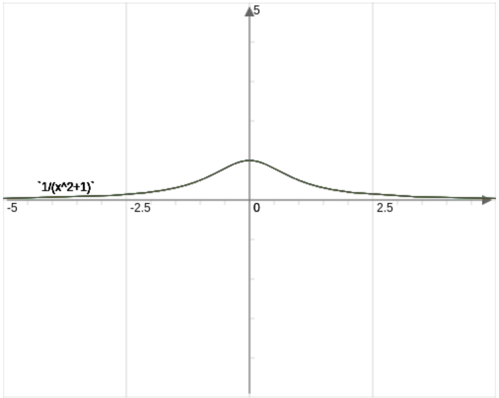
Exemple 4 : I = $\mathbb{R}$ et $f(x)=x^2+4$
2 n’admet pas d’antécédent, 4 admet un unique antécédent (0), 5 admet deux antécédents (-1 et 1)

☞ Courbe ou représentation graphique de f
On se place dans un plan muni d’un repère. La courbe de f sur I est l’ensemble des points du plan définis comme suit : $C_f=\{(x,y),x\in I,y=f(x)\}$

Exemple 5 : I = $\mathbb{R}$ et $f(x)=x^2$



Exemple 6 : I = $\mathbb{R}$ et $f(x)=\frac{1}{x^2+1}$



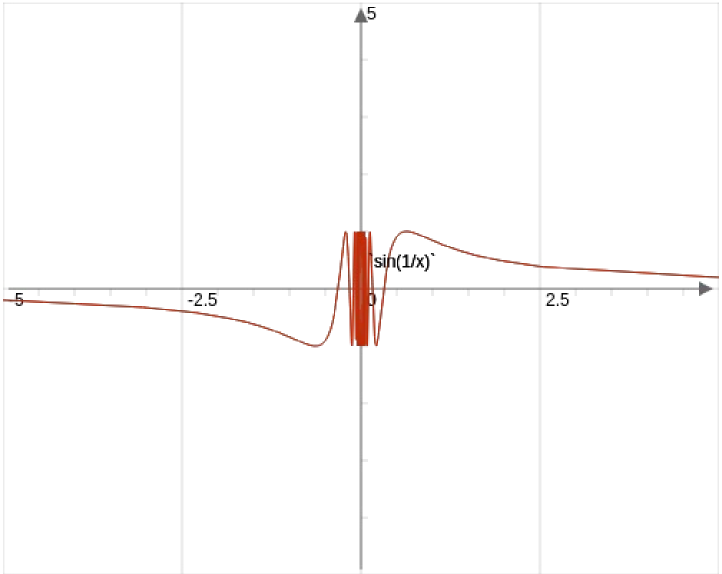
☞ Sens de variation : soit I un intervalle de $\mathbb{R}$ et a et b deux réels de I, on dira que
- f est croissante sur I si : $a<b \Rightarrow f(a) \leq f(b)$
- f est strictement croissante sur I si : $a<b \Rightarrow f(a)<f(b)$
- f est décroissante sur I si : $a<b \Rightarrow f(a) \geq f(b)$
- f est strictement décroissante sur I si : $a<b \Rightarrow f(a)>f(b)$

L’exemple précédent de la fonction définie sur I = $\mathbb{R}$ par $f(x)=\frac{1}{x^2+1}$ illustre la croissance stricte sur $\mathbb{R}-$ et la décroissance stricte sur $\mathbb{R}+$

☞ étude de la fonction f : étudier une fonction, c’est déterminer les lieux sur lesquels elle est croissante ou décroissante. Cet étude peut s’avérer complexe voire impossible. Voici quelques exemples que l’on se contentera d’évoquer

Exemple 7 : I = $\mathbb{R}$ et $f(x)$ vaut 1 si x est un nombre rationnel, 0 sinon
Dans la mesure où le sous-ensemble des rationnels et le sous-ensemble des irrationnels est dense dans $\mathbb{R}$ (tout intervalle ouvert contient au moins un rationnel ou un irrationnel), il n’existe aucun intervalle ouvert sur lequel cette fonction est croissante ou décroissante

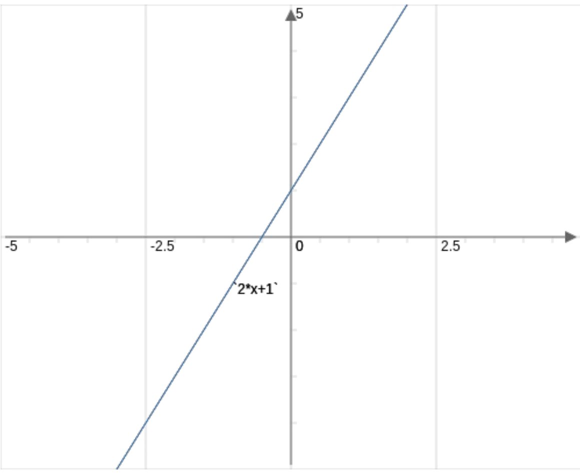
Exemple 8 : I = $]0;+\infty[$ et $f(x)=\sin(\frac{1}{x})$
Sur tout intervalle de la forme $]0;a[\ ,a>0$, la fonction n'est ni croissante , ni décroissante.
En effet, lorsque x se rapproche de 0 en étant positif, $\frac{1}{x}$ devient très grand et prend donc toutes les valeurs multiples de π qui annulent la fonction f avec changement de signe et ainsi de sens de variation



2 – fonctions usuelles

☞ fonction affine : a et b sont deux réels quelconques, I = $\mathbb{R}$, $f(x)=ax+b$
- la courbe représentative est une droite
- si $b=0$, cette fonction est dit linéaire et la courbe passe par l'origine
- le nombre a est noté coefficient directeur
- lorsque $a<0$, la fonction est strictement décroissante
- lorsque $a=0$, la fonction est constante
- lorsque $a>0$, la fonction est strictement croissante
- la courbe représentative coupe l'axe horizontale des abscisses au point d'abscisse $-\frac{b}{a}$

Exemple 9 : I = $\mathbb{R}$ et $f(x)=2x+1$



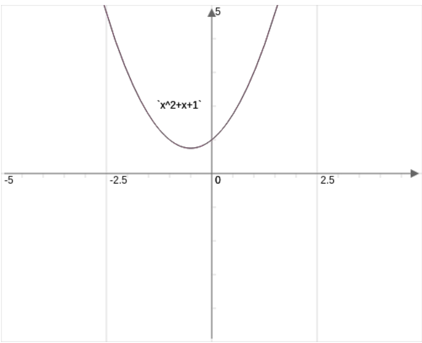
☞ fonction du second degré : $a \neq 0$, b , c sont des réels quelconques, I = $\mathbb{R}$, $f(x)=ax^2+bx+c$

- la courbe représentative est notée parabole
- lorsque $a<0$, la parabole est orientée vers le bas (forme de montagne)
- lorsque $a>0$, la parabole est orientée vers le haut (forme de vallée)
- on note discriminant associé, le nombre $\Delta=b^2-4ac$

CAS 1 : $\Delta<0$

- pas d'intersection avec l'axe des abscisses, donc pas de solution à l'équation $f(x)=ax^2+bx+c=0$
- autrement dit, $f(x)$ ne change pas de signe et la courbe se situe d'un côté de l'axe des abscisses

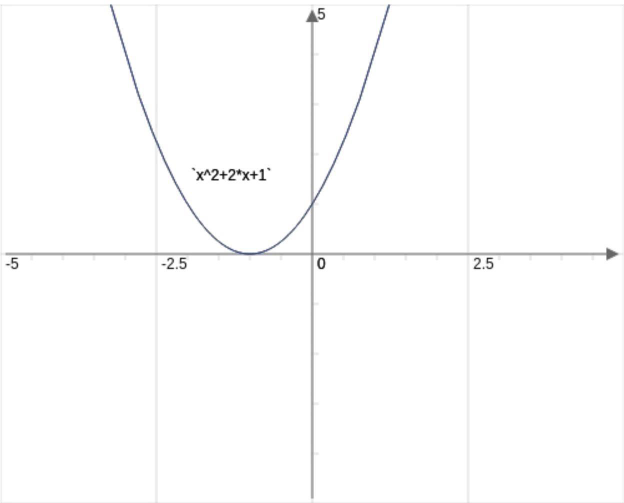
Exemple 10 : I = $\mathbb{R}$ et $f(x)=x^2+x+1$



CAS 2 : $\Delta=0$

- une seule intersection avec l'axe des abscisses, donc une seule solution à l'équation $f(x)=ax^2+bx+c=0$, définie par $x_0=-\frac{b}{2a}$
- autrement dit, $f(x)$ s'annule sans changer de signe et la courbe se situe d'un côté de l'axe des abscisses

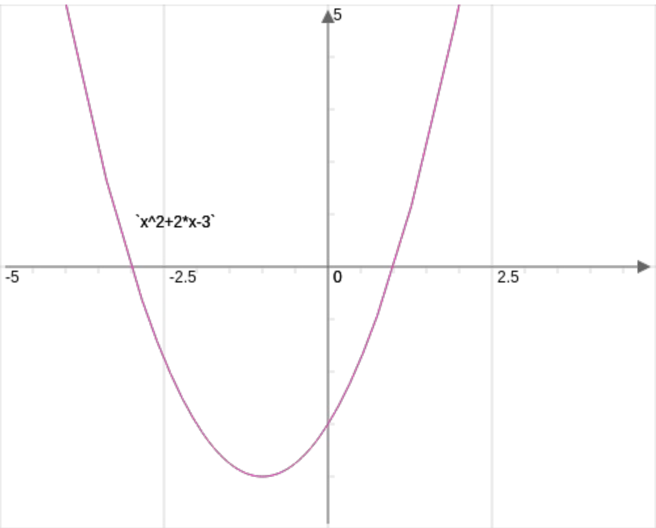
Exemple 11 : I = $\mathbb{R}$ et $f(x)=x^2+2x+1$



