

PIETRO CORETTO
CURRICULUM VITAE ET STUDIORUM

Thursday 29th June, 2023

PERSONAL DATA

- **Address:** University of Salerno, Department of Economics and Statistics; Via Giovanni Paolo II, no.132, 84084, Fisciano (SA), Italy
- **Contacts:** e-mail: pcoretto@unisa.it; phone: (+39)-089-96.22.13; fax: (+39)-089-96.20.49
- **Citizenship:** Italian
- **Language:** Italian (mother tongue), English (fluent), French (basic)
- **Position:** Associate professor in statistics

EDUCATION

- **2004–2008:** Ph.D. in Statistical Sciences, University College London (UCL, UK). Primary supervisor: Christian Hennig (Department of Statistical Science, UCL). Secondary Supervisor: A. Philip Dawid (Department of Pure Mathematics and Mathematical Statistics, University of Cambridge)
- **2003–2004:** Master of Sciences in Economics, University of Pennsylvania (Philadelphia, Usa)
- **2002–2003:** Master in Economics and Finance (MEF), University of Napoli “Federico II” (Italy)
- **1998–2002:** Laurea in Economia e Commercio (B.A. in Economics and Business Studies), University of Salerno (Italy). Final score: summa cum laude with distinction. Thesis supervisor: Prof. Marco Pagano

HONORS, PRIZES AND SCHOLARSHIPS

- **2018:** gold medal for the “*Best Teaching Award*”, MEF/LMEF graduate program in Economics and Finance, University of Naples “Federico II”
- **2009:** nominated for the “*Classification Society Distinguished Dissertation Award*” by the IFCS (International Federation of Classification Societies)
- **2007:** “*Costas Goutis Prize*”. For the best research work leading to the transfer to PhD registration. Awarded by the research committee of the Department of Statistical Sciences, University College London (UCL)
- **2006–2007:** “*Graduate School Scholarship*”, from University College London (UCL)
- **2003–2005:** “*Giovanna Crivelli*” scholarship from Unicredito Italiano Spa
- **2002:** “*Angelo Costa Prize*” (VI edition, International). Best of the papers selected by an international scientific committee comprising R. Solow (Nobel 1987), R. Mundell (Nobel 1999), P. Diamond, et al.
- **2002:** Banca Nazionale delle Comunicazioni Prize (for high quality BA dissertations selected in Italy). From B.N.C. Foundation
- **2002–2003:** Scholarship from Italian Society for Investment Professionals (ISIP)

PROFESSIONAL EXPERIENCES AND RESEARCH PROJECTS

- **2006–present:** Principal investigator in FARB (local) research projects
- **2013–2017:** Research collaborator, project “*A multicriterion approach for cluster validation*” (EP/K033972/1) and funded by Engineering and Physical Sciences Research Council (EPSRC, UK). Funded 98,024 GBP.
- **2012–2015:** Research collaborator, project (PRIN 2010J3LZEN) “*Forecasting economic and financial time series: understanding the complexity and modeling structural change*”, founded by Ministero dell’Istruzione, dell’Università e della Ricerca (MIUR). Funded: 506,134 Euro.
- **2009–2012:** Research collaborator, project “*High Performance Computing*”, funded by University of Salerno (ASSA098434)
- **2008–2011:** Research collaborator, project “*Evolutionary Computation in Statistics*” (PRIN 20077P5AWA), founded by Ministero dell’Istruzione, dell’Università e della Ricerca (MIUR). Funded 52,315 Euro.
- **2007–2009:** Research collaborator, project “*Economic consequences of the demographic transition in Europe*” (INTERLINK II04C0ECL72007). Partners: Institute of Fiscal Studies (IFS, London), Tilburg University, Goethe University Frankfurt.
- **2006:** Consultancy for Cirrus Capital, London, UK
- **2005–2006:** Participation to the project: “*Quantitative Analysis of Financial Markets*”, University of Salerno
- **2002–2002:** Research assistant for the international project “*Financial Market Integration, Corporate Financing and Economics Growth*”. Funded by the European Commission

