

# STT 1000 - STATISTIQUES

ARTHUR CHARPENTIER



# Notations

alpha	$\alpha$	significativité			
beta	$\beta$	erreur			
gamma	$\gamma, \Gamma$		sigma	$\sigma, \Sigma$	variance $\sigma^2$
delta	$\delta, \Delta$		tau	$\tau$	
epsilon	$\epsilon, \varepsilon$		upsilon	$\upsilon$	
zeta	$\zeta$		phi	$\phi, \varphi, \Phi$	loi $\mathcal{N}(0, 1)$
eta	$\eta$		chi	$\chi$	loi $(\chi^2)$
theta	$\theta, \Theta$	paramètre	psi	$\psi, \Psi$	
iota	$\iota$		omega	$\omega, \Omega$	évènement
kappa	$\kappa$		<i>exempli gratia</i>	e.g.	par exemple
lambda	$\lambda, \Lambda$	$\mathcal{P}(\lambda)$	<i>id est</i>	i.e.	c'est-à-dire
mu	$\mu$	moyenne			
nu	$\nu$	degrés de liberté			
xi	$\xi, \Xi$				
pi	$\pi, \Pi$	3.1415...			
rho	$\rho$	corrélacion			

# Notations

**1** indicatrice  $\mathbf{1}_A(x) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{si } x \notin A \end{cases}$  (parfois noté  $\mathbf{1}(x \in A)$ )

$\mathbb{N}$  ensemble des entiers naturels,  $\{0, 1, 2, \dots\}$

$\mathbb{R}$  ensemble des réels

$\cap$  intersection, e.g.  $\{A, B\} \cap \{B, C\} = \{B\}$

$\cup$  union, e.g.  $\{A, B\} \cup \{B, C\} = \{A, B, C\}$

$\forall$  quel que soit

$\exists$  il existe

$\sum_{i=1}^n$  somme,  $\sum_{i=1}^n x_i = x_1 + x_2 + \dots + x_n$

$\prod_{i=1}^n$  produit,  $\prod_{i=1}^n x_i = x_1 \cdot x_2 \cdot \dots \cdot x_n$

**!** factorielles,  $n! = \prod_{k=1}^n k = n(n-1)(n-2) \dots 3 \cdot 2 \cdot 1$

**log** logarithme népérien  $\log(e) = 1$

# Notations

$\lfloor x \rfloor$	partie entière de $x$ , $\lfloor x \rfloor \leq x < \lfloor x \rfloor + 1$
$\lceil x \rceil$	partie entière supérieure de $x$ , $\lceil x \rceil - 1 < x \leq \lceil x \rceil$
$\cdot^{-1}$	inverse, i.e. $\mathbf{M}^{-1}$ , matrice telle que $\mathbf{M}\mathbf{M}^{-1} = \mathbf{I}$
	inverse, i.e. $f^{-1}$ , fonction telle que $f[f^{-1}(x)] = x$ , $\forall x$
$\frac{d}{dx}$ ou $\frac{\partial}{\partial x}$	dérivation (cf plus loin)
$\min_{x \in \mathcal{X}} \{f(x)\}$	minimum, i.e. valeur $f^* = f(x^*)$
$\operatorname{argmin}_{x \in \mathcal{X}} \{f(x)\}$	argument du minimum, i.e. valeur $x^*$
$\mathbb{P}$	probabilité
$\mathbb{E}$	espérance mathématique
$\operatorname{Var}$	variance

# Définitions, propriétés, formules, etc

Dans le cours, nous allons définir des objets (avec des définitions)

## Variance (définition)

La variance d'une variable aléatoire  $X$  est

$$\text{Var}[X] \stackrel{\text{def}}{=} \mathbb{E}[(X - \mathbb{E}[X])^2]$$

## Erreur quadratique moyenne (définition)

L'erreur quadratique moyenne (ou MSE, *mean squared error*) d'un estimateur  $\hat{\theta}$  d'un paramètre  $\theta$  est

$$\text{EQM}[\hat{\theta}] \stackrel{\text{def}}{=} \mathbb{E}[(\hat{\theta} - \theta)^2]$$

... une définition ne s'invente pas !