CAS 3 : $\Delta>0$

- deux intersections avec l’axe des abscisses, donc deux solutions à l’équation $f(x)=ax^2+bx+c=0$, définies par $x_1=\frac{-b-\sqrt{\Delta}}{2a}$ et $x_2=\frac{-b+\sqrt{\Delta}}{2a}$
- autrement dit, $f(x)$ s’annule en changeant de signe autour de x_1 et x_2 et la courbe coupe l’axe des abscisses en deux points

Exemple 12 : I = $\mathbb{R}$ et $f(x)=x^2+2x-3$



☞ monômes de degré n (entier naturel strictement positif) : I = $\mathbb{R}$, $f(x)=x^n$

- le comportement de cette fonction aux bornes de son intervalle de définition dépend de la parité de n

CAS 1 : n pair

on admettra que $f(x)$ tend vers $+\infty$ quand x tend vers $-\infty$ ou $+\infty$

CAS 2 : n impair

on admettra que $f(x)$ tend vers $-\infty$ quand x tend vers $-\infty$ et vers $+\infty$ quand x tend vers $+\infty$

☞ fonctions trigonométriques : , I = $\mathbb{R}$, $f(x)=\sin(x)$ ou $f(x)=\cos(x)$
- ces fonctions sont périodiques de période 2π
- dans tous les cas, $|f(x)|\leq 1$

☞ fonction exponentielle : I = $\mathbb{R}$, $f(x)=\exp(x)$

☞ fonction logarithme népérien : I = $\mathbb{R}$, $f(x)=\ln(x)$

Ces deux dernières fonctions qui sont liées seront étudiées explicitement

3 – notion de dérivée

Contexte : dans le cadre de l'étude des fonctions, on définit la notion de dérivée qui permet de relier le sens de variation d'une fonction f définie sur une partie I de $\mathbb{R}$ au signe d'une nouvelle fonction notée f '

☞ Notion de dérivée au point A($x_0;f(x_0)$) de la courbe de f : on associe à x_0 le coefficient directeur (noté nombre dérivé $f'(x_0)$) de la droite tangente (non verticale) à la courbe au point A. On dira que la fonction f est dérivable en x_0 . Lorsque cette association peut être réalisée sur I, on définit ainsi une nouvelle fonction, notée fonction dérivée de f

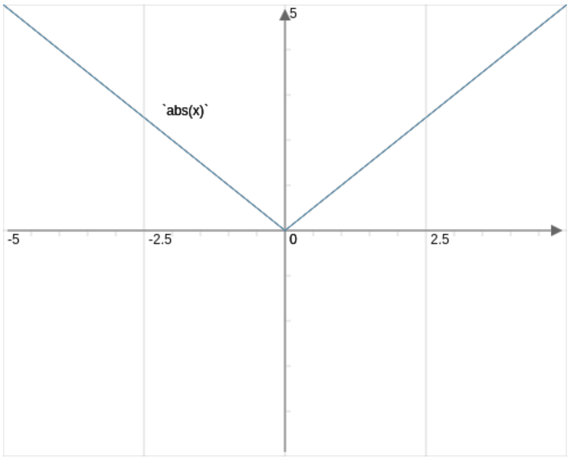
Remarque : par définition la tangente en A est une droite qui épouse la forme de la courbe en A

Remarque : il existe des cas où cette correspondance ne peut pas avoir lieu

Exemple 13 : I = $\mathbb{R}+$ et $f(x)=\sqrt{x}$

Cette fonction est définie en 0 mais n'est pas dérivable en 0, la tangente à la courbe en 0 est une droite verticale

Exemple 14 : I = $\mathbb{R}$ et $f(x)=|x|$



Cette fonction est définie en 0 mais n'est pas dérivable en 0, le point A(0 ; 0) est un point dit anguleux, il n'existe pas de droite épousant la forme de la courbe en A

☞ Condition d'existence de la dérivée : sauf cas particulier, les fonctions fabriquées à partir des fonctions usuelles (par addition, multiplication, division, composition) sont généralement dérivables sur leur ensemble de définition

☞ Rappel : composition de fonction (opérateur "rond"), $(f \circ g)(x)=f(g(x))$

Exemple 15 : $f(x)=\frac{2x}{x^2+1}$; $g(x)=x+1$; $(f \circ g)(x)=\frac{2(x+1)}{(x+1)^2+1}$; $(g \circ f)(x)=\frac{2x}{x^2+1}+1$

☞ Tableau des dérivées usuelles (1)

$f(x)$	$f'(x)$
1	0
x	1
x^2	$2x$
x^n ; $n \in \mathbb{Q}$	$n x^{n-1}$
$\exp(x)$	$\exp(x)$
$x>0$; $\ln(x)$	$\frac{1}{x}$

☞ Tableau des règles de dérivation (2)

$f(x)$	$f'(x)$
$u(x)+v(x)$	$u'(x)+v'(x)$
$k \in \mathbb{R}$; $k \times u(x)$	$k \times u'(x)$
$\frac{1}{v(x)}$	$-\frac{v'(x)}{v^2(x)}$

$u(x)\cdot v(x)$	$u'(x)\cdot v(x)+u(x)\cdot v'(x)$
$\dfrac{u(x)}{v(x)}$	$\dfrac{u'(x)\cdot v(x)-u(x)\cdot v'(x)}{v^2(x)}$
$\exp(v(x))$	$v'(x)\cdot\exp(v(x))$
$v(x)>0$; $\ln(v(x))$	$\dfrac{v'(x)}{v(x)}$
$(u\circ v)(x)$	$u'(v(x))\cdot v'(x)$

Exemple 16 : $f(x)=5x^3-4x^2+6x+45$; $f'(x)=15x^2-8x+6$

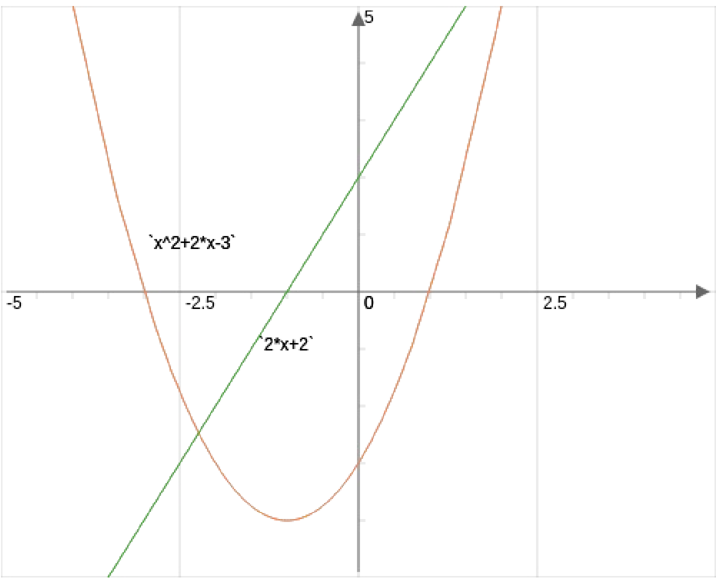
Exemple 17A : $f(x)=\ln(x^2+1)$; $f'(x)=\dfrac{2x}{x^2+1}$ (voir $\ln(v(x))$ du tableau 2)

Exemple 17B : $f(x)=\exp(-5x)$; $f'(x)=-5\exp(-5x)$ (voir $\exp(v(x))$ du tableau 2)

Exemple 18A : $f(x)=x\ln(x)$; $f'(x)=1\cdot\ln(x)+x\cdot\dfrac{1}{x}=\ln(x)+1$ (voir $u(x)\cdot v(x)$ du tableau 2)

Exemple 18B : $f(x)=\dfrac{3x+10}{-x+6}$; $f'(x)=\dfrac{3(-x+6)-(3x+10)(-1)}{(-x+6)^2}=\dfrac{28}{(-x+6)^2}$ (voir $\dfrac{u(x)}{v(x)}$ du tableau 2)

Exemple 19 : illustration du lien entre le sens de variation de la fonction et le signe de la dérivée, $f(x)=x^2+2x-3$; $f'(x)=2x+2$



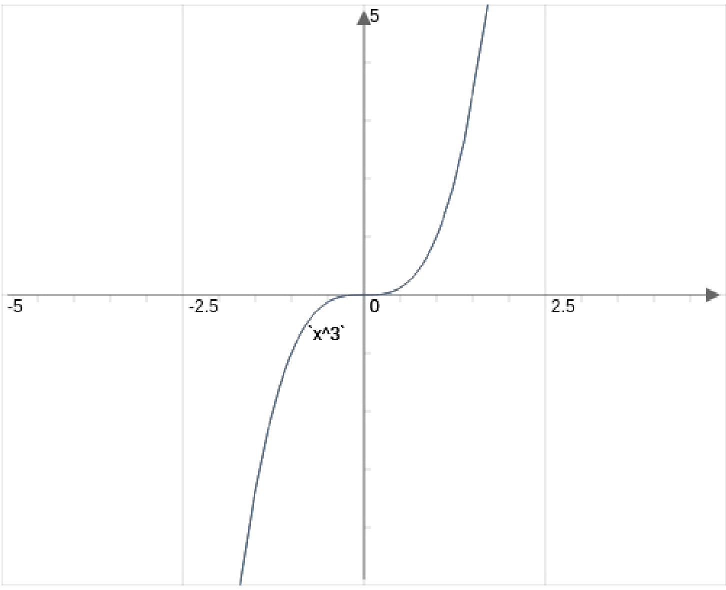
Afin d’étudier les fonctions, on s’appuiera sur les théorèmes suivants :

☞ TH1 : on suppose que f est dérivable sur un intervalle I
CAS 1 : si pour tout x de I, $f'(x)\geq 0$, on suppose que f' peut s’annuler uniquement en

des point isolés, alors f est strictement croissante sur I
CAS 2 : si pour tout x de I, $f'(x)\leq 0$, on suppose que f' peut s'annuler uniquement en des point isolés, alors f est strictement décroissante sur I
CAS 3 : si f est nulle pour tout x de I, alors f est constante sur I

