

## Oral UC2 Big Data

Joseph  
ALLYNDREE,  
Thomas  
CLERC, Louis  
LACOSTE,  
Tom  
LORTHIOS

Sélection des  
métabolites  
par la méthode  
Lasso

Infos sur les données  
de métabolisme

Infos sur les données  
de métabolisme

Matrices Y et X

Test de Blancheur

Structure du bruit des  
résidus

Sélection de variables

Exportation des  
Métabolites  
sélectionnés par GLM  
Lasso

# Oral UC2 Big Data

Joseph ALLYNDREE, Thomas CLERC, Louis LACOSTE,  
Tom LORTHIOS

Oral UC2 Big  
Data

Joseph  
ALLYNDREE,  
Thomas  
CLERC, Louis  
LACOSTE,  
Tom  
LORTHIOS

Sélection des  
métabolites  
par la méthode  
Lasso

Infos sur les données  
de métabolisme

Infos sur les données  
de métabolisme

Matrices Y et X

Test de Blancheur

Structure du bruit des  
résidus

Sélection de variables

Exportation des  
Métabolites  
sélectionnés par GLM  
Lasso

# Sélection des métabolites par la méthode Lasso

# Sélection des métabolites par la méthode Lasso

Oral UC2 Big  
Data

Joseph  
ALLYNDREE,  
Thomas  
CLERC, Louis  
LACOSTE,  
Tom  
LORTHIOS

Sélection des  
métabolites  
par la méthode  
Lasso

Infos sur les données  
de métabolisme

Infos sur les données  
de métabolisme

Matrices Y et X

Test de Blancheur

Structure du bruit des  
résidus

Sélection de variables

Exportation des  
Métabolites  
sélectionnés par GLM  
Lasso

Seuil défini à 0.7

Le chargement a nécessité le package : Matrix

Loaded glmnet 4.1-3

## Oral UC2 Big Data

Joseph  
ALLYNDREE,  
Thomas  
CLERC, Louis  
LACOSTE,  
Tom  
LORTHIOS

Sélection des  
métabolites  
par la méthode  
Lasso

Infos sur les données  
de métabolisme

Infos sur les données  
de métabolisme

Matrices Y et X

Test de Blancheur

Structure du bruit des  
résidus

Sélection de variables

Exportation des  
Métabolites  
sélectionnés par GLM  
Lasso

# Infos sur les données de métabolisme

# Infos sur les données de métabolisme

Oral UC2 Big  
Data

Joseph  
ALLYNDREE,  
Thomas  
CLERC, Louis  
LACOSTE,  
Tom  
LORTHIOS

Sélection des  
métabolites  
par la méthode  
Lasso

Infos sur les données  
de métabolisme

Infos sur les données  
de métabolisme

Matrices Y et X

Test de Blancheur

Structure du bruit des  
résidus

Sélection de variables

Exportation des  
Métabolites  
sélectionnés par GLM  
Lasso

```
[1] "Nombre de colonne dont la moyenne est nulle :"  
named integer(0)
```

```
[1] "Nombre de colonne où la variable est constante :"  
named integer(0)
```

```
[1] "Nombre de NAs dans le dataframe :"  
integer(0)
```

## Oral UC2 Big Data

Joseph  
ALLYNDREE,  
Thomas  
CLERC, Louis  
LACOSTE,  
Tom  
LORTHIOS

Sélection des  
métabolites  
par la méthode  
Lasso

Infos sur les données  
de métabolisme

**Infos sur les données  
de métabolisme**

Matrices Y et X

Test de Blancheur

Structure du bruit des  
résidus

Sélection de variables

Exportation des  
Métabolites  
sélectionnés par GLM  
Lasso

# Infos sur les données de métabolisme

# Infos sur les données de métabolisme

Oral UC2 Big  
Data

Joseph  
ALLYNDREE,  
Thomas  
CLERC, Louis  
LACOSTE,  
Tom  
LORTHIOS

Sélection des  
métabolites  
par la méthode  
Lasso

Infos sur les données  
de métabolisme

Infos sur les données  
de métabolisme

Matrices Y et X

Test de Blancheur

Structure du bruit des  
résidus

Sélection de variables

Exportation des  
Métabolites  
sélectionnés par GLM  
Lasso

- Nombre de colonne dont la moyenne est nulle : 0
- Nombre de colonne où la variable est constante : 0
- Nombre de NAs dans le dataframe : 0

