

Statistiques, Probabilités, Analyse Spatiale

PRÉSENTATION DU MODULE

Paul Chapron¹ & Yann Ménéroux¹ & Juste Raimbault

2021-2022

¹IGN-ENSG-UGE



Introduction

Citations en vrac

«If you torture data long enough, it will confess»

Ronald Coase

«It is a capital mistake to theorize before one has data.»

Sherlock Holmes

«All models are wrong, but some are useful.»

George Box

«Data scientist (noun): Person who is better at statistics than any software engineer and better at software engineering than any statistician.»

Josh Wills

«There are lies, damned lies and statistics»

Mark Twain

«Les statistiques sont une forme d'accomplissement de désir, tout comme les rêves.»

Jean Baudrillard

«Ce qui est simple est faux, ce qui est compliqué est inutile.»

Paul Valéry

«La statistique est la première des sciences inexactes.»

Édouard et Jules de Goncourt

- **Statistique** : Démarche scientifique visant à acquérir des connaissances sur l'état d'un objet difficilement perceptible sur le plan cognitif :
 - objet **massif** (population, big data...) : résumer, synthétiser, visualiser, échantillonner, compresser...
 - objet **incertain** (observations bruitées) : moyenner, borner, encadrer, compenser...

- **Statistique** : Démarche scientifique visant à acquérir des connaissances sur l'état d'un objet difficilement perceptible sur le plan cognitif :
 - objet **massif** (population, big data...) : résumer, synthétiser, visualiser, échantillonner, compresser...
 - objet **incertain** (observations bruitées) : moyenner, borner, encadrer, compenser...

En général, la statistique sert à **décrire** des états présents ou passés.

- **Probabilités** : théorie mathématique traitant de l'aléatoire, sans justification a priori (même si généralement fondée sur des besoins pratiques) et opérant dans un cadre contrôlé (information parfaite).

- **Probabilités** : théorie mathématique traitant de l'aléatoire, sans justification a priori (même si généralement fondée sur des besoins pratiques) et opérant dans un cadre contrôlé (information parfaite).

En général, les probabilités servent à **prédir** des états futurs.

- **Analyse Spatiale** : Étude de la répartition et de l'organisation d'objets localisés pour «déceler en quoi la localisation apporte un élément utile à la connaissances des objets étudiés et peut en expliquer les caractéristiques»

Pumain, Saint – Julien1997

- **Analyse Spatiale** : Étude de la répartition et de l'organisation d'objets localisés pour «déceler en quoi la localisation apporte un élément utile à la connaissances des objets étudiés et peut en expliquer les caractéristiques»

Pumain, Saint – Julien1997

Nous couvrirons principalement l'analyse spatiale **statistique** qui utilisent de nouvelles variables : les coordonnées des objets et les distances qui les séparent.

Programme du module

Semaine	Jour	Date	Commentaire / cours	Enseignant
semaine 43 du au 29/10/2022	lundi	24/10/2022	Introduction	Paul Chapron
	lundi	24/10/2022	Théorie des probabilités	Yann Meneroux
	mardi	25/10/2022	Analyse univariée, bases de R	Paul Chapron
	mardi	25/10/2022	Analyse univariée, bases de R	Paul Chapron
	mercredi	26/10/2022	Analyse bivariée	Paul Chapron
	mercredi	26/10/2022	Théorie des probabilités	Yann Meneroux
	jeudi	27/10/2022	Théorie des probabilités	Yann Meneroux
	vendredi	28/10/2022	ACP	Paul Chapron
	vendredi	28/10/2022	Théorie des probabilités	Yann Meneroux
	lundi	14/11/2022	Géostatistiques	Yann Meneroux
semaine 46 du 26/11/2022	mardi	15/11/2022	Intro	Sébastien Mustière
	mardi	15/11/2022	Géostatistiques	Yann Meneroux
	mercredi	16/11/2022	Statistiques spatiales	Juste Rimbault
	mercredi	16/11/2022	Statistiques spatiales	Juste Rimbault
	vendredi	18/11/2022	Intro	Sébastien Mustière
	lundi	21/11/2022	Statistiques spatiales	Juste Rimbault
	lundi	21/11/2022	Statistiques spatiales	Juste Rimbault
semaine 49 du 10/12/2022	mardi	22/11/2022	Lancement des projets	Y. Meneroux / P. Chapron / J. Rimbault
	mercredi	23/11/2022	Apprentissage supervisé	Yann Meneroux
	mercredi	23/11/2022	Apprentissage supervisé	Yann Meneroux
	vendredi	25/11/2022	Réseaux spatiaux	Juste Rimbault
	vendredi	25/11/2022	Réseaux spatiaux	Juste Rimbault
	lundi	05/12/2022	Apprentissage non supervisé	Yann Meneroux
	lundi	05/12/2022	Apprentissage non supervisé	Yann Meneroux
semaine 50 au 16/12/2022	mardi	06/12/2022	projet en autonomie	
	mardi	06/12/2022	projet en autonomie	
	mercredi	07/12/2022	Méthodologie de la validation	Yann Meneroux
	mercredi	07/12/2022	Méthodologie de la validation	Yann Meneroux
	jeudi	08/12/2022	projet en autonomie	
	jeudi	08/12/2022	suivi de projet	Y. Meneroux / P. Chapron / J. Rimbault
	vendredi	09/12/2022	Deep Learning	
semaine 51 au 20/12/2022	vendredi	09/12/2022	Deep Learning	
	lundi	12/12/2022	projet en autonomie	
	lundi	12/12/2022	projet en autonomie	
	mardi	13/12/2022	projet en autonomie	
	mardi	13/12/2022	Modèles de simulation	Juste Rimbault
	jeudi	15/12/2022	Interros "Stats / analyse spatiale / apprentissage"	
	vendredi	16/12/2022	Fin des projets	Y. Meneroux / P. Chapron / J. Rimbault

