

**STATISTIQUE NON-PARAMÉTRIQUE  
POUR DONNÉES FONCTIONNELLES:  
QUELQUES IDÉES DE BASE**

Philippe Vieu ([vieu@math.univ-toulouse.fr](mailto:vieu@math.univ-toulouse.fr))

En collaboration avec [Frédéric Ferraty](#)

Institut de Mathématiques, Toulouse, FRANCE

Page web du groupe STAPH: <http://www.math.univ-toulouse.fr/staph/>

# PLAN DE L'EXPOSÉ

I. Introduction

II. Quelques éléments historiques

→ Statistique non-paramétrique uni-dimensionnelle

→ Statistique non-paramétrique multi-dimensionnelle

III. Statistique non-paramétrique fonctionnelle

→ Un exemple d'estimateur

→ Un exemple de résultat asymptotique

IV. Quelques exemples applications

V. Quelques éléments bibliographiques

# PLAN DE L'EXPOSÉ

## I. Introduction

## II. Quelques éléments historiques

- Statistique non-paramétrique uni-dimensionnelle
- Statistique non-paramétrique multi-dimensionnelle

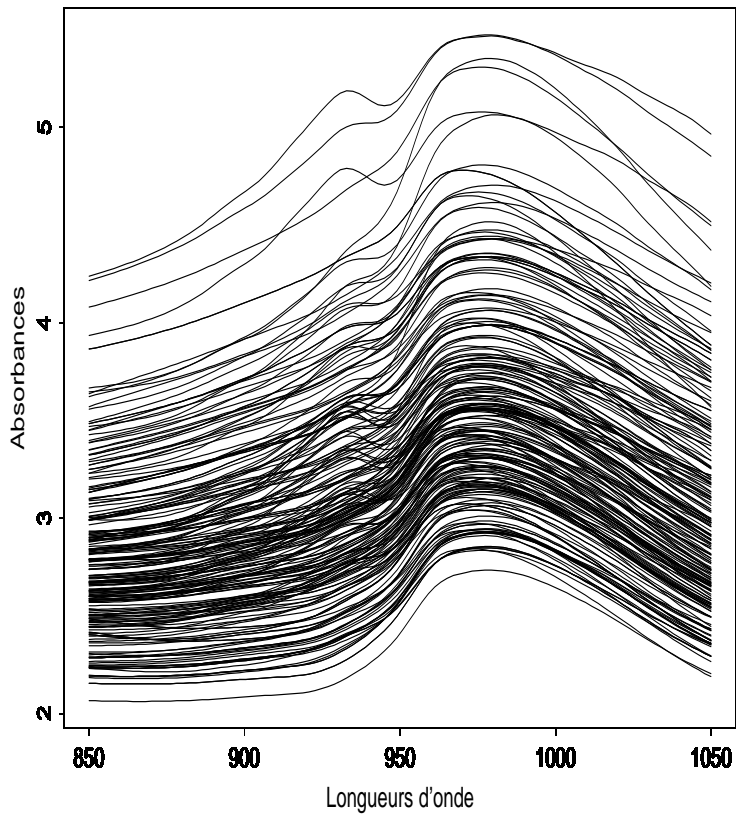
## III. Statistique non-paramétrique fonctionnelle

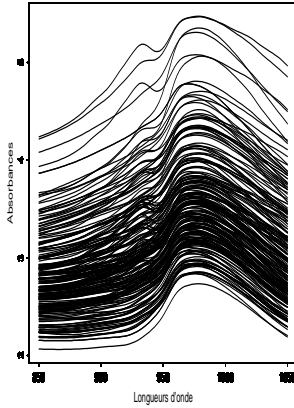
- Un exemple d'estimateur
- Un exemple de résultat asymptotique