## Oral UC2 Big Data

Joseph  
ALLYNDREE,  
Thomas  
CLERC, Louis  
LACOSTE,  
Tom  
LORTHIOS

Sélection des  
métabolites  
par la méthode  
Lasso

Infos sur les données  
de métabolisme

Infos sur les données  
de métabolisme

**Matrices  $Y$  et  $X$**

Test de Blancheur

Structure du bruit des  
résidus

Sélection de variables

Exportation des  
Métabolites  
sélectionnés par GLM  
Lasso

# Matrices $Y$ et $X$



# Matrices Y et X

Oral UC2 Big  
Data

Joseph  
ALLYNDREE,  
Thomas  
CLERC, Louis  
LACOSTE,  
Tom  
LORTHIOS

Sélection des  
métabolites  
par la méthode  
Lasso

Infos sur les données  
de métabolisme

Infos sur les données  
de métabolisme

**Matrices Y et X**

Test de Blancheur

Structure du bruit des  
résidus

Sélection de variables

Exportation des  
Métabolites  
sélectionnés par GLM  
Lasso

```
[1] "Lignes de X"
```

```
[1] 9
```

```
[1] "Colonnes de Y"
```

```
[1] 199
```

## Oral UC2 Big Data

Joseph  
ALLYNDREE,  
Thomas  
CLERC, Louis  
LACOSTE,  
Tom  
LORTHIOS

Sélection des  
métabolites  
par la méthode  
Lasso

Infos sur les données  
de métabolisme

Infos sur les données  
de métabolisme

Matrices  $Y$  et  $X$

**Test de Blancheur**

Structure du bruit des  
résidus

Sélection de variables

Exportation des  
Métabolites  
sélectionnés par GLM  
Lasso

# Test de Blancheur

# Test de Blancheur

Oral UC2 Big  
Data

Joseph  
ALLYNDREE,  
Thomas  
CLERC, Louis  
LACOSTE,  
Tom  
LORTHIOS

Sélection des  
métabolites  
par la méthode  
Lasso

Infos sur les données  
de métabolisme

Infos sur les données  
de métabolisme

Matrices Y et X

**Test de Blancheur**

Structure du bruit des  
résidus

Sélection de variables

Exportation des  
Métabolites  
sélectionnés par GLM  
Lasso

P-valeur du test de blancheur : 0.0303858

## Oral UC2 Big Data

Joseph  
ALLYNDREE,  
Thomas  
CLERC, Louis  
LACOSTE,  
Tom  
LORTHIOS

Sélection des  
métabolites  
par la méthode  
Lasso

Infos sur les données  
de métabolisme

Infos sur les données  
de métabolisme

Matrices Y et X

Test de Blancheur

**Structure du bruit des  
résidus**

Sélection de variables

Exportation des  
Métabolites  
sélectionnés par GLM  
Lasso

# Structure du bruit des résidus

# Structure du bruit des résidus

Oral UC2 Big  
Data

Joseph  
ALLYNDREE,  
Thomas  
CLERC, Louis  
LACOSTE,  
Tom  
LORTHIOS

Sélection des  
métabolites  
par la méthode  
Lasso

Infos sur les données  
de métabolisme

Infos sur les données  
de métabolisme

Matrices Y et X

Test de Blancher

Structure du bruit des  
résidus

Sélection de variables

Exportation des  
Métabolites  
sélectionnés par GLM  
Lasso

	Pvalue	Decision
AR1	0.035	NO WHITE NOISE
nonparam	0.741	WHITE NOISE
ARMA 2 1	0.029	NO WHITE NOISE

On voit donc que lorsque l'on applique la structure non paramétrique la valeur p de 0.741 nous indique que parmi les fonctions testées c'est la structure non paramétrique qui permet de blanchir la matrice.