Naissance au XVIIe ?



Siège de Platée (-430 av J.C.)

Naissance au XVIIe ?

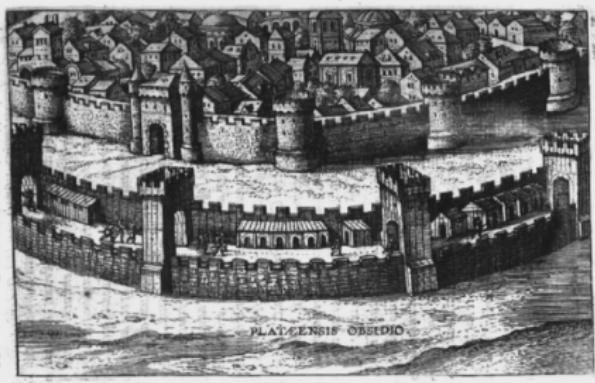


Siège de Platée (-430 av J.C.)

Naissance au XVIIe ?



Thucydide

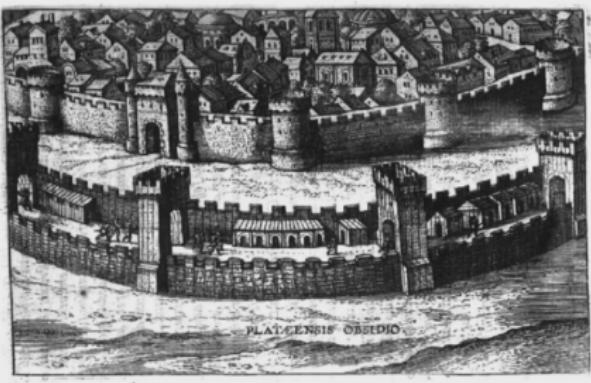


Siège de Platée (-430 av J.C.)

Naissance au XVIIe ?



Thucydide



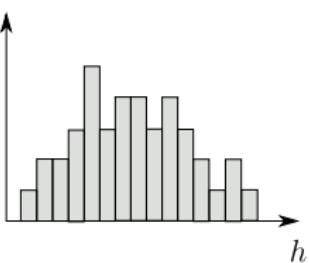
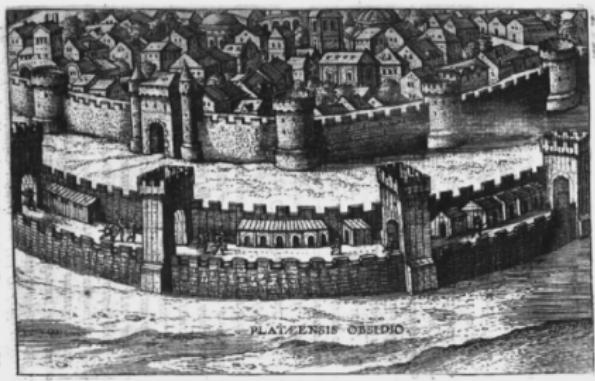
Siège de Platée (-430 av J.C.)

