RESEARCH PUBLICATIONS

Pietro Coretto

Thursday 29th June, 2023

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.