## Oral UC2 Big Data

Joseph  
ALLYNDREE,  
Thomas  
CLERC, Louis  
LACOSTE,  
Tom  
LORTHIOS

### Sélection des métabolites par la méthode Lasso

Infos sur les données  
de métabolisme

Infos sur les données  
de métabolisme

Matrices  $Y$  et  $X$

Test de Blancheur

Structure du bruit des  
résidus

#### Sélection de variables

Exportation des  
Métabolites  
sélectionnés par GLM  
Lasso

# Sélection de variables

# Sélection de variables

Oral UC2 Big  
Data

Joseph  
ALLYNDREE,  
Thomas  
CLERC, Louis  
LACOSTE,  
Tom  
LORTHIOS

Sélection des  
métabolites  
par la méthode  
Lasso

Infos sur les données  
de métabolisme

Infos sur les données  
de métabolisme

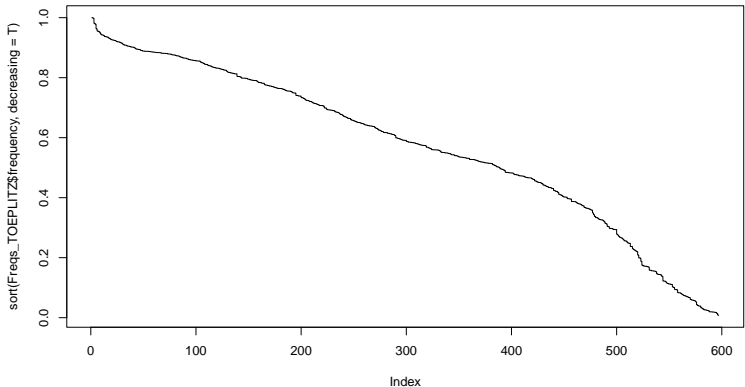
Matrices Y et X

Test de Blancheur

Structure du bruit des  
résidus

Sélection de variables

Exportation des  
Métabolites  
sélectionnés par GLM  
Lasso



# Les 50 métabolites les plus fréquents

Oral UC2 Big  
Data

Joseph  
ALLYNDREE,  
Thomas  
CLERC, Louis  
LACOSTE,  
Tom  
LORTHIOS

Sélection des  
métabolites  
par la méthode  
Lasso

Infos sur les données  
de métabolisme

Infos sur les données  
de métabolisme

Matrices Y et X

Test de Blancher

Structure du bruit des  
résidus

Sélection de variables

Exportation des  
Métabolites  
sélectionnés par GLM  
Lasso

1, 0.998, 0.9812, 0.9784, 0.9628, 0.9554, 0.9544, 0.9516,  
0.9458, 0.9428, 0.9422, 0.9402, 0.9372, 0.9366, 0.9356, 0.9346,  
0.9312, 0.9298, 0.9272, 0.9258, 0.9256, 0.924, 0.9238, 0.9206,  
0.9196, 0.9194, 0.9184, 0.9166, 0.9152, 0.912, 0.9098, 0.9084,  
0.9082, 0.9062, 0.905, 0.9048, 0.9036, 0.9022, 0.9018, 0.9016,  
0.9008, 0.897, 0.8954, 0.895, 0.8942, 0.894, 0.8914, 0.8906,  
0.8888, 0.8888

