# Matteo Sesia

Bridge Hall 401H, Marshall School of Business, University of Southern California, Los Angeles, CA 90089
sesia@marshall.usc.edu — `https://msesia.github.io`

## ACADEMIC POSITIONS

**Assistant Professor** (tenure-track). June 2020–present. Department of Data Sciences and Operations, Marshall School of Business, University of Southern California.

## EDUCATION

**Ph.D. in Statistics**, 2020. Stanford University. Advisor: Emmanuel Candès.
Thesis title: New methods for variable importance testing with applications to genetic studies.

**M.S. in Physics of Complex Systems**, 2015. Politecnico di Torino & Université Paris-Sud (joint).
Graduated *cum laude* (top final grade and highest honors).

**M.A. in Statistics and Applied Mathematics**, 2015. Collegio Carlo Alberto.
Graduated *with distinction* (highest honors).

**B.S. in Engineering Physics**, 2013. Politecnico di Torino.
Graduated *cum laude* (top final grade and highest honors).

## TEACHING

University of Southern California:

BUAD 310: Applied Business Statistics (undergraduate), Fall 2020, Fall 2021.

DSO 621: Research Forum (PhD), Spring 2021 (contributed).

Stanford University:

Stats 390: Consulting Workshop (graduate), Summer 2018.

Stats 195: Introduction to R (undergraduate), Spring 2018, Spring 2020.

## PUBLICATIONS AND PREPRINTS

Publications[*]

[1] J. Hoelzen, K. Sander, M. Sesia, D. Roy, E. Rijcken, A. Schnabel, B. Strücker, M. Juratli, A. Pascher. Robotic-assisted esophagectomy leads to significant reduction in postoperative acute pain: A retrospective clinical trial. *Ann. Surg. Oncol.* (2022) https://doi.org/10.1245/s10434-022-12200-0.

[2] N. Fingerhut, M. Sesia, Y. Romano. Coordinated double machine learning. *ICML* (2022). https://proceedings.mlr.press/v162/fingerhut22a.html

[3] A. Fayazi, M. Sesia, K. S. Anand. Hyperoxemia among pediatric intensive care unit patients receiving oxygen therapy. *J. Pediatr. Intensive Care* (2021). https://doi.org/10.1055/s-0041-1740586

[4] S. Li,[*] M. Sesia,[*] Y. Romano, E. Candès, C. Sabatti. Searching for robust associations with a multi-environment knockoff filter. *Biometrika* (2021). https://doi.org/10.1093/biomet/asab055

[5] M. Sesia, Y. Romano. Conformal regression with conditional histograms. *NeurIPS* (spotlight) (2021). https://arxiv.org/abs/2105.08747

[6] M. Sesia, S. Bates, E. Candès, J. Marchini, C. Sabatti. False discovery rate control in genome-wide association studies with population structure. *Proc. Natl. Acad. Sci. U.S.A.*, 118 (40) (2021). https://doi.org/10.1073/pnas.2105841118

---

[*]Asterisks indicate equal contributions.

[7] C. Chia,* M. Sesia,* C.-S. Ho, S. Jeffrey, J. Dionne, E. Candès, R. Howe. Interpretable classification of bacterial Raman spectra with knockoff wavelets. *IEEE J. Biomed. Health. Inform.* (2021). https://doi.org/10.1109/JBHI.2021.3094873

[8] Y. Romano,* M. Sesia,* E. Candès. Classification with valid and adaptive coverage. *NeurIPS* (spotlight) (2020). https://arxiv.org/abs/2006.02544

[9] S. Bates, M. Sesia, C. Sabatti, E. Candès. Causal inference in genetic trio studies. *Proc. Natl. Acad. Sci. U.S.A.*, 117 (39) 24117-24126 (2020). https://doi.org/10.1073/pnas.2007743117