# Définitions, propriétés, formules, calculs

Mais nous allons aussi avoir des propriétés, théorèmes

## Variance (propriété)

La variance d'une variable aléatoire  $X$  vérifie

$$\text{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

## Erreur quadratique moyenne (propriété)

L'erreur quadratique moyenne d'un estimateur  $\hat{\theta}$  d'un paramètre  $\theta$  vérifie

$$\text{EQM}[\hat{\theta}] = \text{Biais}(\hat{\theta})^2 + \text{Var}(\hat{\theta})$$

... une propriété se démontre (et peut se retrouver)

## Définitions, propriétés, formules, calculs

par exemple, pour la propriété de la variance

$$\begin{aligned}\text{Var}(X) &= \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2 - 2X\mathbb{E}[X] + \mathbb{E}[X]^2] \\ &= \mathbb{E}[X^2] - 2\mathbb{E}[X]\mathbb{E}[X] + \mathbb{E}[X]^2 = \mathbb{E}[X^2] - \mathbb{E}[X]^2\end{aligned}$$

Dans le cours, je donnerai des exemples de calculs,  
par exemple, si  $X$  suit une loi exponentielle,  $f(x) = \lambda e^{-\lambda x}$

$$\begin{aligned}\mathbb{E}[X] &= \int_0^\infty \lambda x e^{-\lambda x} dx = \frac{1}{\lambda} \\ \mathbb{E}[X^2] &= \int_0^\infty \lambda x^2 e^{-\lambda x} dx = \left[-x^2 e^{-\lambda x}\right]_0^\infty + \int_0^\infty 2x e^{-\lambda x} dx \\ &= 0 + \frac{2}{\lambda} \mathbb{E}[X] = \frac{2}{\lambda^2}.\end{aligned}$$

de telle sorte que

$$\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}.$$

# Définitions, propriétés, formules, calculs

Certaines preuves ne seront pas faites, mais pour la plupart des résultats, j'essayerais de donner une explication heuristique

**Note:** tous les concepts et résultats de ce cours se trouvent dans n'importe quel livre de statistique (cf slides # 1)

L'accent sera aussi mis sur les outils numériques / informatiques

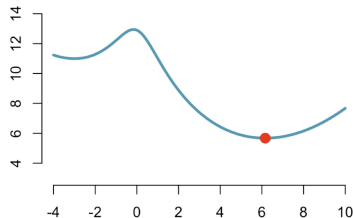
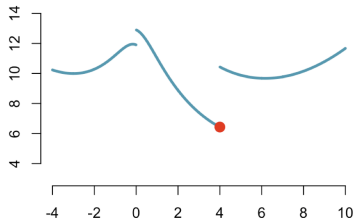
```
1 > mean(Davis$height)
2 [1] 170.565
3 > var(Davis$height)
4 [1] 79.7847
5 > mean(Davis$height^2) - mean(Davis$height)^2
6 [1] 79.3857
```

*“Dans toute statistique, l'inexactitude du nombre est compensée par la précision des décimales” (Alfred Sauvy)*



# Hypothèses, conditions nécessaires & conditions suffisantes

En statistique on va chercher le *meilleur modèle* ou le *meilleur estimateur*. Le terme “*meilleur*” signifie qu’on cherche à maximiser un critère (quantitatif).



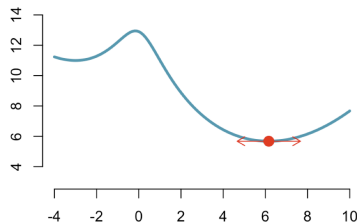
On cherche soit  $x^* = \underset{x \in \mathcal{X}}{\operatorname{argmin}}\{f(x)\}$ , soit  $f^* = \min_{x \in \mathcal{X}}\{f(x)\}$

Si le problème est trop général, difficile de trouver des conditions que doit/devrait/pourrait satisfaire  $x^*$

On va alors faire des *hypothèses*, c’est à dire imposer des conditions (réalistes) que doit/devrait/pourrait satisfaire  $f$ .