Naissance au XVIIe ?



Siège de Platée (-430 av J.C.)

Thucydide

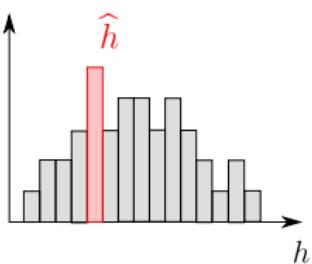
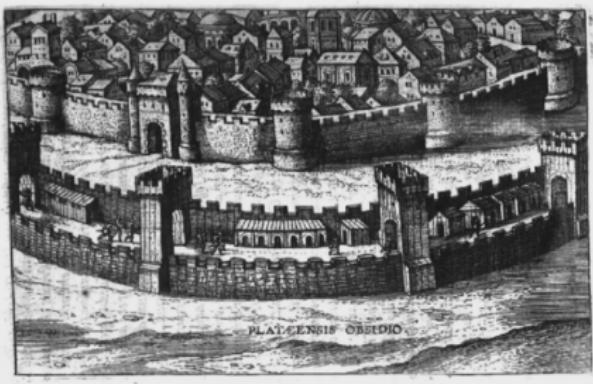


Naissance au XVIIe ?



Siège de Platée (-430 av J.C.)

Thucydide



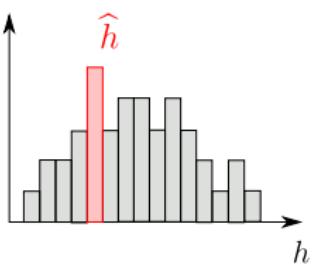
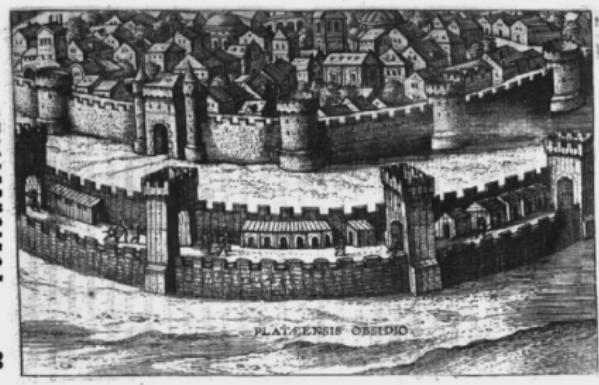
Naissance au XVIIe ?



Siège de Platée (-430 av J.C.)

$$L(h, x)$$

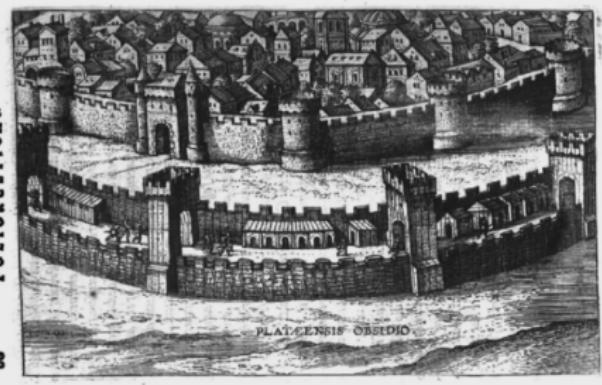
Thucydide



Naissance au XVIIe ?

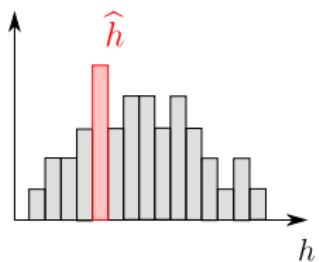


Thucydide



Siege de Platée (-430 av J.C.)

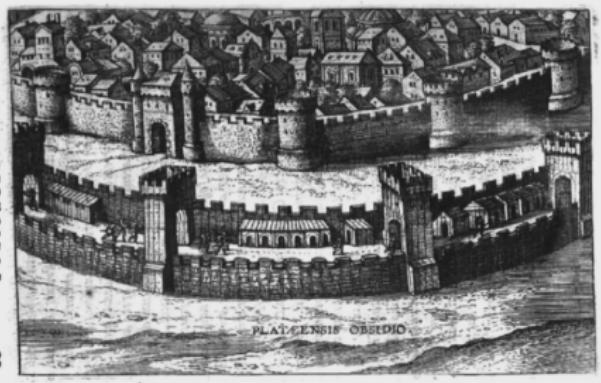
$$L(h, x) = 1 - \mathbf{1}_h(x) = \begin{cases} 0 & \text{si } x = h \\ 1 & \text{sinon.} \end{cases}$$



Naissance au XVIIe ?



