

Ngoc Dung Nguyen

Curriculum Vitae

 Vietnamese
 Department of Statistical Sciences,
University of Padova, Italy
 via Cesare Battisti 241,
35121 Padova, Italy
 (+39) 338 770 5721
 ngocdung.nguyen@unipd.it
 [NgocDung-NGUYEN.github.io](https://github.com/NgocDung-NGUYEN)



Academic Position

2022–2023 **Postdoctoral Research Fellow**, Department of Statistical Sciences, University of Padova, Italy.
Research Supervisor: [Professor Alberto Roverato](#).
Topic: Structure learning of multiple graphical models for dependent samples.

Tertiary Education

2018–2021 **Ph.D. in Statistics**, Department of Statistical Sciences, University of Padova, Italy.
Advisor: [Professor Alberto Roverato](#).
Thesis title: Model selection for colored graphical models for paired data.

2017–2018 **Master of Science, Technology and Health**, Université de Tours, France, GPA: 16.796/20.
Master 2 Program of Mathematics. Major in Applied Mathematics. Mention: “Très Bien”.
Advisor: [Professor Romain Abraham](#).
Thesis title: Limits or large random graphs.

2013–2017 **Bachelor of Science**, Vietnam National University-Ho Chi Minh University of Science, Ho Chi Minh City, Vietnam, GPA: 8.21/10.
Program in Mathematics and Computer Science. Major in Probability and Statistics. Mention: Very Good.
Advisor: [Professor Dang Duc Trong](#).
Thesis title: Multiplicative censoring models.

Research Interests

Statistical learning Supervised and unsupervised learning; model inference, e.g. bootstrap methods, Bayesian methods; high-dimensional problems, e.g. regularization, principal component analysis (PCA), etc.

Applied Statistics Statistical inference, statistical modelling, regression analysis and diagnostics, model selection on model spaces, e.g. greedy search.

Graphical Models Structure learning and model search for (undirected) graphical models, e.g. graphical Lasso, model selection of colored graphical models for paired data, inferring the distributions of quantities of interests given the observations.

Biostatistics Statistical methods and models for large and complex biological data sets, inference for biological/genetic networks.

Conference, Seminar, and Workshop Presentations

Jul. 2019 [Data Research Camp](#) of the Department of Statistical Sciences of the University of Padova supported by the excellence program [Statistical methods and models for complex data](#) in San Servolo Island, Venice, Italy.

Professional Memberships

- 2021–2023 [Institute of Mathematical Statistics \(IMS\)](#).
2018 [COST \(European Cooperation in Science and Technology\) Association](#).

Fellowships and Awards

- 2022 Research grant for the Postdoctoral Fellowship in Statistics in the University of Padova, Italy.
2018–2021 Ph.D. scholarship in Statistical Sciences funded by University of Padova, Italy.
2016–2017 Scholarship of the National Program for the Development of Mathematics 2010–2020 of [Vietnam Institute for Advanced Study in Mathematics \(VIASM\)](#).
2016–2017 Award for excellent students, Department of Mathematics and Computer Science, Vietnam National University–Ho Chi Minh University of Science.

Professional Experiences

Department of Statistical Sciences Paolo Fortunati, University of Bologna, Italy

- Aug.-Oct. ***a short-term research contract for three months.***
2021
 - Objectives of the assignment: (1) develop of efficient algorithms to carry out greedy searches in the space of graphical models with symmetries; (2) implementations of the algorithms through functions to be developed in the statistical environment R; (3) application to a set of real data.

Department of Statistical Sciences, University of Padova, Italy

- 2018–2021 ***Ph.D. student in Statistics.***
 - First year: completed the foundation courses on Functional Analysis, Probability Theory, Programming Methodologies, Theory and Methods of Inference, Statistical Models and specialised short courses and seminars.
 - Second and third years: research and thesis preparation; Statistical Consulting course.***Ph.D. research: Model selection for colored graphical models for paired data.***
 - Advisor: [Professor Alberto Roverato](#) at the Department of Statistical Sciences, University of Padova.
 - Objectives: develop a computationally efficient model search algorithm for the space of colored graphical models for paired data:
 - Review of the literature concerning model selection criteria for the tuning parameter of the graphical lasso (gLasso) algorithm, and analyse the relationship existing between the degree of sparsity of the graph and value taken by the tuning parameter.
 - Get the ideas of partial orders, lattices, graph colorings, RCON models, model inclusion orders and review the structural properties of the lattice structure of models spaces such as the completeness, distributivity, etc.
 - Identify the family of colored graphical models for paired data and investigate the associated structure of the model space, and show that, under model inclusion order, it forms a complete non-distributive lattice.
 - Introduce a novel partial order for the class of colored graphical models for paired data, called by the *twin order*, which would be feasible in practice and easily implemented in model selection.
 - Investigate the structure properties of model space of graphical models for paired data and show that under the twin order, the family of colored graphical models for paired data forms a complete distributive lattice.
 - Adapt the idea of greedy search method to construct a backward elimination stepwise procedure on the twin lattice, which can exploit the structural property of the twin lattice into the model search and respect the principle of coherence.
 - Perform the numerical experiments for the proposed model selection procedure on the twin lattice for the simulated data sets, and show that this is more efficient than the similar approach on the model inclusion lattice and it can be applied with higher variables.
 - Implement the backward elimination stepwise procedure on the twin lattice on the graph through functions developed in the statistical environment R to investigate the dynamic activity of each brain region between two hemispheres through changes in the blood oxygenation level dependent (BOLD) signal via functional MRI data.