# Hypothèses, conditions nécessaires & conditions suffisantes

**Example:** Supposons que  $f$  soit dérivable (de dérivée continue)



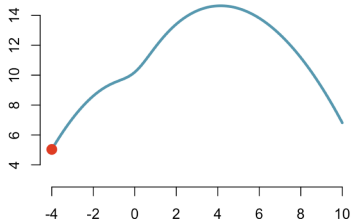
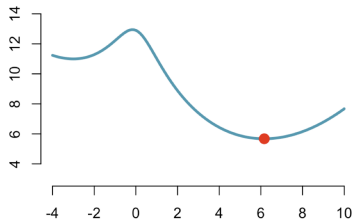
Condition possible :  $\left. \frac{df(x)}{dx} \right|_{x=x^*} = 0$ , ou  $f'(x^*) = 0$ .

$$x^* = \operatorname{argmin}_{x \in \mathcal{X}} \{f(x)\} \stackrel{?}{\Leftrightarrow} f'(x^*) = 0$$

**Problème (1)**  $x^* = \operatorname{argmin}_{x \in \mathcal{X}} \{f(x)\} \not\Leftrightarrow f'(x^*) = 0$

**Problème (2)**  $f'(x^*) = 0 \not\Leftrightarrow x^* = \operatorname{argmin}_{x \in \mathcal{X}} \{f(x)\}$

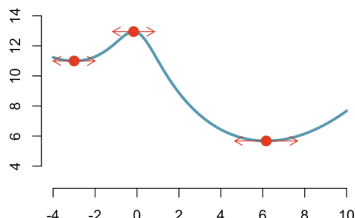
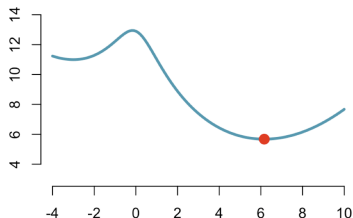
# Hypothèses, conditions nécessaires & conditions suffisantes



On peut avoir  $f'(x^*) \neq 0$  quand  $x^*$  est au bord du support  $\mathcal{X}$   
Condition possible (1) : on exclut les solutions de bords

Si  $\mathcal{X}$  ouvert,  $x^* = \operatorname{argmin}_{x \in \mathcal{X}} \{f(x)\}$  et  $x \in \mathcal{X} \Rightarrow f'(x^*) = 0$   
(condition d'Euler, ou de Fermat, 1629)

# Hypothèses, conditions nécessaires & conditions suffisantes



On peut avoir  $f'(x^*) = 0$  et  $x^* \neq \underset{x \in \mathcal{X}}{\operatorname{argmin}}\{f(x)\}$

Cas (1) :  $x^*$  peut être un maximum, on va imposer  $f'(x) > 0$

Cas (2) :  $x^*$  peut être un minimum local

mais ne suffit pas...

$f'(x^*) = 0 \not\Rightarrow x^* = \underset{x \in \mathcal{X}}{\operatorname{argmin}}\{f(x)\}$

$f'(x^*) = 0$  : condition nécessaire (moyennant quelques hypothèses de régularité) mais pas suffisante...

# Logarithme & Exponentielle

## Logarithme & Exponentielle

L'exponentielle est l'unique fonction  $f$  dérivable telle que  $f(0) = 1$  et  $f(x) = f'(x) \forall x \in \mathbb{R}$ ,  $f(x) = \exp[x] = e^x$ .