SERVICE TO PROFESSION

Local management duties

- **2022–present:** scientific director of the “D2LAB” computing laboratory, University of Salerno
- **2018–present:** member of the committee for student orientation, degree in *Statistics for Big Data*, University of Salerno
- **2012–present:** member of the academic board of the PhD program in *Economics and Policy of Markets and Firms*, University of Salerno
- **2012–present:** member of the editorial committee; Department of Economics and Statistics, University of Salerno
- **2018–2019:** member of the local committee for the project “Piano Nazionale Lauree Scientifiche” (PLS, founded by MIUR)

Editorial Activity. Referee for (in alphabetical order): *Advances in Classification and Data Analysis*, *BMC Bioinformatics*, *Bioinformatics*, *Biometrika*, *Biostatistics*, *Communications in Statistics*, *Computational Statistics and Data Analysis*, *Electronic Journal of Statistics*, *IEEE Transactions on Cybernetics*, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *Journal of Computational and Graphical Statistics*, *Journal of Machine Learning Research*, *Journal of Multivariate Analysis*, *Journal of the Korean Statistical Society*, *Machine Learning*, *Mathematical and Computational Applications*, *Metron*, *Statistical Methods and Applications*, *Statistical Modelling*, *Statistical Papers*, *Statistics and Computing*.

Organization of Scientific Events

- **2023:** member of organizing committee of the “*14-th Scientific Meeting Classification and Data Analysis Group*” (CLADAG-2023); organized by SIS-CLADAG; September 11–13; Salerno, Italy
- **2018:** member of organizing committee of the “*International Summer School on Clustering, Data Analysis and Visualization of Complex Data*”; organized by ERS-IASC, ECAS and SIS-CLADAG; May 21–25; Catania, Italy

- **2018:** member of the local organizing committee for the “*4th Conference of the International Society for Nonparametric Statistics*” (ISNPS2018), June 11–15 2018, Salerno, Italy
- **2016:** member of the local organizing committee for the “*48th Meeting of the Italian Statistical Society*” (SIS2016), June 8–10 2016, Salerno, Italy
- **2014:** member of the local organizing committee for the “*6th International Conference on Mathematical and Statistical Methods for Actuarial Sciences and Finance*” (MAF2014), April 22–24, 2014, Vietri sul Mare (SA) - Italy
- **2010:** member of the local organizing committee for the “*4th International Conference on Mathematical and Statistical Methods for Actuarial Sciences and Finance*” (MAF2010), April 7–9, 2010, Ravello (SA) - Italy

CONFERENCES, WORKSHOPS AND SEMINARS

- Invited talk at the *17th conference of the International Federation of Classification Societies*, 19-23 July 2022, Porto, Portugal
- Invited talk at the *European Conference on Data Analysis* (ECDA2022), September 14–16, 2022, Napoli, Italy
- Invited talk at the *Workshop on Model-Based Clustering and Classification* (MBC²-2022), August 31–September 2, 2022, Catania, Italy
- Invited talk at the *14th International Conference of the ERCIM WG on Computational and Methodological Statistics* (CMStatistics 2021) December 18–20, 2021, London, UK
- Seminar at Department of Statistics “P. Fortunati”, University of Bologna, April 1, 2021, Bologna, Italy
- Co-organizer of the special conference session “*Time series clustering*”, Conference: *13-th Scientific Meeting, Classification and Data Analysis Group* (CLADAG 2021), September 9–11, 2021, Firenze, Italy
- Invited talk at the *International Conference on Mathematical and Statistical Methods for Actuarial Sciences and Finance* (MAF2020), April 15–17, 2020, Geneva, Switzerland
- Invited talk at the *13th International Conference of the ERCIM WG on Computational and Methodological Statistics* (CMStatistics 2019), December 19–21, 2020, London, UK
- Organizer of the conference session: “*Cluster validation and the selection of a cluster*”, conference: *12-th Scientific Meeting Classification and Data Analysis Group* (CLADAG 2019). September 11–13, 2019, Cassino, Italy
- Invited talk at the *12-th Scientific Meeting Classification and Data Analysis Group* (CLADAG 2019), September 11–13, 2019, Cassino, Italy
- Co-organizer of the conference session: “*Nonparametric inference for complex data*”, conference: *Smart Statistics for Smart Applications* (SIS 2019). June 18–21, 2019, Milan, Italy
- Invited talk at the *Workshop on Model-Based Clustering and Classification* (MBC²-2018), September 5–7, 2018, Catania, Italy
- Invited talk at the *Conference of International Federation of Classification Societies* (IFCS-2017), August 8–10, 2017, Tokyo, Japan
- Invited talk at “*Statistics and Data Science: New challenges, new generations*” (SIS 2017), Conference of the Italian Statistical Society (SIS), June 28-30, 2017, Firenze, Italy
- Invited talk at the *Workshop on Model-Based Clustering and Classification* (MBC2), Università di Catania, Italy, September 2016
- Discussant for the invited session: “*Clustering in time series*”, at the *10th Scientific Meeting of the CLADAG Group of the Italian Statistical Society* (CLADAG 2015), Santa Margherita di Pula, Italy, October 2015
- Invited talk at the *10th Scientific Meeting of the CLADAG Group of the Italian Statistical Society* (CLADAG 2015), Santa Margherita di Pula, Italy, October 2015
- Invited talk at the *2nd Workshop PRIN 2010J3LZEN group*, Università Bologna, Italy, November 2015