## IV. Quelques exemples applications

## V. Quelques éléments bibliographiques

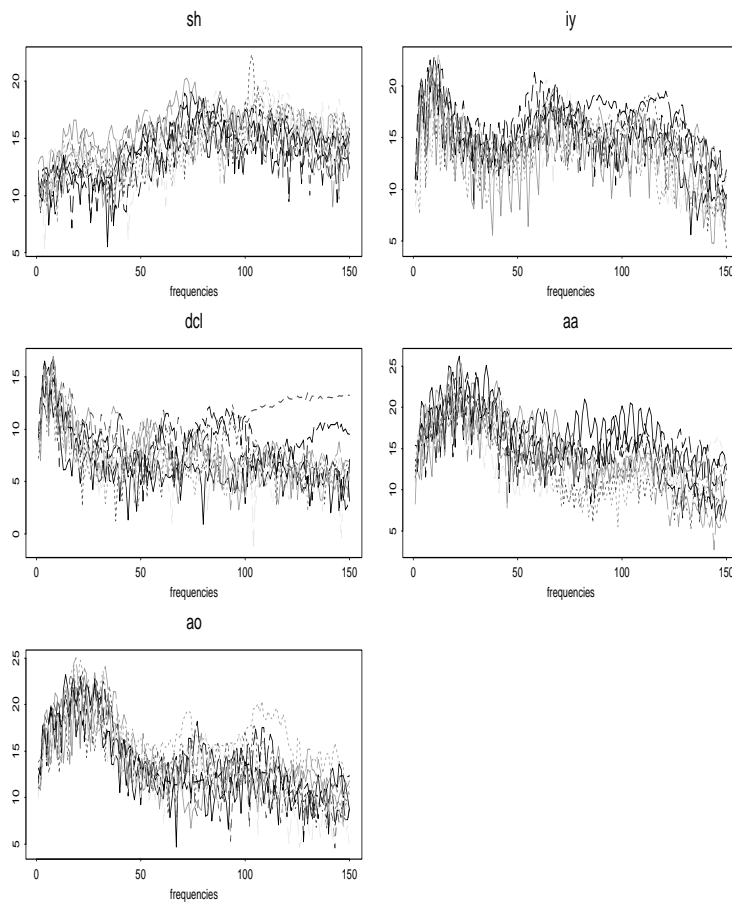
# EXEMPLE 1: Spectrométrie



$X = \text{Courbe Spectr.}$  $Y = \text{Taux de graisse}$  $Y_1, \dots, Y_n$ 

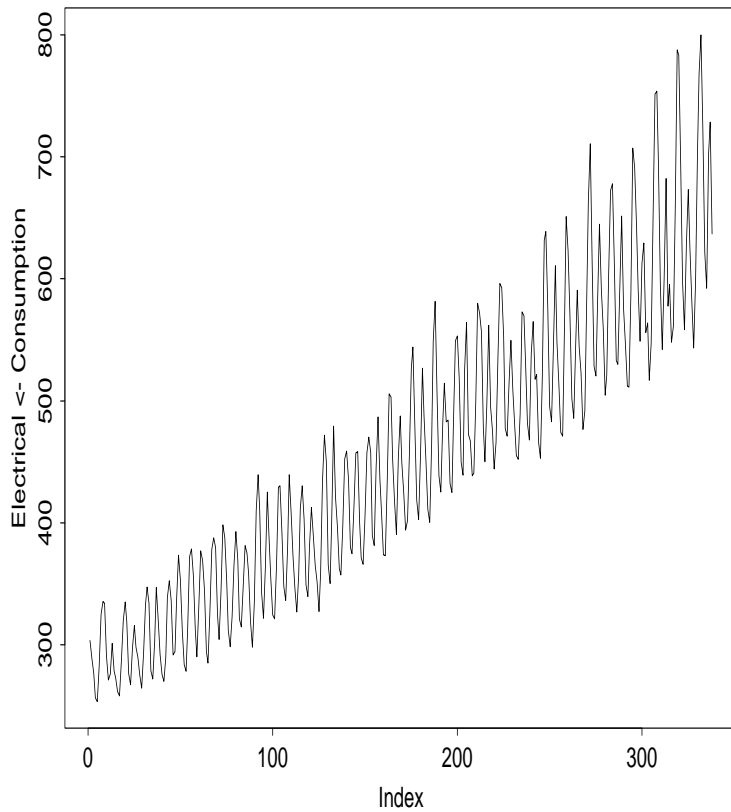
Peut-on lier (et estimer ce lien) la caractéristique chimique  $Y$  à la courbe spectrométrique  $X$ ?