Le logarithme est la fonction inverse de l'exponentielle,

$$\log(x) = \int_1^x \frac{dt}{t} \text{ pour } x \in \mathbb{R}_+ \text{ et } \log[e^x] = x, \text{ pour } x \in \mathbb{R}$$

$$\frac{d \log(x)}{dx} = \frac{1}{x}, \text{ et si } u \text{ est différentiable, } \frac{d \log(u(x))}{dx} = \frac{u'(x)}{u(x)}$$

$$\frac{d \exp(x)}{dx} = \exp(x), \text{ et } \frac{d \exp(u(x))}{dx} = u'(x) \cdot \exp[u(x)]$$

$$\begin{cases} \log(ab) = \log(a) + \log(b), \forall a, b \in \mathbb{R}_+ \\ \exp[a + b] = \exp[a] \cdot \exp[b], \forall a, b \in \mathbb{R} \end{cases}$$

# Logarithme & Exponentielle

Pour  $n \in \mathbb{N}$ ,

$$\lim_{x \rightarrow 0^+} x^n \log(x) = 0 \text{ et } \lim_{x \rightarrow +\infty} \frac{\log(x)}{x^n} = 0$$

$$\lim_{x \rightarrow -\infty} x^n \exp(x) = 0 \text{ et } \lim_{x \rightarrow +\infty} \frac{\exp(x)}{x^n} = +\infty$$

Suite géométrique,  $u_n - u_{n-1} = k \cdot u_{n-1}$ ,

$$u_n = (1 + k) \cdot u_{n-1} = (1 + k)^n \cdot u_0$$

Version continue (taux d'accroissement)

$$\frac{u(x+h) - u(x)}{h} = k \cdot u(x) \text{ ou } u'(x) = k \cdot u(x)$$

alors  $u(x) = \exp[kx] \cdot u(0) = \exp[k]^x \cdot u(0)$ .

# Notations

## Fonction Gamma

$$\Gamma : z \mapsto \int_0^{+\infty} t^{z-1} e^{-t} dt \text{ pour } z \in \mathbb{R}_+$$

On peut montrer que  $\Gamma(z+1) = z \Gamma(z)$

$$\begin{aligned}\Gamma(z+1) &= \int_0^{\infty} x^z e^{-x} dx \\ &= \left[ -x^z e^{-x} \right]_0^{\infty} + \int_0^{\infty} z x^{z-1} e^{-x} dx \\ &= \lim_{x \rightarrow \infty} (-x^z e^{-x}) - (-0^z e^{-0}) + z \int_0^{\infty} x^{z-1} e^{-x} dx.\end{aligned}$$

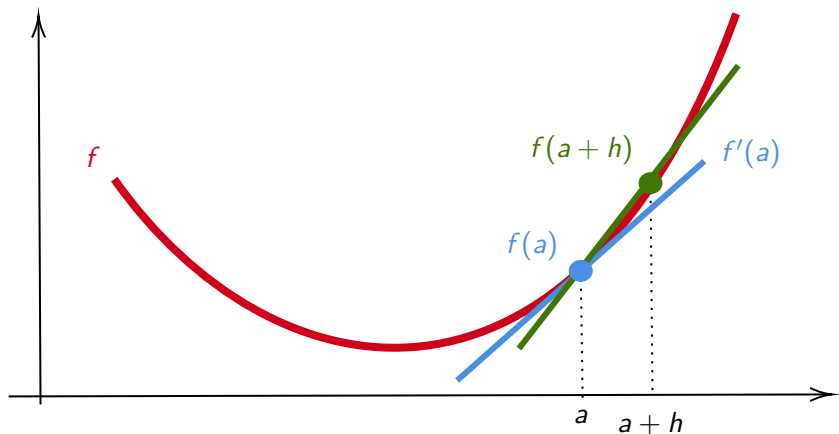
comme  $-x^z e^{-x} \rightarrow 0$  lorsque  $x \rightarrow \infty$

$$\Gamma(z+1) = z \int_0^{\infty} x^{z-1} e^{-x} dx = z \Gamma(z)$$

Pour  $z \in \mathbb{N}$ ,  $\Gamma(z+1) = z!$ ,  $\Gamma(1) = 1$  et  $\Gamma(1/2) = \sqrt{\pi}$ .

