



HAL
open science

Les écoles d'astrostatistique ” Statistics for Astrophysics ”

Julyan Arbel, Didier Fraix-Burnet, Stéphane Girard

► **To cite this version:**

Julyan Arbel, Didier Fraix-Burnet, Stéphane Girard. Les écoles d'astrostatistique ” Statistics for Astrophysics ”. CFIES 2017 - 5ème Colloque Francophone International sur l'Enseignement de la Statistique, Sep 2017, Grenoble, France. hal-01583854

HAL Id: hal-01583854

<https://inria.hal.science/hal-01583854>

Submitted on 8 Sep 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

LES ÉCOLES D'ASTROSTATISTIQUE "STATISTICS FOR ASTROPHYSICS"

Julyan Arbel¹, Didier Fraix-Burnet² & Stéphane Girard¹

¹ *Univ. Grenoble Alpes, Inria, LJK, F-38000 Grenoble, France,
julyan.arbel@inria.fr, stephane.girard@inria.fr*

² *Univ. Grenoble Alpes, CNRS, IPAG, F-38000 Grenoble, France,
didier.fraix-burnet@univ-grenoble-alpes.fr*

Résumé. L'astrophysique est une science physique basée sur l'observation, s'appuyant sur des instruments de détection toujours plus gigantesques et perfectionnés. La quantité de données recueillies impose des traitements statistiques fiables et efficaces. L'objectif des écoles d'astrostatistique est de donner les compétences nécessaires aux participants astrophysiciens pour entreprendre par eux-mêmes des analyses statistiques. Depuis 2013, les trois sessions de cette école ont été consacrées successivement aux méthodes de régression, classification, puis aux approches bayésiennes. Nous présentons dans cette communication les retours de notre expérience d'organisateur.

Mots-clés. Astrostatistique, Formation continue.

Abstract. Astrophysics is a physical science based on observations through ever bigger and more sophisticated detection tools. The amount of data collected requires reliable and effective statistical modeling. These astrostatistic schools aim at providing the astrophysicist participants with the required skills in order to implement statistical analyses by themselves. Since 2013, the three editions of the school were successively dedicated to regression methods, to classification, and to the Bayesian approach. In this talk, we present feedback of the lessons learnt as organizers.

Keywords. Astrostatistics, Continuing education.

1 Introduction

L'astrophysique est une science physique basée sur l'observation. Les instruments de détection sont toujours plus gigantesques et perfectionnés, la quantité de données recueillies impose à l'évidence des traitements fiables et efficaces. Les problématiques physiques sont pointues, les objets étudiés très complexes, de sorte que les analyses et interprétations nécessitent des méthodologies sophistiquées. Sans entrer dans les détails, tous les domaines de l'astrophysique sont concernés par une question statistique sous-jacente, que ce soit au niveau du détecteur (traitement du signal), de l'information recueillie (traitement d'image), de l'extraction d'une quantité physique (bruits, distributions) et de leur exploitation (analyses de données, fouille de données, bases aux millions d'entrées, réduction de dimension, classification multivariée, ajustements de modèles, tests statistiques, etc).

Les astrophysiciens ont donc un énorme besoin de compétences en statistique, et ce de plus en plus, mais ne font pas suffisamment appel aux statisticiens et restent donc trop souvent avec des méthodes un peu dépassées voire inadaptées. Cependant, on sent depuis quelques années un engouement certain, notamment chez les jeunes chercheurs qui sont souvent bien seuls, pour découvrir et apprendre les techniques appropriées. Par ailleurs, les problématiques astrophysiques sont un défi passionnant pour la recherche en statistique en lui offrant des données souvent uniques. D'un côté des méthodes éprouvées existent que les astrophysiciens ignorent, d'un autre côté des méthodes sont à tester et améliorer, voire inventer, sans lesquelles les astrophysiciens ne pourraient extraire l'information pertinente de leurs observations.

Il existe donc à l'évidence un intérêt commun à débrider cette thématique interdisciplinaire qui concerne également d'autres communautés comme l'informatique, le traitement du signal et de l'image, ainsi que les systèmes complexes. Suivant cette ambition, nous avons mis en place depuis 2013 une série d'écoles d'astrostatistique.

2 Objectifs et mise en œuvre

L'objectif des écoles d'astrostatistique est de donner les compétences nécessaires aux participants pour entreprendre par eux-mêmes des analyses statistiques. Les sessions de 2013 et 2015 ont été consacrées respectivement aux méthodes de régression et classification, la session de 2017 sera dédiée aux approches bayésiennes. Nous insistons à chaque édition sur l'importance d'établir des collaborations entre astrophysiciens et statisticiens pour des développements algorithmiques et méthodologiques souvent requis par la spécificité des données astrophysiques et se situant au-delà des compétences des astrophysiciens, même en ayant participé à cette école. Les deux sessions précédentes de cette école nous permettent de proposer un format stable et efficace, à savoir :

- les intervenants sont statisticiens, les participants astrophysiciens ;

- chaque session comporte cinq journées et se concentre sur un thème précis de la statistique ;
- une session est décomposée en séances d'une demi-journée, moitié cours moitié TP ;
- des temps de discussions sont réservés autour de problématiques concrètes amenées par les participants ;
- la publication de comptes-rendus est assurée, un exemplaire étant envoyé à chaque participant.

Autant que cela est possible, les exemples et travaux pratiques s'effectuent en partie sur des données astrophysiques. Le logiciel utilisé (R ou Python) pour tous les travaux pratiques est unique pour toute la session, il fait l'objet d'une séance d'initiation en début d'école et fait donc aussi partie de l'apport pédagogique. La publication des cours et travaux pratiques sous forme de comptes-rendus nous paraît remplir trois rôles :

- i) fournir un support pérenne aux participants,
- ii) assurer une diffusion et une visibilité large des sessions,
- iii) exiger un niveau élevé de pédagogie et de présentation que les seules diapositives ne peuvent atteindre.

Les comptes-rendus précédents ont été publiés sous format papier, voir Fraix-Burnet et Valls-Gabaud (2014), Fraix-Burnet et Girard (2016). Pour plus de détails, nous renvoyons également le lecteur aux sites internet de chaque édition :

<https://ecastrostat2013.sciencesconf.org>

<https://stat4astro2015.sciencesconf.org>

<https://stat4astro2017.sciencesconf.org>

Bibliographie

- [1] *Statistics for Astrophysics: Clustering and Classification*, 2016. D. Fraix-Burnet & S. Girard (eds), EDP Sciences, EAS Publications Series 77 (DOI:978-2-7598-9001-9).
- [2] *Statistics for Astrophysics: Methods and Applications of the Regression*, 2014. D. Fraix-Burnet & D. Valls-Gabaud (eds), EDP Sciences 66 (DOI:978-2-7598-1729-0).