- Invited talk at the *7th International Conference of the ERCIM WG on Computational and Methodological Statistics* (ERCIM 2014), University of Pisa, Italy, December 2014
- Participation to the *Workshop on Model-Based Clustering and Classification* (MBC2), University of Catania, Italy, September 2014
- Invited talk the *1st Workshop PRIN 2010J3LZEN*, University of Tor Vergata, Rome, Italy, June 2014
- Invited talk at the *47th Scientific meeting Italian Statistical Society* (SIS 2014), Cagliari, Italy, June 2014
- Invited talk at the *6th International Conference of the ERCIM WG on Computing and Statistics* (ERCIM 2013), University of London, London, UK, December 2013
- Invited talk at the *46th Scientific meeting Italian Statistical Society* (SIS 2012) Rome, Italy, June 2012
- Invited talk at the *4th International Conference of the ERCIM WG on Computing and Statistics* (ERCIM 2011), University of London, London, UK, December 2011
- Invited talk at the *International conference on Evolutionary Computation in Statistics*, University of Tor Vergata, Roma, Italy, June 2011
- Invited talk the *3rd International Conference of the ERCIM WG on Computing and Statistics* (ERCIM 2010), University of London, London, UK, December 2010
- Invited talk at the *International Conference on Mathematical and Statistical Methods for Actuarial Sciences and Finance* (MAF10), Ravello, Italy, April 2010
- Invited talk at the *International conference on Evolutionary Computation in Statistics*, University Ca' Foscari, Venezia, Italy, March 2010
- Invited talk at the *7th Scientific Meeting of the CLADAG Group of the Italian Statistical Society*, Catania, Italy, September 2009
- Invited talk a the *JSM-ASA, Joint Statistical Meeting of the American Statistical Association*, Washington DC, USA, August 2009
- Invited talk at the *Model-Based Clustering Summer Session Workshop*, Institut Henri Poincaré, Paris, France, July 2009
- Seminar at Dpto. Estadística e Investigación Operativa, Universidad de Valladolid, Valladolid, Spain, November 2008.
- Invited talk a the *1st Workshop of the ERCIM Working Group on Computing & Statistics* (ERCIM 2008), Neuchâtel, Switzerland, June 2008
- Invited talk at the *1st joint meeting of the Société Francophone de Classification and the Classification and Data Analysis Group* (CLADAG 2008) of the Italian Statistical Society, Caserta, Italy, June 2008
- Invited talk at the *International Conference on Mathematical and Statistical Methods for Actuarial Sciences and Finance* (MAF08), Venice, Italy, March 2008
- Invited talk at the *6th Scientific Meeting of the Classification and Data Analysis Group of the Italian Statistical Society* (CLADAG 2007), Macerata, Italy, September 2007
- Participation to the *31st Annual Conference of the German Classification Society on Data Analysis, Machine Learning, and Applications*, Freiburg, Germany, March 2007
- Seminar a the Department of Statistical Sciences, University College London, London, UK, September 2006
- Participation to the *29th Research Students' Conference in Probability and Statistics* (RSC 2006), Glasgow, UK, April 2006
- Participation to the *28th Research Students' Conference in Probability and Statistics* (RSC 2005), Cambridge, UK, April 2005