☞ TH2 : on suppose que f est dérivable sur un intervalle ouvert I
- Si f admet un extremum (minimum ou maximum) en a , alors $f'(a)=0$
- (RÉCIPROQUE) si $f'(a)=0$ **avec** un changement de signe de f' de part et d'autre de a , alors f admet un extremum en a

Exemple 20-A à méditer : $f(x)=x^3$; $f'(x)=3x^2$



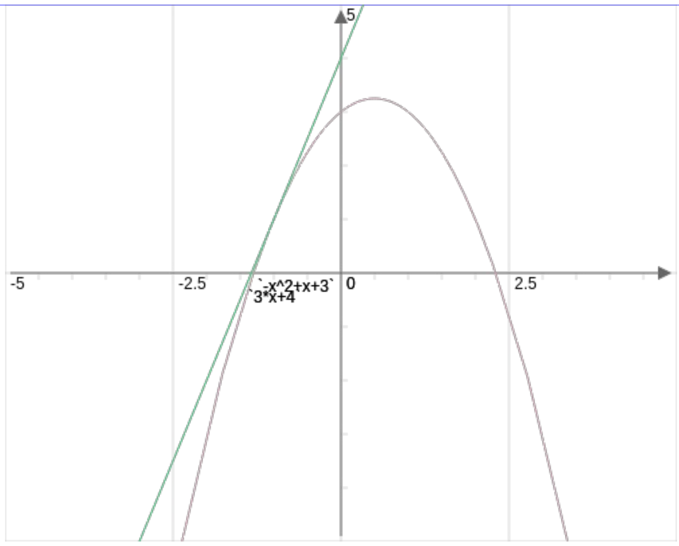
La dérivée s'annule en 0, il y a bien une tangente horizontale en 0, mais la fonction reste strictement croissante sur $\mathbb{R}$. Tout s'explique en analysant le signe de la dérivée, il est évidemment positif

☞ équation de la tangente
Soit f une fonction dérivable en x_0 . La tangente au point $A(x_0, f(x_0))$ vérifie les deux conditions suivantes :

- le coefficient directeur est égal à $f'(x_0)$
- le point A appartient à la tangente

Une équation de cette tangente est donc : $y=f'(x_0)\cdot(x-x_0)+f(x_0)$

Exemple 20-B : $f(x)=-x^2+x+3$
Recherche de la tangente au point d'abscisse $x_0=-1$
$f'(x)=-2x+1$
$f(x_0)=-(-1)^2+(-1)+3=1$
$f'(x_0)=-2\cdot(-1)+1=3$
équation de la tangente : $y=3(x-(-1))+1=3x+4$



Application directe : pour une valeur de x proche de x_0 , on approxime la valeur de $f(x)$ par l'expression de la tangente en x_0 , c'est-à-dire $f'(x_0)\cdot(x-x_0)+f(x_0)$

Exemple 20-C : on recherche une valeur approchée de $A=\frac{1}{1.017}$

On pose $f(x)=\frac{1}{1+x}$, on a $A=f(0.017)$. Déterminons l'expression de la tangente en $x_0=0$

$f'(x)=-\frac{1}{(1+x)^2}$, $f(0)=1$, $f'(0)=-1$, équation de la tangente en 0 : $y=-x+1$.

On déduit $A\approx-0.017+1=0.983$, un calcul exact à 10^{-5} près donne $A\approx0.98328$

4 – dérivée d'ordre supérieur à 1

☞ La dérivée nième (notée $f^{(n)}(x)$) ou d'ordre n de f est obtenue en dérivant f , n fois

Exemple 21 : $f(x)=5x^3-4x^2+6x+45$; $f'(x)=15x^2-8x+6$; $f''(x)=30x-8$; $f'''(x)=f^{(3)}(x)=30$

☞ Point d'inflexion : on considère une fonction f admettant des dérivées de tout ordre sur un intervalle ouvert I, on note C la courbe de cette fonction. Un point d'inflexion de C est un point particulier où la concavité de la courbe change et on admet que la courbe C traverse sa tangente en ce point

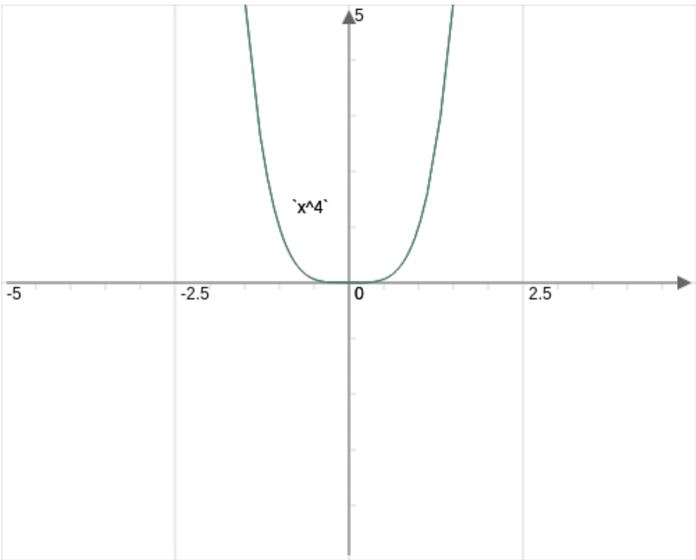
Soit x_0 de I, on note le point A($x_0;f(x_0)$) de C.
Condition nécessaire et suffisante d'existence d'un point d'inflexion en A :

- $f''(x_0)=0$
- k est impair, où k est le premier entier naturel supérieur strictement à 2 tel que $f^{(k)}(x_0)\neq0$

Exemple 22 : $f(x)=x^3$; $f'(x)=3x^2$; $f''(x)=6x$; $f^{(3)}(x)=6$

A(0 ; 0) est un point d'inflexion de C dessinée précédemment (exemple 20-A)

Exemple 23-A : $f(x)=x^4$; $f'(x)=4x^3$; $f''(x)=12x^2$; $f^{(3)}(x)=24x$; $f^{(4)}(x)=24$
on est dans le cas où $f''(0)=0$ mais k = 4, A(0 ; 0) n'est pas un point d'inflexion



Remarque : Tout polynôme de degré 3 admet un unique point d'inflexion. En effet, la dérivée seconde est une fonction affine s'annulant une seule fois, et de plus la dérivée troisième est une fonction constante non nulle

Remarque : lien entre extremum et signe de f''
Soit f une fonction deux fois dérivable sur un intervalle ouvert I, sur lequel :

- $x_0 \in I$ et $f'(x_0)=0$
- f'' ne s'annule pas sur I et garde un signe constant

CAS 1 : $f''(x_0)<0$, f' est donc décroissante, s'annulant en x_0 . f' est positive à gauche de x_0 et négative à droite, on en déduit que f admet un maximum en x_0
CAS 2 : $f''(x_0)>0$, par un raisonnement analogue, on montre que f admet un minimum en x_0
Cette méthode peut être appliquée lorsque la détermination de f'' est simple

Exemple 23-B : $f(x)=3x^2+x+7$; $f'(x)=6x+1$; $f''(x)=6$
Vu le signe positif constant de f'' , toute valeur qui annule la dérivée correspondant à un minimum de la fonction f

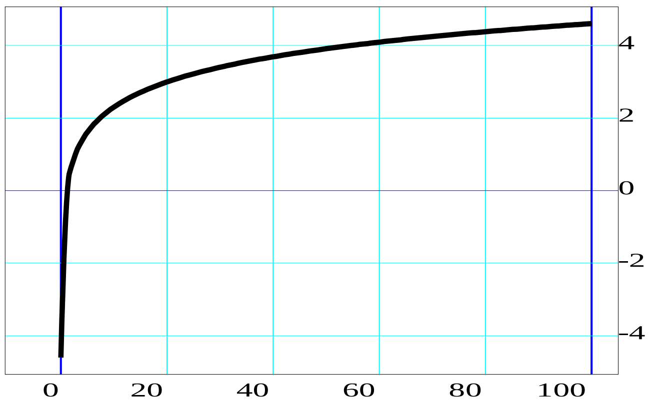
5 – fonction logarithme népérien ln

☞ Définition : on admet qu'il existe une unique fonction dérivable et définie sur I = $]0;+\infty[$, notée $\ln$ vérifiant les deux conditions suivantes

- $\ln'(x)=\dfrac{1}{x}$
- $\ln(1)=0$

☞ propriétés admises :

- $\ln$ est strictement croissante sur I
- $\lim_{x\to 0, x>0} \ln(x) = -\infty$
- $\lim_{x\to +\infty} \ln(x) = +\infty$



☞ propriété caractéristique fondamentale

a et b sont deux réels strictement positifs quelconques : $\ln(a\cdot b) = \ln(a) + \ln(b)$

Autrement dit : $\ln$ "transforme" les produits en sommes

Remarque : ébauche de démonstration

a est un réel fixé et on définit sur I la fonction $g(x) = \ln(a\cdot x) - \ln(a)$

- on a immédiatement $g(1) = 0$
- par application des règles de dérivation, on obtient $g'(x) = \frac{1}{x}$

Par unicité (admise lors de la définition de $\ln$) de la fonction $\ln$, on a $g(x) = \ln(x)$ et on en déduit $\ln(a\cdot x) - \ln(a) = \ln(x) \Rightarrow \ln(a\cdot x) = \ln(a) + \ln(x)$ ■

☞ propriétés induites

- $\ln\left(\frac{1}{b}\right) = -\ln(b)$
- $\ln\left(\frac{a}{b}\right) = \ln(a) - \ln(b)$
- $\ln(a^r) = r\cdot\ln(a)$; $r \in \mathbb{Q}$

