

# Marie Perrot-Dockès

Post-doctorante en Statistique

340 rue Saint-Jacques

75005 Paris

☎ 06 84 41 06 71

✉ [marie.perrot-dockes@upmc.fr](mailto:marie.perrot-dockes@upmc.fr)

🌐 <https://marie-perrotdockes.github.io/>

📍 [marie-perrotdockes](#)

## Formation

- 2016 – 2019 **Doctorat en mathématiques**, *AgroParisTech, INRA, Université Paris-Saclay*,  
**Options** : Mathématiques aux interfaces.  
Méthodes régularisées pour l'analyse de données multivariées en grande dimension : théorie et applications.
- 2012-2015 **Institut de Statistique de l'Université de Paris Spécialisation Biostatistiques**, *ISUP*, à l'UPMC(Paris, 75)  
Modèle mixte, modèle linéaire, données de survie, estimation fonctionnelle, bootstrap, théorie des probabilités, classification, techniques de rang, séries temporelles, C, R.
- 2012-2015 **Centre de la statistique appliquée à la médecine et à la biologie médicale**, *CESAM*,  
Statistique en recherche clinique, méthodologie statistique, épidémiologie.
- 2011-2012 **Classe préparatoire MPSI/MP**, au lycée Charlemagne (Paris, 75).

## Expériences professionnelles :

- 2019 – **Post-doctorante en statistique**, *Institut Mathématiques de Toulouse, Sorbonne université*,  
**Sous la direction de** : Etienne Roquain, Pierre Neuvial.
- Recherche de bornes de confiance posthoc dans des cadres de tests multiples sous dépendance
    - \* Tests multiples
    - \* Inférence *posthoc*
    - \* Modèle de Markov Caché
- 2016 – 2019 **Doctorante en mathématiques**, *AgroParisTech, INRA, Université Paris-Saclay*,  
**Encadrants** : Céline Lévy-Leduc, Julien Chiquet, Laure Sansonnet.
- Nouvelle méthode de sélection de variables dans le modèle linéaire multivarié avec prise en compte de la dépendance
    - \* Résultat théorique : consistance en signe de l'estimateur proposé
    - \* Résultat numérique : études des performances par expériences numériques
    - \* Application à des données de métabolomique et d'immunologie
    - \* Mise en place de package R
  - Méthodes d'estimation de matrice de covariance en grande dimension
    - \* Matrice par blocs
    - \* Matrice toeplitz
    - \* Mise en place de package R
- 2015 – 2016 **Biostatisticienne Institut Curie**, *Institut Curie*.
- Modélisation de données immunologiques
    - \* Régression linéaire pénalisée
    - \* Méthode Partial Least Square
    - \* Etude des interactions
    - \* Sélection de modèle
- 2014 - 2015 **Stage de M2 en alternance**, *Institut Curie*,  
**Encadrants** : Frédérique Berger, Vassili Soumelis.
- Modélisation statistique de la réponse des Lymphocytes Th au signaux des Cellules dendritiques
  - Rédaction d'un mémoire : *Utilisation de l'approche SPLS2 sur des données immunologiques corrélées et adaptation d'une méthode de classification biologique des interactions*
- 2014 **Stage de biostatistiques**, *Institut-Curie (Laboratoire d'Immunologie, INSERM U932 et l'équipe de Biostatistiques Cliniques, INSERM U900)*,  
**Encadrants** : Frédérique Berger, Vassili Soumelis.
- Recherche de profils cytokiniques des lymphocytes Th infiltrant les cancers du sein

---

## Enseignements

2019- **Polytech**, *niveau : 5<sup>e</sup> année.*

- Probabilités et Statistiques ( 30 heures de cours)
  - \* Espace de Probabilité
  - \* Variable aléatoire
  - \* Statistique inférentielle

2016 - 2019 **AgroParisTech**, *niveau : 1<sup>re</sup> et 2<sup>e</sup> années.*

- Modélisation mathématique et ses applications : statistique ( 33 heures de TD par an)
  - \* Estimation par maximum de vraisemblance
  - \* Intervalle de confiance
  - \* Test paramétrique
  - \* Régression linéaire simple
- Base du modèle linéaire ( 18 heures de TP par an)
  - \* Introduction à R
  - \* Régression linéaire avec R

---

## Communication orales

Avril 2017 **Rencontres des jeunes statisticiens**, *Porquerolles*,

Méthode de sélection de variable dans le modèle linéaire généralisé : Application aux données de métabolomique.