TEACHING

Current courses

- ECONOMETRICS (graduate); *Master in Economics and Finance* (MEF), level II master degree, University of Naples Federico II (Since 2016)
- ADVANCED STATISTICAL LEARNING (graduate); degree in *Statistics for Finance/Big Data Analytics*; University of Salerno (Since 2020)
- DATA SCIENCE (graduate); PhD in *Economics and Policy of Markets and Firms*, University of Salerno (Since 2015)
- STATISTICAL LEARNING (undergraduate); degree in *Statistics for Big Data*; University of Salerno (Since 2020)
- DATA ANALYSIS AND VISUALIZATION (undergraduate); degree in *Statistics for Big Data*; University of Salerno (Since 2018)

Past courses

- **2018:** lectures on ADVANCED TOPICS IN CLUSTER ANALYSIS (graduate level) at the “*International Summer School on Clustering, Data Analysis and Visualization of Complex Data*”; organized by ERS-IASC, ECAS and SIS-CLADAG; Catania (Italy), May 21-25
- **2009–2022** COMPUTATIONAL STATISTICS (undergraduate); degree in *Management and Computer Science*; University of Salerno
- **2009–2020** STATISTICS (undergraduate); degree in *Management and Computer Science*; University of Salerno
- **2013–2019** PROBABILITY AND INFERENCE (graduate); *Master in Economics, Finance and Risk Management* (MEFiRM), level I master degree, University of Salerno
- **2013–2014:** MATHEMATICAL STATISTICS (graduate level); PhD in “Economics, Law and Statistics for Policy, Markets, and Firms”, University of Salerno
- **2008–2013:** MATHEMATICAL STATISTICS (graduate level); PhD in “Engineering and Economics of Innovations”, University of Salerno
- **2008–2009:** SAMPLING THEORY AND METHODS FOR BUSINESS APPLICATIONS (undergraduate level); Faculty of Economics, University of Salerno
- **2006–2008:** COMPUTATIONAL STATISTICS (graduate level); PhD in “Engineering and Economics of Innovations”, University of Salerno
- **2006–2008:** DATA ANALYSIS (undergraduate level); Faculty of Economics, University of Salerno
- **2008:** STATISTICS (undergraduate level); GO.TER, II level Master (equivalent to MA), University of Salerno

RESEARCH PUBLICATIONS

Papers in journals (with peer-review)

- Coretto, P. and C. Hennig (2010). A simulation study to compare robust clustering methods based on mixtures. *Advances in Data Analysis and Classification*, vol. 4(2): pages 111–135. URL: <https://www.doi.org/10.1007/s11634-010-0065-4>.
- Coretto, P. and C. Hennig (2011). Maximum likelihood estimation of heterogeneous mixtures of gaussian and uniform distributions. *Journal of Statistical Planning and Inference*, vol. 141(1): pages 462–473. URL: <https://www.doi.org/10.1016/j.jspi.2010.06.024>.
- Coretto, P. (2013). Discussion (invited) of: How to find an appropriate clustering for mixed-type variables with application to socio-economic stratification; by C. Hennig and T. F. Liao. *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, vol. 62: pages 346–347. URL: <https://doi.org/10.1111/j.1467-9876.2012.01066.x>.