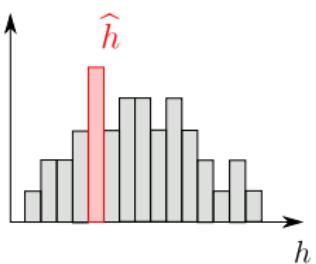
Thucydide



Siège de Platée (-430 av J.C.)

$$L(h, x) = 1 - \mathbf{1}_h(x) = \begin{cases} 0 & \text{si } x = h \\ 1 & \text{sinon.} \end{cases}$$

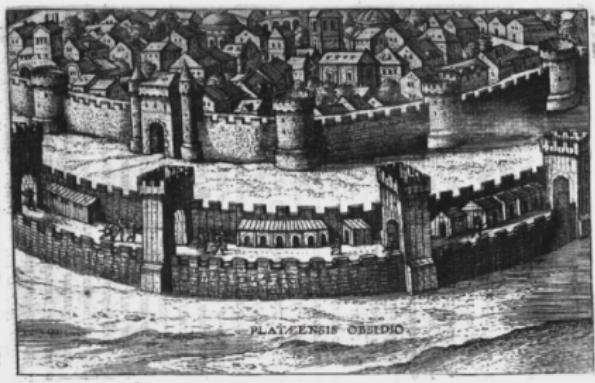
$$\hat{h} = \operatorname{argmin}_h \mathbb{E}[L(h, X)]$$



Naissance au XVIIe ?



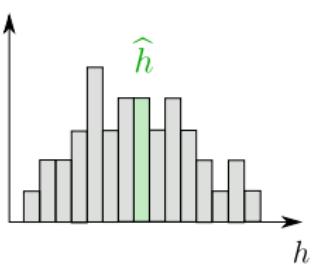
Thucydide



Siège de Platée (-430 av J.C.)

$$L(h, x) = |h - x| \quad (\text{Norme L1})$$

$$\hat{h} = \operatorname{argmin}_h \mathbb{E}[L(h, X)]$$



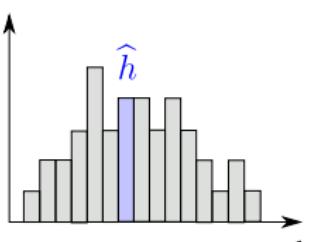
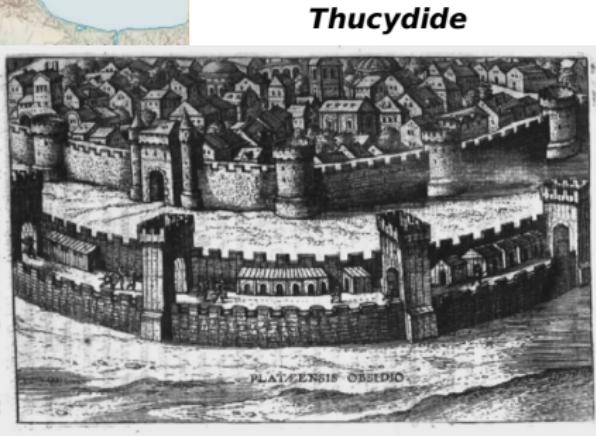
Naissance au XVIIe ?



Siège de Platée (-430 av J.C.)

$$L(h, x) = (h - x)^2 \quad (\text{Norme L2})$$

$$\hat{h} = \operatorname{argmin}_h \mathbb{E}[L(h, X)]$$



Correspondance de Pascal et Fermat (1654) : *Deux joueurs déposent chacun une mise m pour jouer une partie en 3 manches, mais décident de se quitter après deux victoires de l'un et une victoire de l'autre. Comment devraient-ils se répartir la mise ?*

- probabilisme (XVI^{ème} siècle)
- Pas d'utilisation en dehors des jeux
- Leibniz : outil de connaissance et de raisonnement sur le réel.
La Statistique devient un instrument scientifique (1677)
- Paradoxe de Saint-Petersbourg (1713)
- Bayes : théorie des causes (1763)
- Fin XVIII^{ème} : Arithmétique politique (démographie, tables de mortalité, etc)