# Dérivées

$$f'(a) = \left. \frac{df(x)}{dx} \right|_{x=a} = \lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{f(a+h) - f(a)}{h}$$





## Dérivée

Si  $f : \mathbb{R} \rightarrow \mathbb{R}$  est dérivable en  $a \in \mathbb{R}$ ,

$$f'(a) = \left. \frac{df(x)}{dx} \right|_{x=a} = \lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{f(a+h) - f(a)}{h}$$

## Dérivée seconde

Si  $f : \mathbb{R} \rightarrow \mathbb{R}$  est deux fois dérivable en  $a \in \mathbb{R}$ ,

$$f''(a) = \left. \frac{d^2 f(x)}{dx^2} \right|_{x=a} = \left. \frac{df'(x)}{dx} \right|_{x=a} = \lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{f'(a+h) - f'(a)}{h}$$

## Gradient

Si  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  est dérivable en  $\mathbf{a} = (a_1, a_2) \in \mathbb{R}$ ,

$$\nabla f(\mathbf{a}) = \left( \frac{\partial f(x_1, x_2)}{\partial x_1}, \frac{\partial f(x_1, x_2)}{\partial x_2} \right) \Big|_{(x_1, x_2) = (a_1, a_2)}$$

$$\frac{\partial f(x_1, x_2)}{\partial x_1} \Big|_{(x_1, x_2) = (a_1, a_2)} = \lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{f(a_1 + h, a_2) - f(a_1, a_2)}{h}$$

$$\frac{\partial f(x_1, x_2)}{\partial x_2} \Big|_{(x_1, x_2) = (a_1, a_2)} = \lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{f(a_1, a_2 + h) - f(a_1, a_2)}{h}$$

## Hessienne

Si  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  est dérivable en  $\mathbf{a} = (a_1, a_2) \in \mathbb{R}$ ,

$$\mathbb{H}(\mathbf{a}) = \left( \begin{array}{cc} \frac{\partial^2 f(x_1, x_2)}{\partial x_1^2} & \frac{\partial^2 f(x_1, x_2)}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f(x_1, x_2)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x_1, x_2)}{\partial x_2^2} \end{array} \right) \bigg|_{(x_1, x_2) = (a_1, a_2)}$$

# Dérivées

**Exemple:**  $f(x) = \frac{1}{2}x^4 - x\sqrt{3} + 3\sqrt{x},$

$$f'(x) = \frac{4x^3\sqrt{x} - 2\sqrt{3x} + 3}{2\sqrt{x}}$$

**Exemple:**  $f(x) = (x^2 + 3)x^5, f'(x) = x^4(7x^2 + 15)$

**Exemple:**  $f(x) = x^2\sqrt{x}, f'(x) = \frac{5}{2}x\sqrt{x}$

**Exemple:**  $f(x) = \left(x + \frac{1}{x}\right)x, f'(x) = 2x$

**Exemple:**  $f(x) = \frac{3}{x+1}, f'(x) = \frac{-3}{(x+1)^2}$

**Exemple:**  $f(x) = \frac{x^2 + \frac{3}{x}}{x^2 + \frac{x}{3}}, f'(x) = 3\frac{x^3 - 27x - 6}{x(3x^2 + x)^2}$

**Exemple:**  $f(x) = \left(1 + \frac{1}{\sqrt{x}}\right)\left(1 - \frac{1}{\sqrt{x}}\right), f'(x) = \frac{1}{x^2}$

# Primitives & Intégrales

## Intégrale

Soit  $f$  définie sur  $[a, b]$ , et  $F$  telle que  $F' = f$ , alors

$$\int_a^b f(x) dx = F(b) - F(a)$$

ou, pour une intégrale impropre

$$\int_a^b f(x) dx = \lim_{x \rightarrow b} F(x) - \lim_{x \rightarrow a} F(x)$$

Linéarité de l'intégrale :

$$\int_a^b (f(x) + g(x)) dx = \int_a^b f(x) dx + \int_a^b g(x) dx$$

(si les deux intégrales existent)