☞ Croissances comparées (admises) aux bornes de l'intervalle de définition

$\lim_{x\to 0, x>0} x\cdot\ln(x) = 0$, $\lim_{x\to +\infty} \frac{\ln(x)}{x} = 0$

Exemple 24 : $\ln(40) = \ln(8\times 5) = \ln(8) + \ln(5) = \ln(2^3) + \ln(5) = 3\ln(2) + \ln(5)$

Exemple 25 : au début d'une année que l'on notera 0, Riop dépose à la banque YAUNDOOT une somme notée C_0 qui doit rapporter des intérêts composés annuels à un taux en pourcentage noté t. Riop souhaite savoir au bout de quelle année n, ce capital aura triplé mais il ne souvient plus si son conseiller lui avait dit t = 0.8 % ou 8 %.
On admet (propriétés des suites géométriques) que le capital au bout de l'année n est $C_n=C_0(1+\frac{t}{100})^n$. On recherche la première valeur de n pour laquelle $C_n \geq 3C_0$

Ce qui donne : $C_0(1+\frac{t}{100})^n \geq 3C_0$

par simplification : $(1+\frac{t}{100})^n \geq 3$

par croissance de la fonction $\ln$: $\ln((1+\frac{t}{100})^n) \geq \ln(3)$

par propriété de ln : $n\ln(1+\frac{t}{100}) \geq \ln(3)$

évidemment $1+\frac{t}{100}>0$, d'où : $n \geq \frac{\ln(3)}{\ln(1+\frac{t}{100})}$

Application numérique pour t = 0.8 % : n = 138 ; Riop devrait se poser la question de la pertinence d'un tel placement

Application numérique pour t = 8 % : n = 15 ; Riop devrait se méfier d'un taux aussi généreux

6 – fonction exponentielle exp

☞ Lien entre $\ln$ et $\exp$: faisant suite à l'étude précédente de la fonction $\ln$ aux bornes de son intervalle de définition et de son sens de variation, on admet que la fonction $\ln$ prend TOUTE valeur réelle une unique fois, ce qui peut s'écrire de la façon suivante

Pour tout réel y , il existe un unique réel x strictement positif tel que : $y=\ln(x)$

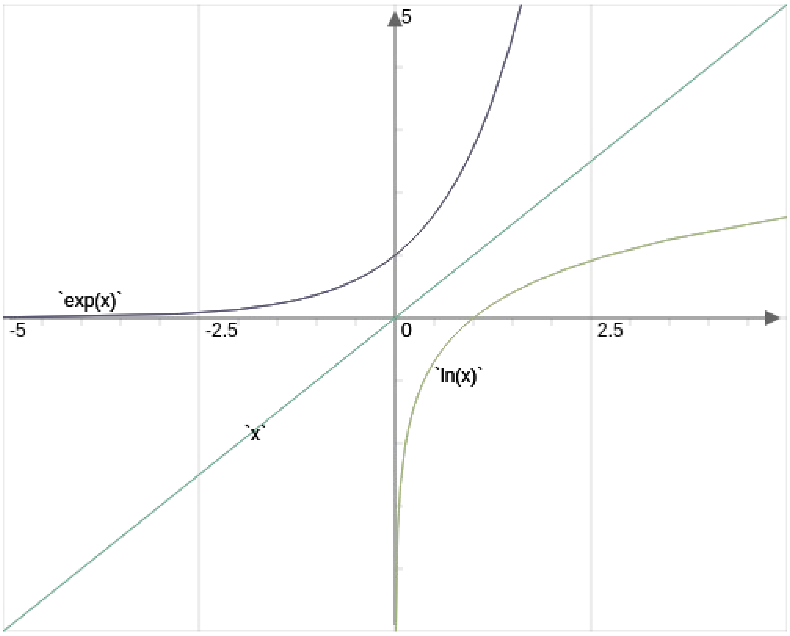
La correspondance de x vers y est définie par la fonction $\ln$ et la correspondance de y vers x décrit une nouvelle fonction qui est par définition la fonction RÉCIPROQUE de la fonction $\ln$, autrement dit : $x>0, y\in\mathbb{R}, y=\ln(x) \Leftrightarrow x=\exp(y)$

Remarque : la notion de continuité non abordée ici permettrait de justifier la construction de $\exp$

☞ propriétés induites et admises :

- $x>0, \exp(\ln(x))=x$
- $y\in\mathbb{R}, \ln(\exp(y))=y$
- $\exp$ est strictement croissante sur $\mathbb{R}$
- $\lim_{y\to-\infty} \exp(y)=0$
- $\lim_{y\to+\infty} \exp(y)=+\infty$
- $y\in\mathbb{R}, \exp(y)>0$

- les courbes de $\ln$ et $\exp$ sont symétriques par rapport à la droite d'équation $y=x$ (première bissectrice)
- $\exp(0)=1$, ceci découle de $\ln(1)=0$



Remarque : dorénavant, on définira la nouvelle fonction $\exp$ à l'aide de la notation classique $\exp(x)$

☞ Définition : on pose $\exp(1)=e$. Il en découle $\ln(e)=1$

☞ On admet que la fonction $\exp$ est dérivable sur $\mathbb{R}$ et $(\exp(x))'=\exp(x)$

☞ propriété caractéristique fondamentale
a et b sont deux réels quelconques : $\exp(a+b)=\exp(a)\cdot\exp(b)$

Autrement dit : $\exp$ "transforme" les sommes en produits

Remarque : ébauche de démonstration
$a\in\mathbb{R}, b\in\mathbb{R}, \ln(\exp(a+b))=a+b=\ln(\exp(a))+\ln(\exp(b))=\ln(\exp(a)\cdot\exp(b))$
d'où $\exp(a+b)=\exp(a)\cdot\exp(b)$ ■

☞ propriétés induites

- $\exp(-b)=\dfrac{1}{\exp(b)}$
- $\exp(a-b)=\dfrac{\exp(a)}{\exp(b)}$
- $\exp(r\cdot a)=(\exp(a))^r$; $r\in\mathbb{Q}$

☞ Nouvelle expression de $\exp$:
En posant $a=1$ dans la dernière propriété, on a $\exp(r\cdot 1)=\exp(r)=(\exp(1))^r=e^r$

Par des considérations de continuité et de densité non abordées ici, on étend cette définition sur $\mathbb{R}$, ce qui donne : $x \in \mathbb{R}$, $\exp(x) = e^x$

☞ Extension de la fonction exponentielle

On pose : $a > 0$, $a^x = e^{x \ln(a)}$

☞ Croissances comparées (admises) aux bornes de l'intervalle de définition

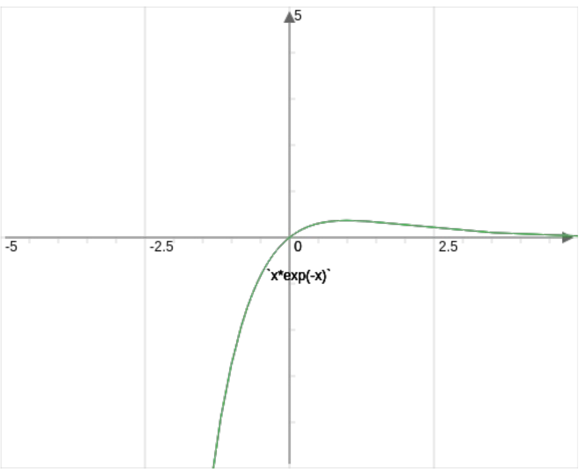
$\lim_{x \to -\infty} x e^x = 0$, $\lim_{x \to +\infty} \frac{e^x}{x} = +\infty$

voici quelques applications immédiates :

Exemple 26 : Résolution de l'équation $e^{x+4} = 10$
on a $x+4 = \ln(10)$ d'où $x = \ln(10) - 4$, ainsi $S = \{ \ln(10) - 4 \}$

Exemple 27 : Résolution de l'équation $\ln(2x+6) = 100$
On doit s'assurer que $2x+6 > 0$ ou encore $x > -3$
on a $2x+6 = \exp(100)$ d'où $x = \frac{e^{100} - 6}{2}$, ainsi $S = \{ \frac{e^{100} - 6}{2} \}$

Exemple 28 : étude de la fonction $f(x) = x e^{-x}$, $x \geq 0$
on a $f'(x) = 1 \cdot e^{-x} + x \cdot (-e^{-x}) = (1-x) e^{-x}$
QUESTION : $f'(x) > 0$?
$(1-x) e^{-x} > 0$, or $e^{-x} > 0$, d'où $1 - x > 0$
RÉPONSE : $0 \leq x < 1$
f est strictement croissante sur [0 ; 1]
f est strictement croissante sur [1 ; $+\infty$ [
$f(0) = 0, f(1) = e^{-1} = \frac{1}{e}$, , $\lim_{x \to +\infty} f(x) = 0$



7 – Applications en économie gestion

☞ Rappel : taux d'accroissement entre a et b pour une fonction f

Ce taux est défini par $\frac{f(b)-f(a)}{b-a}$. On admet les points suivants :

- si f est dérivable en a , par définition, ce taux tend vers $f'(a)$ lorsque b tend vers a
- dans le cas d'une fonction affine, ce taux correspond au coefficient directeur de la droite représentative de la fonction

☞ Rappel : taux d'évolution (%) entre V_I (valeur initiale) et V_F (valeur finale) d'une grandeur

Ce taux est définie par $100\frac{V_F-V_I}{V_I}$

Dans cette partie on se place dans le contexte d'une entreprise fabriquant et vendant un produit. On note x la quantité de produits fabriqués sur une période donnée T.