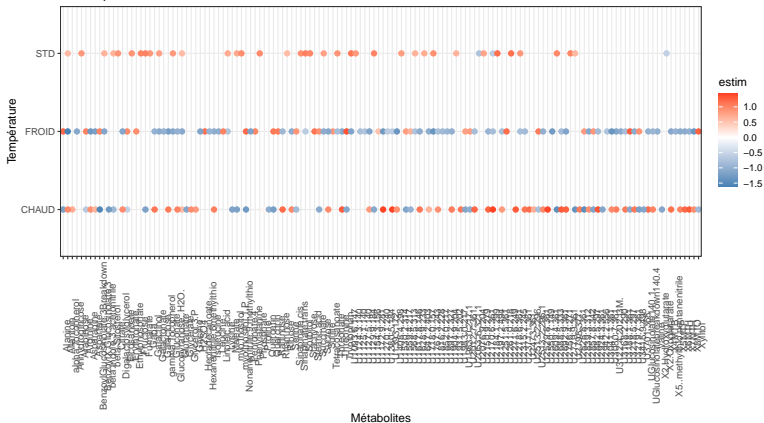


## Oral UC2 Big Data

Joseph  
ALLYNDREE,  
Thomas  
CLERC, Louis  
LACOSTE,  
Tom  
LORTHIOS

## Sélection de variables

Réponse des Métabolites sélectionnés pour les conditions de températures, au seuil 0.7



# Boxplots des réponses des métabolites dépassant le seuil 0.7

Oral UC2 Big Data

Joseph ALLYNDREE,  
Thomas CLERC, Louis LACOSTE,  
Tom LORTHIOS

Sélection des métabolites par la méthode Lasso

Infos sur les données de métabolisme

Infos sur les données de métabolisme

Matrices Y et X

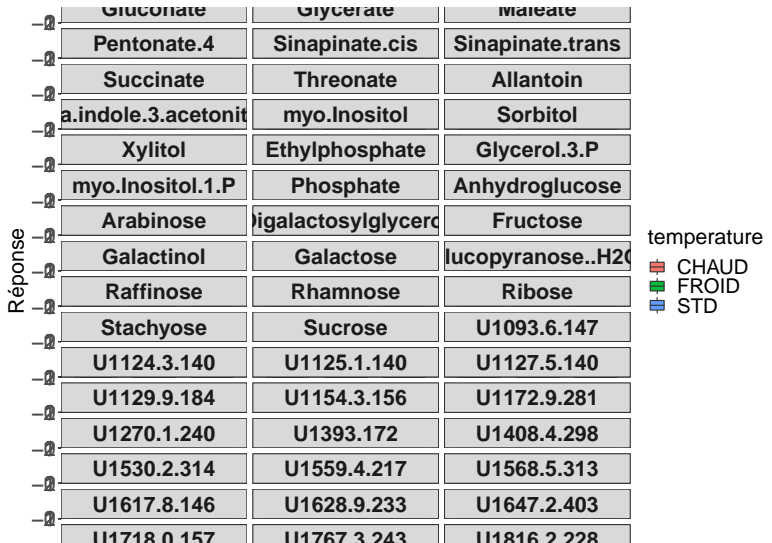
Test de Blancher

Structure du bruit des résidus

Sélection de variables

Exportation des Métabolites sélectionnés par GLM Lasso

Using temperature as id variables



## Oral UC2 Big Data

Joseph  
ALLYNDREE,  
Thomas  
CLERC, Louis  
LACOSTE,  
Tom  
LORTHIOS

Sélection des  
métabolites  
par la méthode  
Lasso

Infos sur les données  
de métabolisme

Infos sur les données  
de métabolisme

Matrices Y et X

Test de Blancheur

Structure du bruit des  
résidus

Sélection de variables

Exportation des  
Métabolites  
sélectionnés par GLM  
Lasso

# Exportation des Métabolites sélectionnés par GLM Lasso

# Exportation des Métabolites sélectionnés par GLM Lasso

Oral UC2 Big  
Data

Joseph  
ALLYNDREE,  
Thomas  
CLERC, Louis  
LACOSTE,  
Tom  
LORTHIOS

Sélection des  
métabolites  
par la méthode  
Lasso

Infos sur les données  
de métabolisme

Infos sur les données  
de métabolisme

Matrices Y et X

Test de Blancheur

Structure du bruit des  
résidus

Sélection de variables

Exportation des  
Métabolites  
sélectionnés par GLM  
Lasso

Fichier **metabolites\_selection\_lasso\_0.7.csv** exporté.