## EXEMPLE 2: Reconnaissance vocale

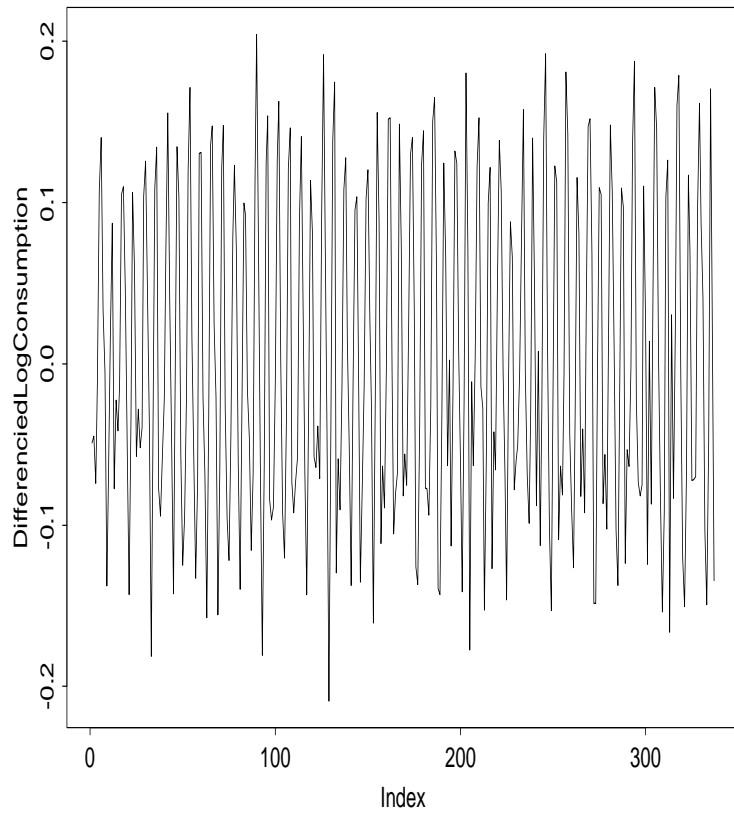


Peut-on retrouver un son à partir d'une courbe spectrométrique  $X$ ?