Juin 2017 **Journées de la Statistique**, *Avignon*,

Modèle linéaire multivarié parcimonieux avec estimation de covariance : une application à des données de métabolomique.

décembre **Journées NETBIO**, *Paris*,

2017 Méthode de sélection de variables dans le modèle linéaire multivarié avec prise en compte de la dépendance : application à des données de métabolomique.

janvier 2018 **SMPGD**, *Montpellier*,

Variable selection in multivariate linear models with high-dimensional covariance matrix estimation .

Mars 2018 **Les sommets de Rochebrune**, *Rochebrune*,

Variable selection in multivariate linear models with high-dimensional covariance matrix estimation.

Juin 2018 **Journées de la Statistique**, *Saclay*,

Sélection de variables dans le modèle linéaire multivarié en grande dimension avec prise en compte de la dépendance.

Octobre 2018 **Atelier StateOfTheR**, *Paris*,

Fancy plotting with ggraph, ganimate.

Octobre 2018 **Regional Workshop of European young researchers in statistics**, *Paris*,

Variable selection in multivariate linear models taking into account the dependence .

Juillet 2019 **UseR**, *Toulouse*,

VariSel : An R package to perform variable selection in multivariate linear model.  
Or how a Gallic village stay irreductible

Novembre **Journées des doctorants et des Post-Doctorants**, *Toulouse*,

2019 Regularized methods to study multivariate data in high dimensional settings : theory and applications..

Novembre **Séminaire AgroCampusOuest**, *Rennes*,

2019 Méthodes régularisées pour l'analyse de données multivariées en grande dimension : théorie et applications..

---

## Production scientifique

### Articles publiés

- 2018 **Journal of Multivariate Analysis**, *M. Perrot-Dockès, C. Lévy-Leduc, L. Sansonnet, J. Chiquet*, Variable selection in multivariate linear models with high-dimensional covariance matrix estimation, 166, 78 – 97.
- 2018 **Statistical Applications in Genetics and Molecular Biology**, *M. Perrot-Dockès, C. Lévy-Leduc, J. Chiquet, L. Sansonnet, M. Brégère, M.-P. Étienne, S. Robin, G. Genta-Jouve*, A variable selection approach in the multivariate linear model : An application to LC-MS metabolomics data, 17(5).
- 2019 **OncoImmunology**, *L Faucheux\*, M Grandclaudon\*, M Perrot-Dockès\*, P Sirven, F Berger, AS Hamy, V Fourchette, A Vincent-Salomon, F Mechta-Grigoriou, F Reyal, and others*, A multivariate Th17 metagene for prognostic stratification in T cell non-inflamed triple negative breast cancer, 8, 1 – 13.  
\* : ces auteurs ont contribué de manière égale à cette publication
- 2019 **Cell**, *M. Grandclaudon\*, M. Perrot-Dockès\*, C. Trichot\*, O. Mostafa-Abouزيد, W. Abou-Jaoudé, F. Berger, P. Hupé, D. Thieffry, L. Sansonnet, J. Chiquet, C. Lévy-Leduc, V. Soumelis*, A quantitative multivariate model of human dendritic cell-T helper cell, 179, 432 – 447.  
\* : ces auteurs ont contribué de manière égale à cette publication

### Article soumis

- 2019 **Soumis**, *M. Perrot-Dockès, C. Lévy-Leduc and Rajjou, Loïc*, Estimation of large block structured covariance matrices : Application to " multi-omic " approaches to study seed quality.

### Chapitre de livre (à paraître prochainement)

- 2019 **Wiley**, *M. Perrot-Dockès, C. Lévy-Leduc, G. Cueff and L Rajjou*, Sélection de variables dans le modèle linéaire général en grande dimension : application à des approches “multi-omiques” pour l’étude de la qualité des graines.

### Package R

- 2019 **CRAN**, *M. Perrot-Dockès, C. Lévy-Leduc and J Chiquet*, MultiVarSel : Variable Selection in a Multivariate Linear Model, version 1.1.3.
- 2019 **CRAN**, *M. Perrot-Dockès and C. Lévy-Leduc*, BlockCov : Estimation of Large Block Covariance Matrices, version 0.1.1.
- 2019 **GitHub**, *M. Perrot-Dockès and J Chiquet*, VariSel, version 0.1.0.

---

## Compétences Informatique et Langues

Langues Français, Anglais  
Langage R  
Outils  $\LaTeX$ , github, SVN  
OS Linux : Ubuntu

---

## Divers

Association Membre du groupe jeune de la SFdS