# Primitives & Intégrales

**Exemple:**  $I_1 = \int_{-1}^3 (5 - 2x) dx, I_1 = 12$

**Exemple:**  $I_2 = \int_0^1 x^2 (x^3 - 1)^5 dx, I_2 = \frac{-1}{18}$

**Exemple:**  $I_3 = \int_0^1 \frac{x}{(x^2 - 4)^2} dx, I_3 = \frac{1}{24}$

**Exemple:**  $I_4 = \int_0^1 e^{-2x} dx, I_4 = \frac{1}{2} (1 - e^{-2})$

**Exemple:**  $I_5 = \int_{-1}^1 2x(8x + 2)^2 dx, I_5 = \frac{128}{3}$

**Exemple:**  $I_6 = \int_0^1 x^2 e^x dx, I_6 = e - 2$

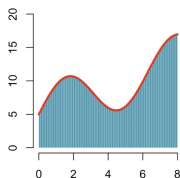
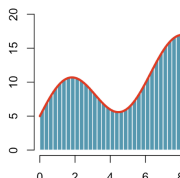
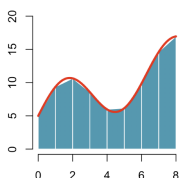
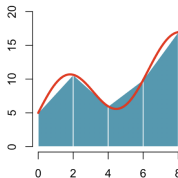
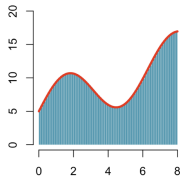
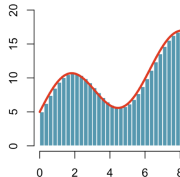
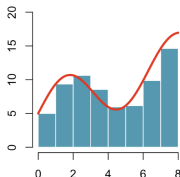
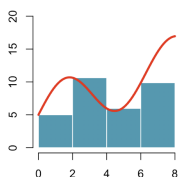
**Exemple:**  $F(y) = \int_{*}^y \frac{1}{x^2} + 3x dx, F(Y) = -\frac{1}{y} + \frac{3}{2}y^2 + \text{cst}$

**Exemple:**  $F(y) = \int_{*}^y \frac{5}{(-2x + 1)^2} + 3dx = \frac{5}{2} \frac{1}{-2y + 1} + 3y + \text{cst}$

# Primitives & Intégrales

On sera a priori intéressé par un calcul numérique d'une intégrale,

```
1 > f = function(x) 5-2*x  
2 > integrate(f, -1, 3)  
3 12 with absolute error < 1.4e-13
```



# Règle de L'Hôpital

## Règle de l'Hôpital

Si  $f$  et  $g$  sont deux fonctions définies sur  $[a, b[$  (dérivables en  $a$ , et telles que  $f(a) = g(a) = 0$  et  $g'(a) \neq 0$ , alors

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \frac{f'(a)}{g'(a)}$$

### Exemple

$$\begin{aligned}\lim_{x \rightarrow 0} \frac{e^x - 1}{2x} &= \frac{1}{2} \\ \lim_{x \rightarrow 0} \frac{e^x - 1}{2x} &= \lim_{x \rightarrow 0} \frac{[e^x - 1]'}{[2x]'} = \lim_{x \rightarrow 0} \frac{e^x}{2} = \frac{1}{2}\end{aligned}$$

Approche alternative:  $e^x \approx 1 + x + \frac{x^2}{2} + \frac{x^3}{6}$

$$\lim_{x \rightarrow 0} \frac{e^x - 1}{2x} = \lim_{x \rightarrow 0} \frac{1}{2x} \left( x + \frac{x^2}{2} + \frac{x^3}{6} \right) = \lim_{x \rightarrow 0} \left( \frac{1}{2} + \frac{x}{4} + \frac{x^2}{12} \right) = \frac{1}{2}$$