## EXEMPLE 3: Série temporelle

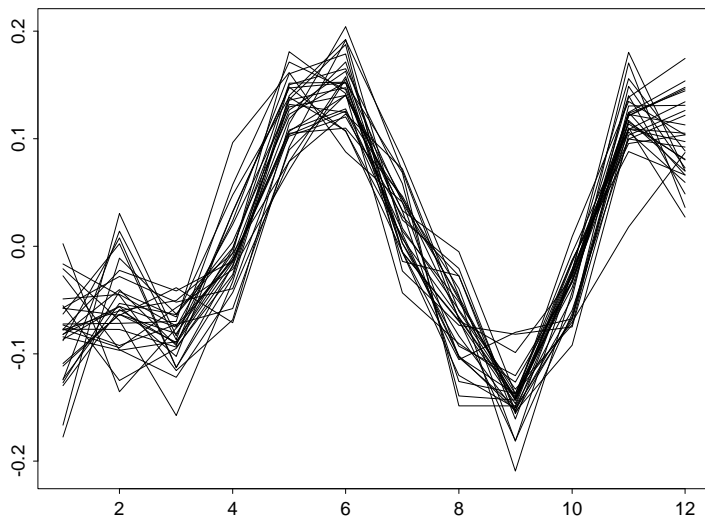


**DONNÉES BRUTES**



**DONNÉES "LOG-DIFFÉRENCIÉES"**

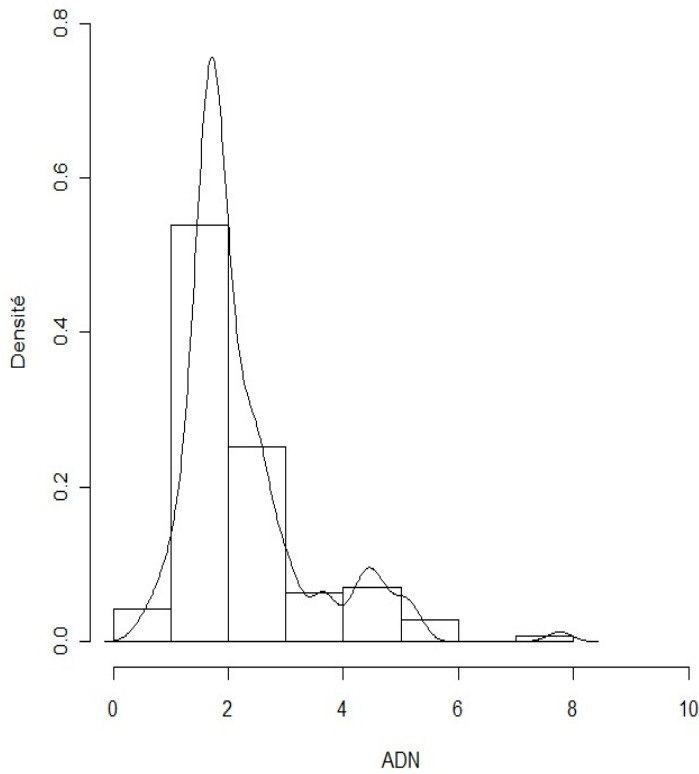




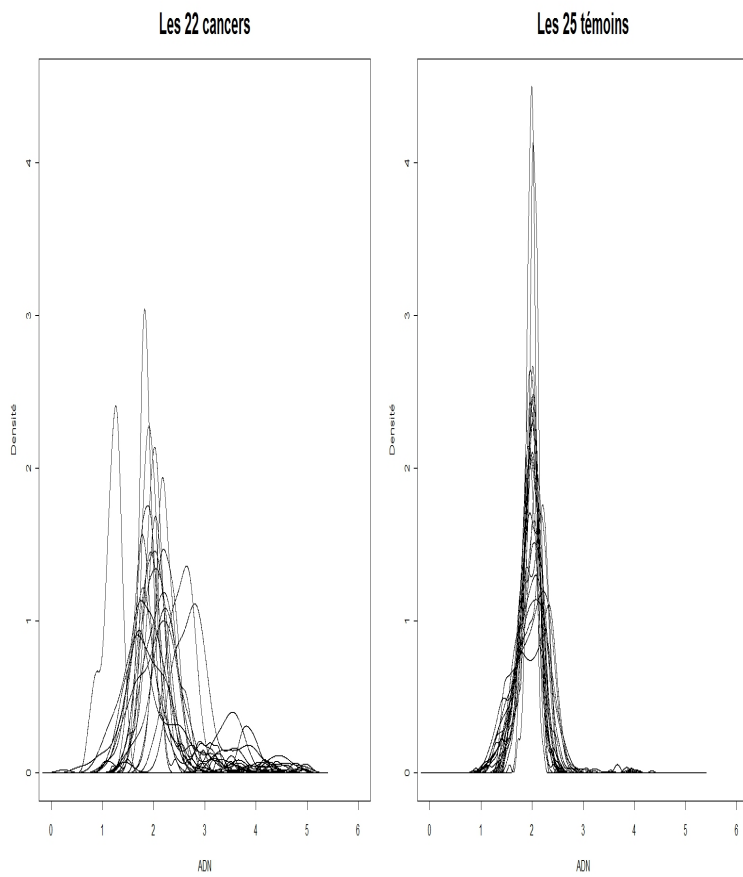
## DONNÉES "LOG-DIFFÉRENCIÉES" ET DÉCOUPÉES EN PÉRIODES

Peut-on prédire une caractéristique  $Y$  du futur à partir d'une période complète du passé  $X$ ?

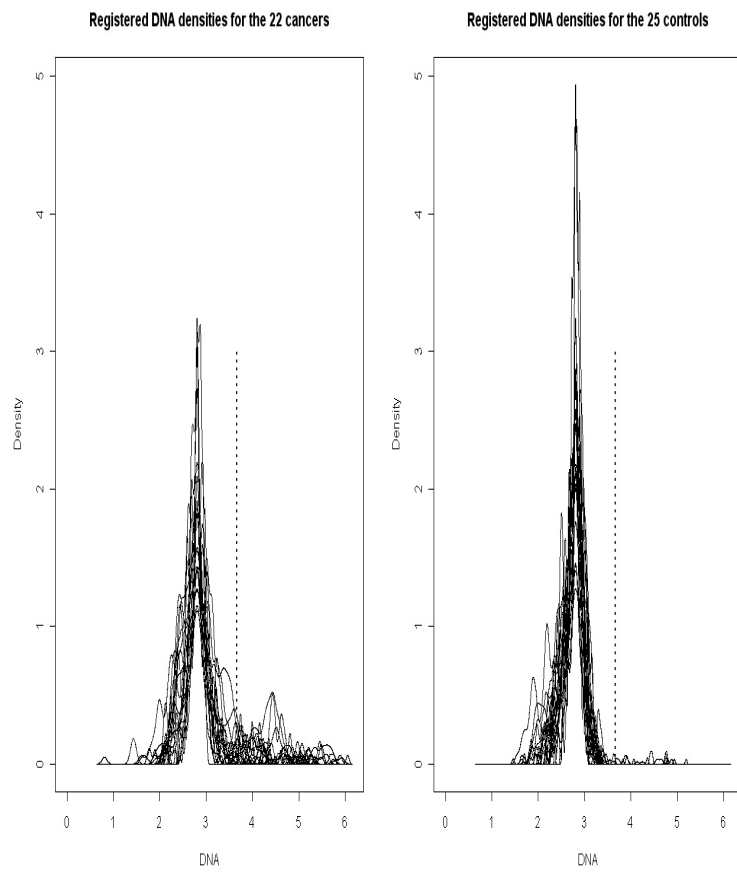
## EXEMPLE 4: Dépistage du cancer broncho-pulmonaire