- ce produit peut être fabriqué à l'unité et dans ce cas x prend des valeurs entières (0, 1, 2, …)
- ou alors x prend toute valeur dans un intervalle donné (à l'unité de mesure pouvant être une longueur, une masse, …)
- dans les deux cas, on pose p le prix de vente à l'unité ou par unité de mesure

☞ On note $C(x)$, le coût total de fabrication d'une quantité x de produits sur la période T

☞ Recette totale associée : $R(x)=p\cdot x$
Dans le cas où p dépend de x , la formule devient $R(x)=p(x)\cdot x$

☞ Bénéfice total associé : $B(x)=R(x)-C(x)$
On dira que le seuil de rentabilité est atteint lorsque $B(x)\geq 0$. À l'aide d'une étude de fonction, il est possible de déterminer ce seuil mais aussi les extrema éventuels

☞ Coût moyen : $C_M(x)=\frac{C(x)}{x}$

☞ Coût marginal : $C_m(x)$

- pour un produit se vendant à l'unité, c'est la variation de coût total pour une unité supplémentaire produite, $C_m(x)=C(x+1)-C(x)$
- sinon, on pose $C_m(x)=C'(x)$

Remarque : les deux définitions se rejoignent, car si on considère que x est grand par rapport à 1, $C(x+1)-C(x)=\frac{C(x+1)-C(x)}{x+1-x}$, c'est donc un taux d'accroissement et on applique la remarque concernant ce taux, $C(x+1)-C(x)\approx C'(x)$

☞ Recette marginale : $R_m(x)$

- pour un produit se vendant à l'unité, c'est la variation de recette totale pour une unité supplémentaire vendue, $R_m(x)=R(x+1)-R(x)$

- sinon, on pose $R_m(x)=R'(x)$

Remarque : les deux définitions se rejoignent, car si on considère que x est grand par rapport à 1, $R(x+1)-R(x)=\frac{R(x+1)-R(x)}{x+1-x}$, c'est donc un taux d'accroissement et on applique la remarque concernant ce taux, $R(x+1)-R(x)\approx R'(x)$

☞ Notion d'élasticité

On considère une fonction f de la variable x et dérivable en a . On cherche à mesurer l'influence de la variation relative de f sur celle de x et réciproquement

- entre $f(a)$ et $f(x)$ le taux d'évolution est $t_f=\frac{f(x)-f(a)}{f(a)}$
- entre a et x le taux d'évolution est $t_x=\frac{x-a}{a}$

Lorsque x est proche de a , on approxime le rapport des taux par sa limite éventuelle en a . Or $\frac{\frac{f(x)-f(a)}{f(a)}}{\frac{x-a}{a}}=\frac{a}{f(a)}\cdot\frac{f(x)-f(a)}{x-a}\rightarrow\frac{a}{f(a)}\cdot f'(a)$

On définit ainsi l'élasticité de f de la façon suivante : $e_f(a)=a\cdot\frac{f'(a)}{f(a)}\approx\frac{t_f}{t_x}$

Exercices d'application : retrouver de nombreux exercices corrigés d'application directe (optimisation du bénéfice, élasticité, …) sur www.acquizitor.net

II – Introduction aux statistiques

1 – Statistique à deux variables quantitatives

a – Rappels

☞ On s’intéresse à des caractères ou variables (notés généralement par des lettres telles que x ou y) issus des disciplines courantes (économie, démographie, biologie, …) tels que le chiffre d’affaires d’une entreprise, l’âge d’une population, la taille des arbres, le poids des animaux, le revenu des ménages, la catégorie socio-professionnelle d’une population, ...

☞ Lorsque les valeurs prises (notées $x_{1,} x_{2,} x_{3,} \cdots$) par la variable étudiée x sont numériques, la variable est dite quantitative, sinon elle est dite qualitative

Exemple 1 : x représente le chiffre d’affaires mensuel de la société SUPRASUN et s’exprime en millions d’euros, elle est quantitative

Exemple 2 : y représente la catégorie socio-professionnelle d’une population, elle est qualitative

On considère une variable quantitative x prenant n valeurs $x_{1,} x_{2,} x_{3,} \cdots, x_n$

☞ moyenne (paramètre de position) : $\bar{x}$

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

☞ variance : $var(x)$
la variance est la moyenne des carrés des écarts à la moyenne

$$var(x)=\frac{(x_1-\bar{x})^2+\ldots+(x_n-\bar{x})^2}{n}=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

☞ écart-type (paramètre de dispersion) : σ_x
l’écart-type donne une indication sur la dispersion des valeurs autour de la moyenne (plus les valeurs prises sont proches de la moyenne et plus il est petit)

$$\sigma_x=\sqrt{var(x)}$$

Remarques : l’écart-type n’est nul que dans le cas où toutes les valeurs x_i sont identiques.
L’écart-type n’est pas exactement la moyenne m des écarts à la moyenne, car cette moyenne m est systématiquement nulle. En effet : $m=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})=\frac{1}{n}((\sum_{i=1}^{n} x_i)-n\bar{x})=\frac{1}{n}(n\bar{x}-n\bar{x})=0$

b - Généralités

Exemple introductif : dans un village, on souhaite étudier le lien entre le temps moyen passé devant la télé et le temps mis pour faire le parcours de santé autour du village de 5 km. On note x la variable notant le temps moyen passé devant la télé et y celle repérant le temps mis pour faire le parcours. Lors d'une enquête, n personnes répondent en révélant leurs valeurs de x et y

On note $(x_1, y_1), \ldots, (x_n, y_n)$ ces valeurs et on suppose que les valeurs x_i sont distinctes deux à deux. On imagine un lien entre les deux variables qui se traduit par l'existence théorique d'une fonction f permettant de passer de l'une à l'autre comme suit : $y = f(x)$

Lorsque cette fonction existe :

- on dira que y est expliquée par x
- on pose $\hat{y}_i = f(x_i)$ la valeur ajustée
- on pose $y_i = \hat{y}_i + e_i$, e_i est noté résidu ou erreur

CAS fondamental : fonction affine $f(x) = ax + b$
la représentation graphique des couples $(x_i, y_i)_{1 \le i \le n}$ donne globalement une droite

CAS exponentiel : fonction de type exponentiel $f(x) = b \exp(ax)$, on suppose $b > 0$ d'où $\ln(f(x)) = \ln(b) + ax$
la représentation graphique des couples $(x_i, \ln(y_i))_{1 \le i \le n}$ donne globalement une droite

CAS logarithme : fonction logarithme népérien $f(x) = b + a \ln(x)$
la représentation graphique des couples $(\ln(x_i), y_i)_{1 \le i \le n}$ donne globalement une droite

CAS puissance : fonction polynomiale simple $f(x) = b x^a$, on suppose $b > 0, x > 0$ d'où $\ln(f(x)) = \ln(b) + a \ln(x)$
la représentation graphique des couples $(\ln(x_i), \ln(y_i))_{1 \le i \le n}$ donne globalement une droite

☞ Conclusion : quitte à appliquer la fonction $\ln$ aux valeurs des variables, on peut se ramener à l'étude du cas fondamental affine

Question importante : comment savoir simplement si on est dans le cas d'un ajustement affine ?
Réponse : représenter graphiquement le nuage des points $(x_i, y_i)_{1 \le i \le n}$ et conclure

c – Régression linéaire par la méthode des moindres carrés

Contexte : On considère la série de n valeurs $(x_1, y_1), \ldots, (x_n, y_n)$ prises par les variables x et y . On recherche la droite D passant au plus près des points $M_1(x_1, y_1), \ldots, M_n(x_n, y_n)$ regroupés sous l'appellation nuage de points. La méthode utilisée est la suivante :

Soit D une droite d'équation réduite $y = ax + b$, i étant un entier compris entre 1 et n , on note $H_i(x_i, a x_i+b)$ le projeté du point M_i sur D parallèlement à l'axe vertical des ordonnées. La distance $M_i H_i$ est égale à $|y_i-(a x_i+b)|$. On détermine la droite D qui passe au plus près des points $M_{1,}\dots, M_n$ en déterminant la valeur de a et b qui minimise l'expression suivante :

$S(a,b)=\sum_{i=1}^{n} M_i H_i^2=\sum_{i=1}^{n}(y_i-a x_i-b)^2$. S est une fonction des variables a et b

On admet que la solution éventuelle est un point critique pour la fonction de deux variables S, c'est-à-dire : $\frac{\partial S(a,b)}{\partial a}=0 ; \frac{\partial S(a,b)}{\partial b}=0$ (les dérivées partielles sont nulles au point critique éventuel). On obtient respectivement deux équations :

$\sum_{i=1}^{n} 2(y_i-a x_i-b)(-x_i)=0$ et $\sum_{i=1}^{n} 2(y_i-a x_i-b)(-1)=0$ (R1)

On pose $\bar{x}=\frac{1}{n}\sum_{i=1}^{n} x_i$ et $\bar{y}=\frac{1}{n}\sum_{i=1}^{n} y_i$. Les deux équations deviennent :

(E1) $\sum_{i=1}^{n} y_i x_i-a\sum_{i=1}^{n} x_i^2-b\sum_{i=1}^{n} x_i=0$

(E2) $\bar{y}=a\bar{x}+b \Leftrightarrow b=\bar{y}-a\bar{x}$

Remarque : la deuxième équation indique que le point moyen $(\bar{x},\bar{y})$ appartient à la droite recherchée

On injecte dans (E1) la valeur de b obtenue dans (E2), ce qui donne :

$(\sum_{i=1}^{n} y_i x_i-a\sum_{i=1}^{n} x_i^2)-(\bar{y}-a\bar{x})\cdot n\cdot\bar{x}=0$ puis $a=\frac{(\sum_{i=1}^{n} y_i x_i)-n\bar{y}\cdot\bar{x}}{(\sum_{i=1}^{n} x_i^2)-n(\bar{x})^2}$

☞ covariance de (x,y) : $cov(x,y)=\frac{1}{n}\sum_{i=1}^{n}(y_i-\bar{y})(x_i-\bar{x})$

Cette définition permet de simplifier l'expression de a . En effet :

$$cov(x,y)=\frac{1}{n}((\sum_{i=1}^{n} y_i x_i)-n\bar{y}\cdot\bar{x}-n\bar{y}\cdot\bar{x}+n\bar{y}\cdot\bar{x})=\frac{1}{n}((\sum_{i=1}^{n} y_i x_i)-n\bar{y}\cdot\bar{x})$$

$$\sigma_x^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2=\frac{1}{n}((\sum_{i=1}^{n} x_i^2)-n(\bar{x})^2)$$