- Coretto, P. and C. Hennig (2016). Robust improper maximum likelihood: tuning, computation, and a comparison with other methods for robust gaussian clustering. *Journal of the American Statistical Association*, vol. 111(516): pages 1648–1659. URL: <https://www.doi.org/10.1080/01621459.2015.1100996>.
- Coretto, P. and C. Hennig (2017). Consistency, breakdown robustness, and algorithms for robust improper maximum likelihood clustering. *Journal of Machine Learning Research*, vol. 18(142): pages 1–39. URL: <http://jmlr.org/papers/v18/16-382.html>.
- Coretto, P. and F. Giordano (2017). Nonparametric estimation of the dynamic range of music signals. *Australian & New Zealand Journal of Statistics*, vol. 59(4): pages 389–412. URL: <https://www.doi.org/10.1111/anzs.12217>.
- Serra, A., P. Coretto, M. Fratello, and R. Tagliaferri (2018). Robust and sparse correlation matrix estimation for the analysis of high-dimensional genomics data. *Bioinformatics*, vol. 34(4): pages 625–634. URL: <https://www.doi.org/10.1093/bioinformatics/btx642>.
- Coretto, P., A. Serra, and R. Tagliaferri (2018). Robust clustering of noisy high-dimensional gene expression data for patients subtyping. *Bioinformatics*, vol. 34(23): pages 4064–4072. URL: <https://www.doi.org/10.1093/bioinformatics/bty502>.
- Serra, A., S. Önlü, P. Coretto, and D. Greco (2019). An integrated quantitative structure and mechanism of action-activity relationship model of human serum albumin binding. *Journal of Cheminformatics*, vol. 11(38): pages 1–10. URL: <https://www.doi.org/10.1186/s13321-019-0359-2>.
- Giordano, F. and P. Coretto (2020). A monte carlo subsampling method for estimating the distribution of signal-to-noise ratio statistics in nonparametric time series regression models. *Statistical Methods & Applications*, vol. 29(3): pages 483–514. URL: <https://www.doi.org/10.1007/s10260-019-00487-5>.
- Coretto, P. (2020). Discussion (invited) of: Multiple-systems analysis for the quantification of modern slavery: classical and bayesian approaches; by Bernard W. Silverman. *Journal of the Royal Statistical Society, Series A (Statistics in Society)*, vol. 183(3): pages 725–726. URL: <https://www.doi.org/10.1111/rssa.12505>.
- Coretto, P., M. La Rocca, and G. Storti (2020). Improving many volatility forecasts using cross-sectional volatility clusters. *Journal of Risk and Financial Management*, vol. 13(4): pages 1–23. URL: <https://www.doi.org/10.3390/jrfm13040064>.
- Storti, G. and P. Coretto (2020). Discussion (invited) of: Linear mixed effects models for non-gaussian continuous repeated measurement data; by O. Asar, D. Bolin, P. J. Diggle and J. wallin. *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, vol. 69(5): page 46. URL: <https://doi.org/10.1111/rssc.12405>.
- Hennig, C. and P. Coretto (2022). An adequacy approach for deciding the number of clusters for OTRIMLE robust Gaussian mixture-based clustering. *Australian & New Zealand Journal of Statistics*, vol. 64(2): pages 230–254. URL: <https://doi.org/10.1111/anzs.12338>.
- Coretto, P. (2022). Estimation and computations for Gaussian mixtures with uniform noise under separation constraints. *Statistical Methods & Applications. Journal of the Italian Statistical Society*, vol. 31(2): pages 427–458. URL: <https://doi.org/10.1007/s10260-021-00578-2>.
- Coraggio, L. and P. Coretto (2023). Selecting the number of clusters, clustering models, and algorithms. A unifying approach based on the quadratic discriminant score. *Journal of Multivariate Analysis*, vol. 196: pages Paper No. 105181, 20. URL: <https://doi.org/10.1016/j.jmva.2023.105181>.

Papers in book collections (with peer-review)

- Hennig, C. and P. Coretto (2008). The noise component in model-based cluster analysis. In C. Preisach, H. Burkhardt, L. Schmidt-Thieme, and R. Decker, editors, *Data Analysis, Machine Learning and Applications*, pages 127–138. Springer Berlin Heidelberg, Berlin, Heidelberg.