**DONNÉES ADN POUR 1 INDIVIDU**



## COURBES DES DENSITÉS D'ADN



## COURBES D'ADN RECALÉES

# PLAN DE L'EXPOSÉ

## I. Introduction

## II. Quelques éléments historiques

- Statistique non-paramétrique uni-dimensionnelle
- Statistique non-paramétrique multi-dimensionnelle

## III. Statistique non-paramétrique fonctionnelle

- Un exemple d'estimateur
- Un exemple de résultat asymptotique

## IV. Quelques exemples applications

## V. Quelques éléments bibliographiques

# PLAN DE L'EXPOSÉ

## I. Introduction

## II. Quelques éléments historiques

→ Statistique non-paramétrique uni-dimensionnelle

→ Statistique non-paramétrique multi-dimensionnelle

## III. Statistique non-paramétrique fonctionnelle

→ Un exemple d'estimateur

→ Un exemple de résultat asymptotique

## IV. Quelques exemples applications

## V. Quelques éléments bibliographiques

# PLAN DE L'EXPOSÉ

## I. Introduction

## II. Quelques éléments historiques

→ Statistique non-paramétrique uni-dimensionnelle

→ Statistique non-paramétrique multi-dimensionnelle

## III. Statistique non-paramétrique fonctionnelle

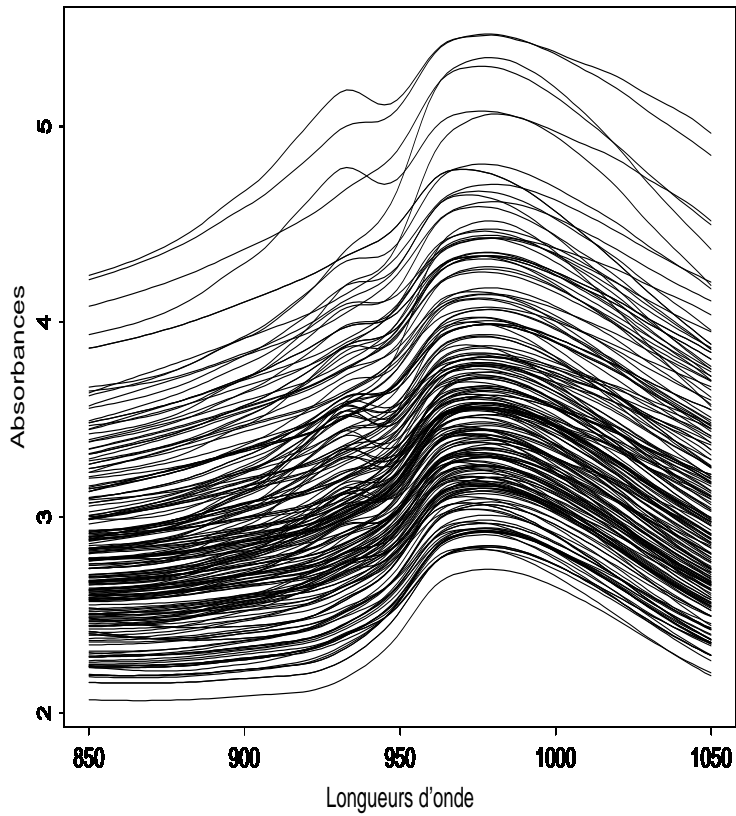
→ Un exemple d'estimateur

→ Un exemple de résultat asymptotique

## IV. Quelques exemples applications

## V. Quelques éléments bibliographiques

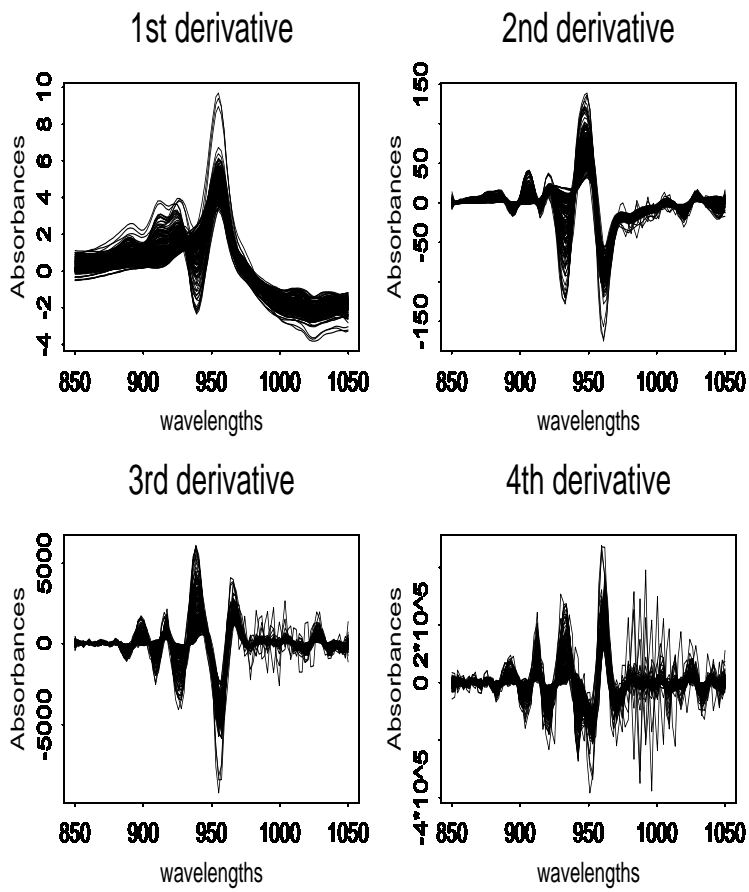
# EXEMPLE 1: Spectrométrie



## COURBES SPECTROMÉTRIQUES



# EXEMPLE 1: Spectrométrie



**DÉRIVÉES DES COURBES SPECTRO.**

## EXEMPLE 1: Spectrométrie

$$d(\mathbf{x}, \mathbf{y}) = \int (\mathbf{x}^{(\mathbf{q})}(\mathbf{t}) - \mathbf{y}^{(\mathbf{q})}(\mathbf{t}))^2 d\mathbf{t}$$

## EXEMPLE 1: Spectrométrie

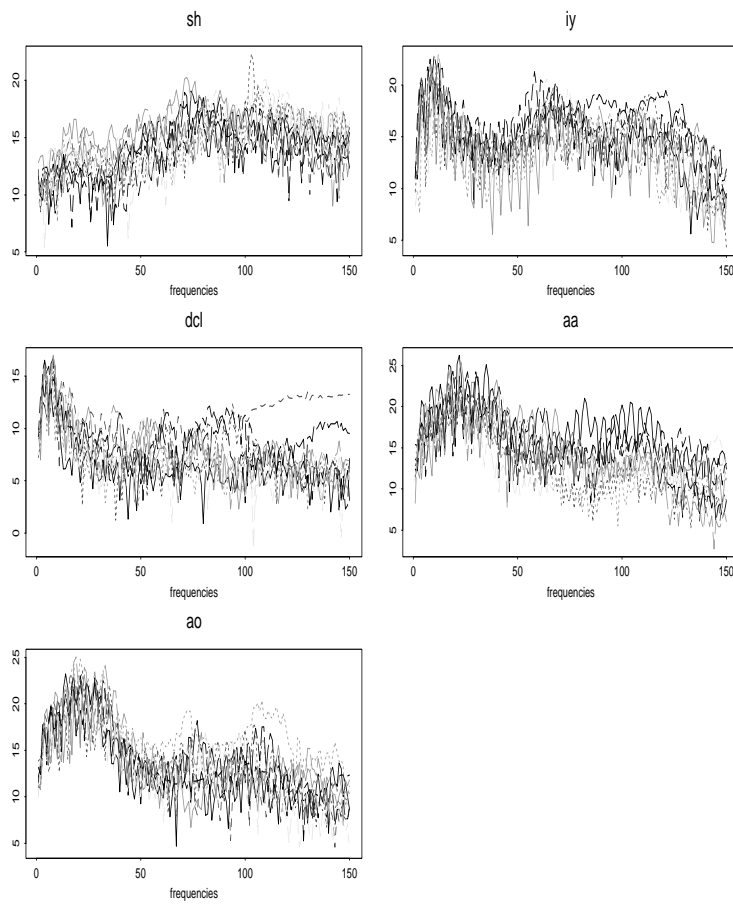
$$\mathbf{d}(\mathbf{x}, \mathbf{y}) = \int (\mathbf{x}^{(q)}(\mathbf{t}) - \mathbf{y}^{(q)}(\mathbf{t}))^2 \mathbf{d}\mathbf{t}$$