Faculty of Mathematics and Computer Science, Vietnam National University-Ho Chi Minh University of Science, Vietnam.

2017–2018 **Master in Applied Mathematics.**

- First stage: completed foundation courses covering a broad range of topics in Probability and Statistics, Optimization, Functional and Numerical Analysis, and Computer Science.
- Second stage: three-month internship in France.

Institut Denis Poisson, Université d'Orléans, France.

Apr.-Jul. **Master Thesis: Limits for large random graphs.**

2018 ○ Advisor: [Professor Romain Abraham](#) in the Institut Denis Poisson, Université d'Orléans, France.

○ Objectives:

- Review of random and large graphs, graph homomorphisms and homomorphism densities.
- Study the limiting object, called by *graphon*, which is a continuum graph where nodes are the elements of $[x, y]$ and the edge between two nodes $x, y \in [0, 1]$ is present with a symmetric probability function $W(x, y) [0, 1]^2$ to $[0, 1]$, and distances between two graphons.
- Make precisely the notion of convergences of growing graphs toward these graphons and exhibit some convergence criteria through the simulation experiments for four types of growing graphs: the growing uniform attachment graphs, the growing ranked attachment graphs, the growing prefix attachment graphs, and the preferential attachment graph on n fixed nodes.

○ Grade: 17/20.

Vietnam National University–Ho Chi Minh University of Science, Vietnam.

2016–2017 **Bachelor Thesis: Multiplicative censoring model.**

- Advisor: [Professor Dang Duc Trong](#) in the Faculty of Mathematics and Computer Science, Vietnam National University-Ho Chi Minh University of Science, Vietnam.
- Abstract: study the nonparametric estimation using orthogonal projection based on the Laguerre basis of both the density f and the corresponding survival function of X'_i 's on the multiplicative censoring model $Y_i = X_i U_i, 1 \leq i \leq n$ where U_i 's are multiplicative noise with uniform density on $[1 - a, 1 + a]$ for some $a \in (0, 1)$. Only $(Y_i)_{1 \leq i \leq n}$ are observed, and $(U_i)_{1 \leq i \leq n}$ and $(X_i)_{1 \leq i \leq n}$ are independent sequences. Then simulation experiments are implemented by R with different scenarios of distributions of X and tuning parameters in estimation for the dimension of the Laguerre basis.
- Grade: 10/10.

Selected Academic Experiences

Oct. 5th-27th **Attended a short course "Statistical aspects of Deep Neural Networks" (28 hours),**

2021 *Programme in Economics, Statistics and Data Science, University of Milano-Bicocca, Italy.*

Invited lecturer: [Professor Omiros Paspaliopoulos](#), Bocconi University, Milano, Italy.

Objectives:

- Covers the foundations of deep neural networks with a special emphasis and priority on more statistical aspects of this research agenda, e.g. high-dimensional regression, the random features models, from PCA to autoencoders, optimisation for neural networks, from deep neural networks to stochastic processes, regularisation, stability and adversarial training, connection between simulation-based inference and neural networks, etc.
- Give an overview of software and implementations in Pytorch.

Sep. 2020 **Attended a course "Regression Modelling with Large Data Sets" (10 hours), Padova, Italy.**

Invited lecturer: [Professor Ioannis Kosmidis](#), Department of Statistics at the University of Warwick and Turing Fellow of The Alan Turing Institute, UK.

Objectives:

- Introduce core theory, methods and software tools for tackling regression problems that involve large data sets in terms of number of observations or in terms of explanatory variables.
- Identify least squares as a core optimisation problem for fitting linear and generalised linear models, and classify the various methods for its solution or regularised versions of it in terms of complexity, memory usage and accuracy.

Apr.-Jun. 2020 **Attended Statistical Consulting course about the group's title "Predicting the economic effects of COVID-19 and of the lockdown measures"**.

Instructors: [Antonio Canale](#), [Bruno Scarpa](#), Department of Statistical Sciences, University of Padova, Italy.

Objectives:

- Investigate the effects and the consequences of the deadly virus COVID-19 outbreak and of the political decisions imposed by governments on the level of economic activities and financial market indices through a scientific approach.
- Apply statistical models on collected time series data to obtain reliable measures of the impact of the outbreak on economy and finance.
- Perform sensitive analysis to consider the possible consequences on different scenarios, such as the scenario of a second wave of Corona infections, with the purpose of helping policymaker to make the right decisions on society and economy in future situations.

Languages

Vietnamese Mother tongue

English Fluent

Excellent in reading, listening, good at writing and speaking.

Italian Basic

Basic words and phrases.

Computer Skills

Programming Languages **Advanced R, Python, MATLAB**

Operating Systems **Ubuntu, macOS**

Referees

Alberto Roverato

Full Professor of Statistical Sciences

University of Padova, Italy

✉ alberto.roverato@unipd.it

Romain Abraham

Full Professor in Institut Denis Poisson

Université d'Orléans, France

✉ Romain.Abraham@univ-orleans.fr

Dang Duc Trong

Professor of Probability and Statistics,

Faculty of Mathematics and Computer Science

University of Science, VNUHCM, Vietnam

✉ ddtrong@hcmus.edu.vn