On obtient : $a=\frac{cov(x,y)}{\sigma_x^2}$ et $b=\bar{y}-a\bar{x}$

Introduisons la notion de variance expliquée :
Pour tout i compris entre 1 et n , $y_i=\hat{y}_i+e_i=a x_i+b+e_i$ et $y_i-\bar{y}=y_i-\hat{y}_i+\hat{y}_i-\bar{y}$. On en déduit $(y_i-\bar{y})^2=(y_i-\hat{y}_i+\hat{y}_i-\bar{y})^2=e_i^2+2(y_i-\hat{y}_i)(\hat{y}_i-\bar{y})+(\hat{y}_i-\bar{y})^2$.
On somme toutes ces relations

$$\sum_{i=1}^{n}(y_i-\overline{y})^2=\sum_{i=1}^{n}(y_i-\hat{y}_i+\hat{y}_i-\overline{y})^2=\sum_{i=1}^{n}e_i^2+2\sum_{1}^{n}(y_i-\hat{y}_i)(\hat{y}_i-\overline{y})+\sum_{i=1}^{n}(\hat{y}_i-\overline{y})^2$$

Simplifions cette relation en appliquant les relations (R1) définies précédemment

$$\sum_{1}^{n}(y_i-\hat{y}_i)(\hat{y}_i-\overline{y})=\sum_{1}^{n}(y_i-a x_i-b)(a x_i+b-\overline{y})=a\sum_{1}^{n}(y_i-a x_i-b)x_i+(b-\overline{y})\sum_{1}^{n}(y_i-a x_i-b)=0$$

On obtient : $\sum_{i=1}^{n}(y_i-\overline{y})^2=\sum_{i=1}^{n}e_i^2+\sum_{i=1}^{n}(\hat{y}_i-\overline{y})^2$

Il est important de noter que par construction (rappel : $\overline{y}=a\overline{x}+b$) : $\overline{e}=0$ et $\overline{\hat{y}}=\overline{y}$

on en déduit : $var(y)=\frac{1}{n}\sum_{i=1}^{n}(y_i-\overline{y})^2=var(e)+var(\hat{y})$

☞ $var(e)$: variance résiduelle, $var(\hat{y})$: variance expliquée par la régression linéaire $\hat{y}$

☞ coefficient de corrélation entre x et y : $\rho_{xy}=\frac{cov(x,y)}{\sigma_x\sigma_y}$

Propriétés admises :

- $-1\leq\rho_{xy}\leq 1$
- le signe du coefficient de corrélation est lié au sens de variation de la droite de régression, en effet ρ_{xy} et a sont de même signe
- ce coefficient mesure la dépendance linéaire entre x et y , lorsque ρ_{xy} est nul ou proche de zéro, il n’y a pas de dépendance linéaire (il peut alors exister une autre dépendance, mais pas linéaire), lorsque ρ_{xy} est égal à 1 ou -1, la corrélation linéaire est parfaite. Hormis ces cas extrêmes, l’interprétation de ρ_{xy} peut s’avérer délicate.

Exercices d’application : retrouver de nombreux exercices corrigés d’application directe (moyenne, écart-type, régression linéaire, …) sur www.acquizitor.net

Que faire de la droite de régression ?

- Mesurer l’écart $(y_i-f(x_i))_{1\leq i\leq n}$ entre la réalité et le modèle obtenu
- faire des prévisions de valeurs en x ou y

Outil de calcul préconisé pour déterminer cette droite de régression : calculatrice ou tableur

Exemple 3 : on considère le tableau des valeurs suivant

x_i	34	36	38	40	41
y_i	45.2	47.8	53.4	57	57 .8

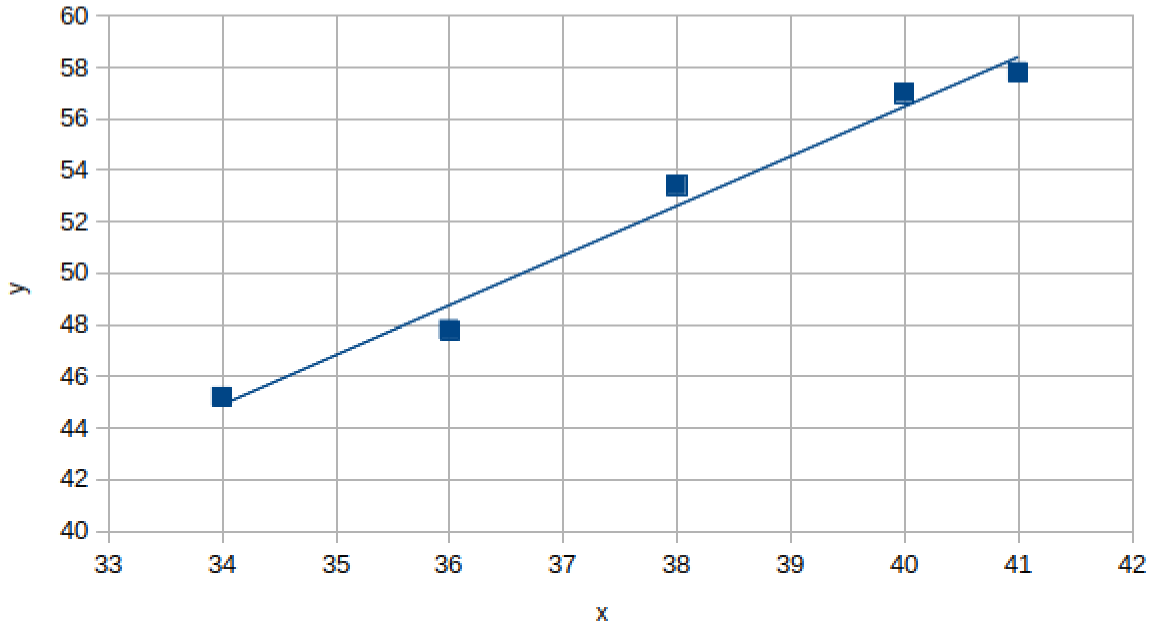
Équation de la droite de régression : $y=ax+b$

$a\approx 1.93$

$b \approx -20.64$

estimation de y pour $x=54$, $y \approx 84$

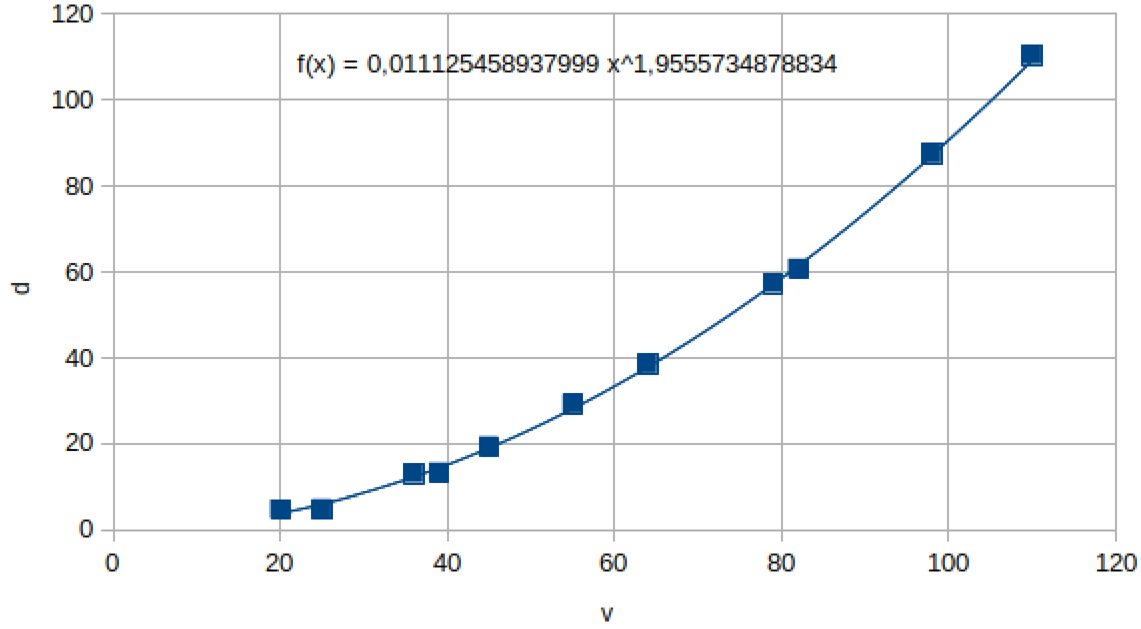
Nuage de points et droite de régression associée :