Résultats en terme de prédiction

$$q = 0 \rightarrow MSE = 70.7, \quad q = 1 \rightarrow MSE = 8.0$$

$$q = 2 \rightarrow MSE = 1.9, \quad q = 3 \rightarrow MSE = 5.9$$

## EXAMPLE 2: Reconnaissance vocale

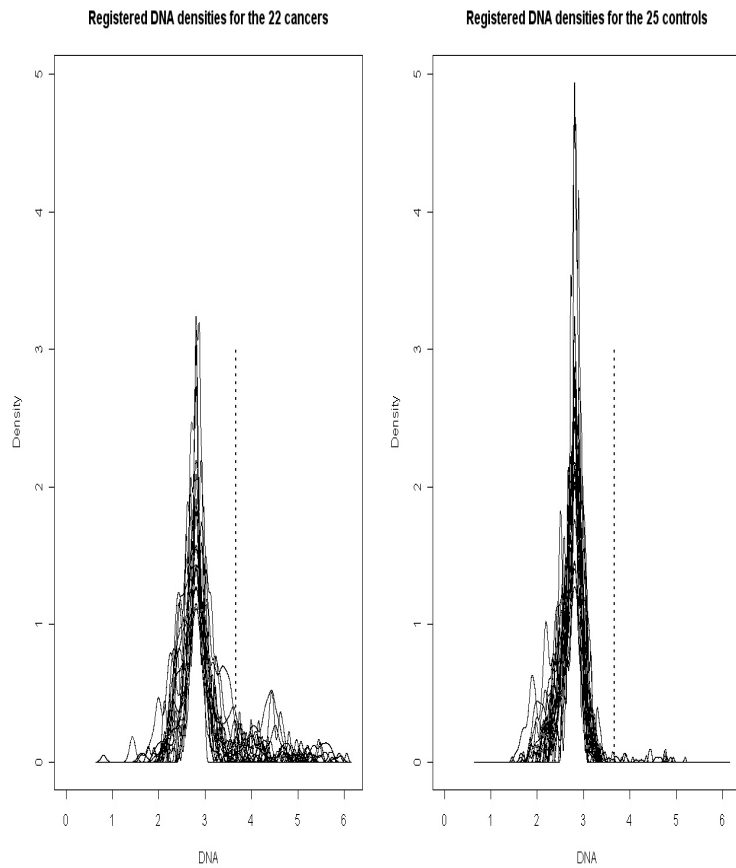


## EXEMPLE 2: Reconnaissance vocale

$$\mathbf{d}(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^q (\mathbf{x}_j - \mathbf{y}_j)^2,$$

$$\text{où } x = \sum_{j=1}^{+\infty} x_j e_j = \sum_{j=1}^{+\infty} \langle x, e_j \rangle e_j$$