- Coretto, P. and M. L. Parrella (2010). Empirical likelihood based nonparametric testing for capm. In M. Corazza and C. Pizzi, editors, *Mathematical and Statistical Methods for Actuarial Sciences and Finance*, pages 103–112. Springer Milan, Milano.
- Coretto, P., M. La Rocca, and G. Storti (2011). Group structured volatility. In S. Ingrassia, R. Rocci, and M. Vichi, editors, *New Perspectives in Statistical Modeling and Data Analysis*, pages 329–335. Springer, Heidelberg Dordrecht London New York.
- Coppola, G. and P. Coretto (2012). Classificazione del territorio provinciale. In A. VV., editor, *Analisi delle caratteristiche strutturali e funzionali delle imprese industriali operanti in provincia di Salerno. Settore Alimentare e Chimico Plastico*, pages 43–68. Edizioni Scientifiche Italiane, Napoli.
- Amendola, A., P. Coretto, and G. Storti (2013). Indagine campionaria. In AA.VV., editor, *Analisi delle Caratteristiche Strutturali e Funzionali delle Imprese Industriali Operanti in Provincia di Salerno: Settore Costruzioni*, pages 39–65. Edizioni Scientifiche Italiane, Napoli.
- Coretto, P., M. La Rocca, and G. Storti (2021). A garch-type model with cross-sectional volatility clusters. In M. Corazza, M. Gilli, C. Perna, C. Pizzi, and M. Sibillo, editors, *Mathematical and Statistical Methods for Actuarial Sciences and Finance*, pages 169–174. Springer International Publishing.

Papers in conference proceedings (with peer-review)

- Coretto, P. and C. Hennig (2007). Choice of the improper density in robust improper ML for finite normal mixtures. In *Classification and Data Analysis 2007*, pages 439–442. Springer, Hildeberg.
- Coretto, P. (2008). Robust mixture modelling. In *Joint Meeting of the SFC and CLADAG*, pages 69–72. Edizioni Scientifiche Italiane, Napoli.
- Coretto, P. and C. Hennig (2009). Robust model-based clustering using an improper constant density selected by a distance optimizatin for non-outliers. In *JSM Proceedings, IMS General Theory Section*, pages 3847–3858. American Statistical Association, Alexandria, VA.
- Coretto, P., M. La Rocca, and G. Storti (2009). Group structured volatility. In *Book of Short Papers CLADAG09 Catania*, pages 119–122. CLEUP, Padova.
- Coretto, P. (2012). Adaptive robust location–scale estimation. In S. I. di Statistica, editor, *Proceedings of the XLVI Scientific Meeting*, pages 1–4. Cleup, Roma.
- Coretto, P. and C. Hennig (2014). Robust covariance matrix estimation with regularization. In S. I. di Statistica, editor, *Proceedings of the 47th Scientific Meeting of the Italian Statistical Society*, pages 1–8. CUEC, Cagliari.
- Coretto, P. and C. Hennig (2015). Robust model-based clustering with covariance matrix constraints. In *10th meeting of Classification and Data Analysis Group*, pages 1–5. CUEC Editrice, Cagliari.
- Serra, A., P. Coretto, and R. Tagliaferri (2017). On the noisy high-dimensional gene expression data analysis. In A. Petrucci and R. Verde, editors, *Statistics and data science: new challenges, new generations*, pages 919–926. Firenze University Press, Firenze.
- Coretto, P., A. Serra, and R. Tagliaferri (2019). Noise resistant clustering of high-dimensional gene expression data. In *Cladag 2019 - Book of Short Papers*, pages 132–135. EUC, Cassino.
- Coraggio, L. and P. Coretto (2019). Likelihood-type methods for comparing clustering solutions. In *Cladag 2019 - Book of Short Papers*, pages 120–123. EUC, Cassino,.
- Coretto, P. and C. Hennig (2021). Non-parametric consistency for the Gaussian mixture maximum likelihood estimator. In *Cladag 2021 - Book of short Papers*, pages 116–119. Firenze University Press, Firenze.
- Coretto, P. and L. Coraggio (2021). In-sample and cross-validated likelihood-type criteria for clustering selection. In *Book of Short Papers of the 5th international workshop on Models and Learning for Clustering and Classification*, pages 27–32. Ledizioni, Milano.