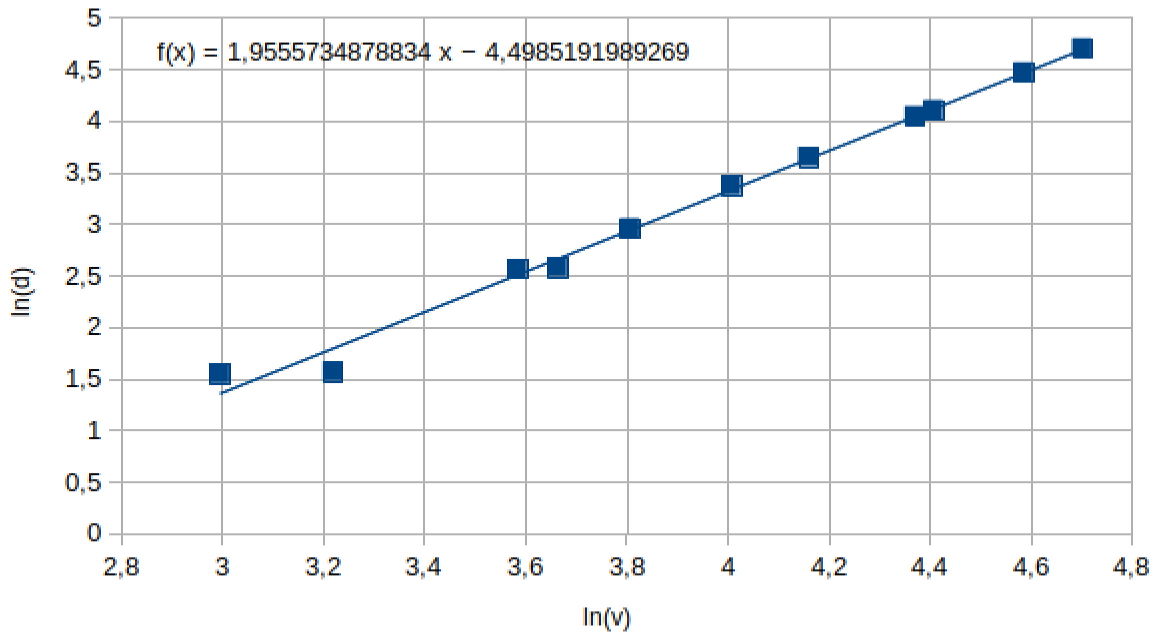


Exemple 4 : lien entre la vitesse et la distance de freinage d'un véhicule donné

v : vitesse en km/h	d : distance de freinage en m	ln(v)	ln(d)
20	4,7	2,9957322736	1,5475625087
25	4,8	3,2188758249	1,5686159179
36	13	3,5835189385	2,5649493575
39	13,2	3,6635616461	2,5802168296
45	19,3	3,8066624898	2,9601050959
55	29,2	4,0073331852	3,3741687093
64	38,5	4,1588830834	3,6506582413
79	57,3	4,3694478525	4,0483006237
82	60,7	4,4067192473	4,1059436981
98	87,5	4,5849674787	4,4716387934
110	110,4	4,7004803658	4,7041101338

La représentation des points (v_i , d_i) suggère que le modèle de régression linéaire n'est pas adapté. La représentation des points ($\ln(v_i)$, $\ln(d_i)$) suggère un modèle de régression de type puissance.
Sur chaque représentation, la fonction f(x) désigne l'expression de la régression appliquée





2 – Série temporelles ou chronologiques

☞ Une série chronologique notée $x(t)$ ou x_t est une série de valeurs prises à des intervalles de temps réguliers par une grandeur quantitative (consommation mensuelle d'électricité pour une région donnée, nombre trimestriel de naissance, nombre mensuel de vente de glace, nombre annuel de vente de voitures, …). $x(t)$ est la valeur prise par la grandeur étudiée à l'instant t, t est un indice de temps exprimant des mois, trimestres, semestres, années, … . On note n le nombre total de valeurs prises $(x_{1,}x_2\ldots,x_n)$

De manière générale, le signal $x(t)$ se décompose de la façon suivante :

- une tendance $m(t)$ ou m_t , c'est l'évolution à long terme
- composante périodique $s(t)$ ou s_t repérant en général un phénomène saisonnier
- composante résiduelle $e(t)$ ou e_t repérant des fluctuations irrégulières ou

accidentelles

Remarque : dans certains cas, une composante cyclique décelable sur plusieurs années et à durée irrégulière peut être ajoutée à la tendance

Le modèle de décomposition est additif lorsque : $x(t)=m(t)+s(t)+e(t)$ et c'est ce modèle qui sera principalement étudié.
Le modèle est multiplicatif lorsque $x(t)=m(t)\cdot s(t)+e(t)$ (1) ou $x(t)=m(t)\cdot s(t)\cdot e(t)$ (2)

Remarque : dans le dernier cas, si toutes les composantes sont strictement positives, on peut passer à un modèle additif à l'aide de l'opération $\ln(x(t))=\ln(m(t))+\ln(s(t))+\ln(e(t))$

Nous admettrons dans un premier temps, qu'il existe des méthodes graphiques et calculatoires qui permettent de choisir le modèle de décomposition (additif ou multiplicatif) et de déterminer la période notée p de la composante périodique.
Exemples :

- si les données sont mensuelles et la périodicité est annuelle : $p=12$
- si les données sont trimestrielles et la périodicité est annuelle : $p=4$
- si les données sont mensuelles et la périodicité est trimestrielle : $p=3$

Citons un critère simple et graphique permettant de choisir le modèle : si la composante périodique semble être indépendante de la tendance, ce qui se traduit par des variations d'amplitudes égales, choisir le modèle additif, sinon choisir le modèle multiplicatif.

Un peu de théorie pour les curieux : on considère une fonction f continue sur $\mathbb{R}$. On suppose qu'il existe deux nombres réels $T>0$, K tels que pour tout réel x :

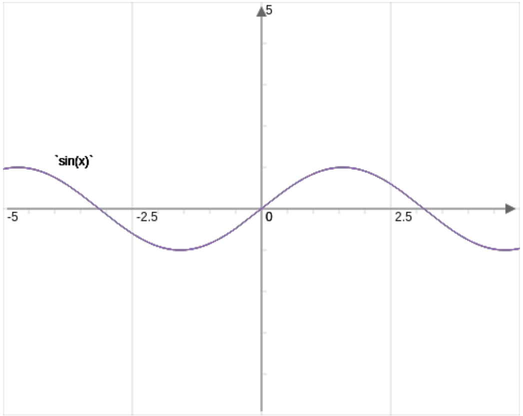
$$\int_x^{x+T} f(t)\,dt=K$$

Interprétation graphique : aire algébrique entre la courbe de f et l'axe des abscisses, entre les abscisses x et $x+T$ est égale à K .

On note F une primitive (fonction de dérivée f) de f , on obtient par définition $F(x+T)-F(x)=K$. En dérivant cette relation et en appliquant les règles classiques de dérivation, on en déduit $f(x+T)-f(x)=0 \Leftrightarrow f(x+T)=f(x)$. f est donc périodique de période T .

On se place dans le cas particulier $K=0$. La relation $\int_x^{x+T} f(t)\,dt=K=0$ indique, par interprétation graphique que l'aire algébrique sur toute période est nulle, en gros : l'aire positive sous la courbe lorsque f est positive compense l'aire négative sur la courbe lorsque f est négative . On imposera cette condition à la composante périodique du signal $x(t)$

Illustrons ceci dans le cas particulier $f(x)=\sin(x)$, $T=2\pi$, $\int_{-\pi}^{\pi} f(t)\,dt=0$



On note $a_0=a, a_1, a_2, \dots, a_{k-1}, a_k=a+T$ une subdivision de l'intervalle $[a;a+T]$. On admet l'approximation $\int_a^{a+T} f(t)\,dt \approx \frac{T}{k}(f(a_1)+f(a_2)+\dots+f(a_k))$ (Somme de Riemann)

Conclusion :

Condition imposée à la composante périodique :

$f(a_1)+f(a_2)+\dots+f(a_k)=0$ (condition notée (C1))

k représente le nombre de sous-intervalles réguliers composant un intervalle de longueur T

a – détermination de la tendance non corrigée des fluctuations saisonnières

En étudiant la courbe de $x(t)$, deux méthodes sont applicables :

Méthode 1 : application d'un modèle de régression (linéaire exponentielle, …). Cette méthode décrite précédemment permet d'obtenir une tendance $m_1(t)$

Méthode 2 : méthode des moyennes mobiles développée ci-dessous
Cette méthode permet de lisser la courbe de x et de supprimer ainsi la composante périodique

On définit la moyenne mobile comme suit.

cas p impair ($p=2i+1$) :

$$m_1(t)=\frac{1}{p}(x(t-i)+x(t-(i-1))+\dots+x(t-1)+x(t)+x(t+1)+\dots+x(t+i-1)+x(t+i))$$

cas p pair ($p=2i$) :

$$m_1(t)=\frac{1}{p}\left(\frac{x(t-i)}{2}+x(t-(i-1))+\ldots+x(t-1)+x(t)+x(t+1)+\ldots+x(t+i-1)+\frac{x(t+i)}{2}\right)$$

Exemple 5 : $p=4$, $i=2$, $m_1(3)=\frac{1}{4}\left(\frac{x(1)}{2}+x(2)+x(3)+x(4)+\frac{x(5)}{2}\right)$

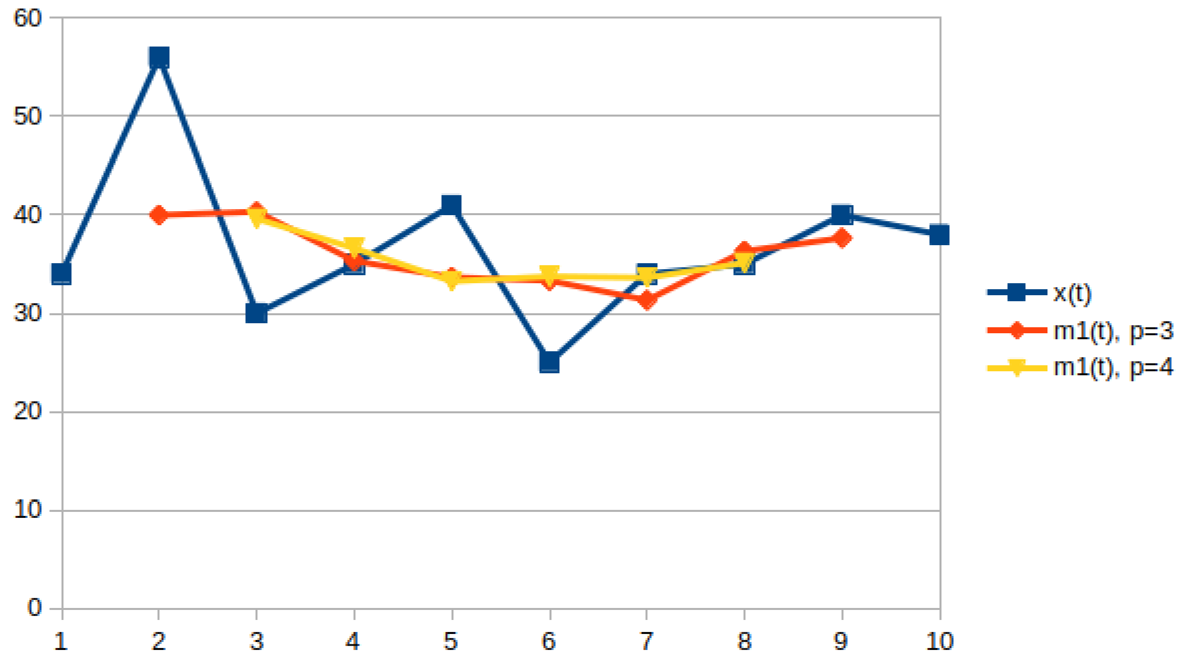
$m_1(t)$ apparaît comme une moyenne pondérée des valeurs de $x(t)$ autour de l'indice t