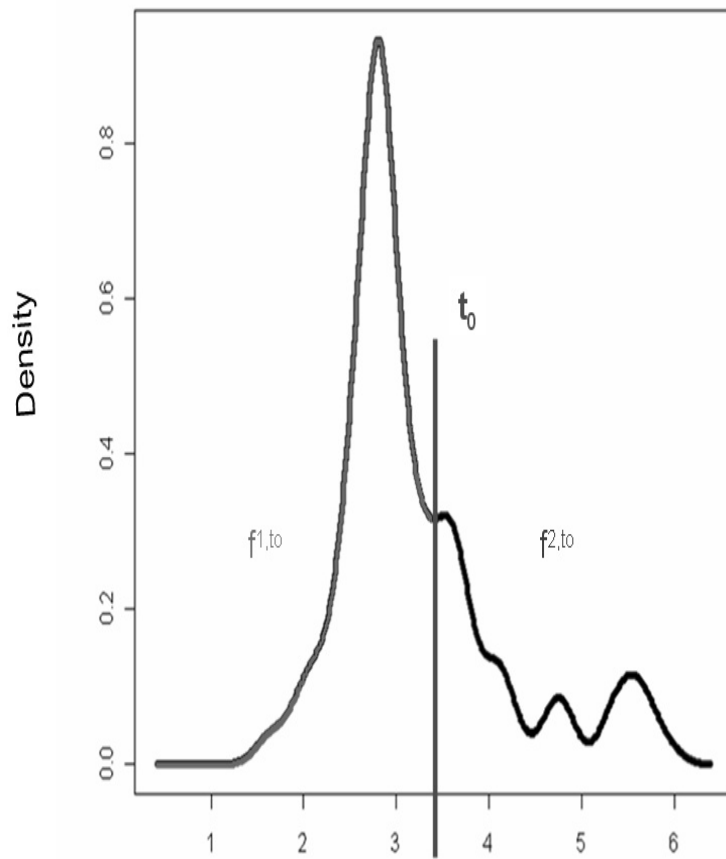
## EXEMPLE 4: Cancer broncho-pulmonaire



**COURBES D'ADN RECALÉES**

## EXAMPLE 4

$$d(\mathbf{x}, \mathbf{y}) = \int_a^b (\mathbf{x}_{\text{recalee}}(\mathbf{t}) - \mathbf{y}_{\text{recalee}}(\mathbf{t}))^2 d\mathbf{t}$$



## MORCEAU OPTIMAL D'UNE COURBE D'ADN



## Quelques familles de semi-métriques

$$\mathbf{d}(\mathbf{x}, \mathbf{y}) = \int (\mathbf{x}^{(\mathbf{q})}(\mathbf{t}) - \mathbf{y}^{(\mathbf{q})}(\mathbf{t}))^2 d\mathbf{t}$$

$$\mathbf{d}(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^{\mathbf{q}} (\mathbf{x}_j - \mathbf{y}_j)^2,$$

où  $x = \sum_{j=1}^{+\infty} x_j e_j = \sum_{j=1}^{+\infty} \langle x, e_j \rangle e_j$

$$\mathbf{d}(\mathbf{x}, \mathbf{y}) = \int_{\mathbf{a}}^{\mathbf{b}} (\mathbf{x}_{\text{recalee}}(\mathbf{t}) - \mathbf{y}_{\text{recalee}}(\mathbf{t}))^2 d\mathbf{t}$$

$$\mathbf{d}(\mathbf{x}, \mathbf{y}) = \dots$$

# PLAN DE L'EXPOSÉ

## I. Introduction

## II. Quelques éléments historiques

→ Statistique non-paramétrique uni-dimensionnelle

→ Statistique non-paramétrique multi-dimensionnelle

## III. Statistique non-paramétrique fonctionnelle

→ Un exemple d'estimateur

→ Un exemple de résultat asymptotique

## IV. Quelques exemples applications

## V. Quelques éléments bibliographiques

## LIVRES

Ramsay-Silverman (1997, 2002, 2005). Functional Data Analysis. Springer.

Ferraty-Vieu (2006). Nonparametric Functional Data Analysis. Springer.

Ferraty-Romain (2010). Handbook of FDA, Oxford Univ. Press.

## ACTES DE CONGRÈS

Ferraty (2011). Recent advances in FDA ... . Springer.

## NUMÉROS SPÉCIAUX

Davidian et al (2004). Statistica Sinica.

Gaonzalez Manteiga et Vieu (2007). Comp. Statist. Data Analysis.

Valderrama (2007). Computational Statistics.

Ferraty (2010). Journal of Multivariate Analysis.

**ÉVÉNEMENT À VENIR:** IWFOS 3ème Edition, Ialje, Juin 2014.

<http://www.math.univ-toulouse.fr/staph>