Remarque : dans les cas pair ou impair, les i premières valeurs et les i dernières valeurs de $m_1(t)$ ne sont pas calculables , à cause d'un dépassement d'indice

On obtient dans les tous les cas, une composante $m_1(t)$, notée tendance non corrigée

Exemple 6 : tableau de valeurs et représentation graphique

t	x(t)	m1(t), p=3	m1(t), p=4
1	34	?	?
2	56	40,000	?
3	30	40,333	39,625
4	35	35,333	36,625
5	41	33,667	33,25
6	25	33,333	33,75
7	34	31,333	33,625
8	35	36,333	35,125
9	40	37,667	?
10	38	?	?



b – détermination des coefficients saisonniers

À ce stade, on considère la composante $m_1(t)$ connue (tendance non corrigée). On en déduit les écarts périodiques $s_1(t)$.

Cas du modèle additif : $s_1(t)=x(t)-m_1(t)$
Cas du modèle multiplicatif (1) : $s_1(t)=\dfrac{x(t)}{m_1(t)}$

Les valeurs de $s_1(t)$ permettent maintenant de déterminer les coefficients saisonniers bruts $c_1(j)$ où $1\leq j\leq p$.

Pour simplifier, on considère que le nombre total n de valeurs $x(t)$ est un multiple de p, période de la composante périodique. On détermine p coefficients saisonniers bruts notés $c_1(j)$ où $1\leq j\leq p$: moyenne (ou médiane si on veut exclure des données fantaisistes) de tous les $s_1(t)$ situés au rang j des périodes successives.

Méthode préconisée : regrouper les $s_1(t)$ dans un tableau en les superposant par LIGNE de p valeurs. Les $c_1(j)$ sont calculés par COLONNE pour toutes les p colonnes.
On note $\bar{c}$ la moyenne de tous les $c_1(j)$.

Exemple 7 : Voyons tout cela sur les ventes de glaces par trimestre sur les années 2000, 2001 et 2002 de la maison AKAPULCO où la périodicité est $p=4$ et le modèle additif

t	x(t)	m1(t)	s1(t)
1	12000	?	?
2	20900	?	?
3	54000	29100	24900
4	30000	29237,5	762,5
5	11000	30250	-19250
6	23000	30875	-7875
7	60000	31000	29000
8	29000	31125	-2125
9	13000	29625	-16625
10	22000	28125	-6125
11	49000	?	?
12	28000	?	?

	regroupement tabulaire des s1(t)			
	?	?	24900	762,5
	-19250	-7875	29000	-2125
	-16625	-6125	?	?
c1	-17937,5	-7000	26950	-681,25
moyenne des c1	332,81			

On fait l'hypothèse que sur l'ensemble de la série des $x(t)$, ce sont les <u>mêmes</u> effets périodiques qui s'applique à chaque période. On adapte donc la condition (C1) évoquée précédemment et on corrige les coefficients $c_1(j)$ de la façon suivante.

Cas du modèle additif : $c_2(j)=c_1(j)-\bar{c}$

Cas du modèle multiplicatif (1) : $c_2(j)=\dfrac{c_1(j)}{\bar{c}}$

☞ les $c_2(j)$ où $1\leq j\leq p$ sont les coefficients saisonniers
Ce sont ces coefficients qui fournissent les valeurs de la composante périodique définitive $s_2(t)$ de $x(t)$ comme suit.

On note $r(t,p)$ le reste de la division euclidienne de t par p

$$s_2(t)=c_2(1+r(t-1,p))$$

Exemple 8 : $p=4$
$s_2(1)=c_2(1)$, $s_2(2)=c_2(2)$, $s_2(3)=c_2(3)$, $s_2(4)=c_2(4)$
$s_2(5)=c_2(1)$, $s_2(6)=c_2(2)$, $s_2(7)=c_2(3)$, $s_2(8)=c_2(4)$
$s_2(9)=c_2(1)$, ...

$s_2(t)$ permet d'extraire la tendance corrigée définitive $m_2(t)$ comme suit.

Cas du modèle additif : $m_2(t)=x(t)-s_2(t)$

Cas du modèle multiplicatif (1) : $m_2(t)=\dfrac{x(t)}{s_2(t)}$

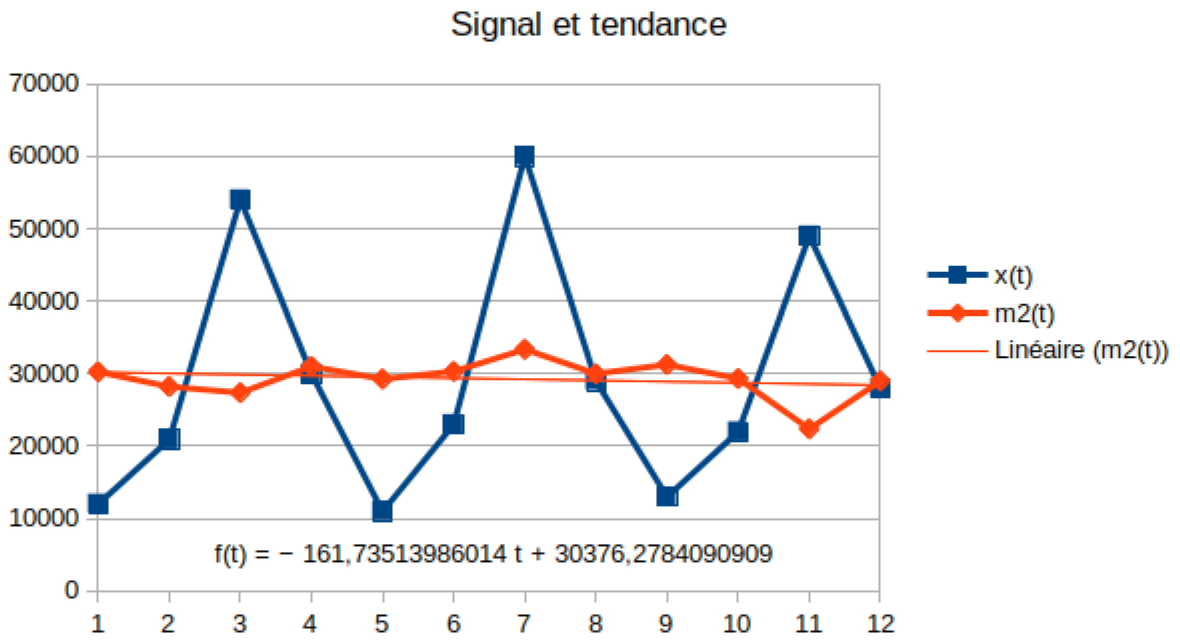
☞ La série $m_2(t)$ est notée série désaisonnalisée, série corrigée des variations saisonnières ou série CVS

Exemple 9 : Voici tous les calculs concernant la vente de glaces chez AKAPULCO

t	x(t)	m1(t)	s1(t)	s2(t)	m2(t)
1	12000	?	?	-18270,31	30270,31
2	20900	?	?	-7332,81	28232,81
3	54000	29100	24900	26617,19	27382,81
4	30000	29237,5	762,5	-1014,06	31014,06
5	11000	30250	-19250	-18270,31	29270,31
6	23000	30875	-7875	-7332,81	30332,81
7	60000	31000	29000	26617,19	33382,81
8	29000	31125	-2125	-1014,06	30014,06
9	13000	29625	-16625	-18270,31	31270,31
10	22000	28125	-6125	-7332,81	29332,81
11	49000	?	?	26617,19	22382,81
12	28000	?	?	-1014,06	29014,06

Il faut bien noter que pour le calcul des $s_2(t)$, ce sont les coefficients $c_2(j)$ que l'on applique de façon cyclique

	regroupement tabulaire des s1(t)					
	?	?	24900	762,5		
	-19250	-7875	29000	-2125		
	-16625	-6125	?	?		
c1	-17937,5	-7000	26950	-681,25		
moyenne des c1	332,81					
c2	-18270,31	-7332,81	26617,19	-1014,06	Moyenne =	0



c – prévision

On suppose connus pour un signal x donné :

- n valeurs $x(1),\dots,x(n)$
- la période p de la composante périodique
- le coefficients saisonniers $c_2(j)$ où $1 \leq j \leq p$
- les valeurs de la tendance $m_2(t)$

Question : comment estimer les valeurs $x(n+1), x(n+2), \dots$?

Réponse : si on peut réaliser une régression (linéaire ou autre) de la tendance, on pose $f(t)$ la fonction associée à cette régression. On en déduit :

$x(n+1) = f(n+1) + c_2(1)$

$x(n+2) = f(n+2) + c_2(2)$ ...

Exemple 10 : toujours chez AKAPULCO, estimons $x(13)$ à l'aide de la régression linéaire de la tendance et les $c_2(j)$

$$x(13) \approx -161{,}74 \times 13 + 30376{,}28 - 18270{,}31 = 10003$$