[10] M. Sesia, E. Katsevich, S. Bates, E. Candès, C. Sabatti. Multi-resolution localization of causal variants across the genome. *Nature Commun.*, 11, 1093 (2020). https://doi.org/10.1038/s41467-020-14791-2

[11] M. Sesia, E. Candès. A comparison of some conformal quantile regression methods. *Stat*, 9:e261 (2020). http://dx.doi.org/10.1002/sta4.261

[12] Y. Romano,* M. Sesia,* E. Candès. Deep knockoffs. *J. Am. Stat. Assoc.* (2019). https://doi.org/10.1080/01621459.2019.1660174

[13] M. Sesia, C. Sabatti, E. Candès. Rejoinder: "Gene hunting with hidden Markov model knockoffs". *Biometrika*, 106, 35–45 (2019). https://doi.org/10.1093/biomet/asy075

[14] M. Sesia, C. Sabatti, E. Candès. Gene hunting with hidden Markov model knockoffs. *Biometrika*, 106, 1–18 (2019). https://doi.org/10.1093/biomet/asy033

Preprints

[1] M. Sesia, Tianshu Sun. Individualized conditional independence testing under model-X with heterogeneous samples and interactions. (2022) (*under review, J. Am. Stat. Assoc.*) https://arxiv.org/abs/2205.08653

[2] B. Einbinder,* Y. Romano,* M. Sesia,* Y. Zhou.* Training uncertainty-aware classifiers with conformalized deep learning. (2022) (*under review, NeurIPS*) https://arxiv.org/abs/2205.05878

[3] M. Sesia, S. Favaro. Conformalized frequency estimation from sketched data. (2022) (*under review, NeurIPS*) https://arxiv.org/abs/2204.04270

[4] S. Li,* Z. Ren,* C. Sabatti,* M. Sesia.* Transfer learning in genome-wide association studies with knockoffs. (2021). (*under review, Sankhya*) https://arxiv.org/abs/2108.08813

[5] S. Bates,* E. Candès,* L. Lei,* Y. Romano,* M. Sesia*. Testing for outliers with conformal p-values. (2021). (*under 2nd round review, Annals of Statistics*) https://arxiv.org/abs/2104.08279

## Patents

M. Sesia and Y. Abbasi-Yadkori (Adobe Inc). "Recommendation system using linear stochastic bandits and confidence interval generation". US 11,100,559. August 24, 2021.

## Awards

Jerome H. Friedman Applied Statistics Dissertation Award (2020).

International Master's Scholarship, Université Paris-Saclay (2014–2015).

Allievi Honors Program, Collegio Carlo Alberto (2011–2015).

## Professional Service

*Journal referee:*[†] Annals of Statistics (3), Bayesian Analysis (1), Biometrika (5), Briefings in Bioinformatics (1), Human Genetics (1), Journal of Machine Learning Research (1), Journal of the American

---

[†]Brackets contain the number of manuscripts reviewed.

Statistical Association (2), Journal of the Royal Statistical Society B (1), Nature Communications (1), Operations Research (1), SIAM Journal on Mathematics of Data Science (1), Statistics and Probability Letters (1), Statistics in Medicine (1), Statistical Science (2).

*Conference referee*: COLT (1), ISIT (1), NeurIPS (8), ICML (8).

*Ad-hoc grant reviewer*: Israel Science Foundation (1).

## DEPARTMENT SERVICE

Statistics seminar organizer (2021–2022); faculty hiring committee (2021–2022); PhD admission committee (2020–2021, 2021–2022)

## RESEARCH ADVISING

Current USC students:

PhD: Cora Liang (Math, 4th year), Yanfei Zhou (Data Sciences and Operations, 1st year), Fang Sun (Math, 1st year).

Past USC students:

Master: Xuanqi Zhang (Financial Engineering, Fall 2021).

## PRESENTATIONS

Invited presentations

*"Individualized conditional independence testing under model-X with heterogeneous samples and interactions"*. International Seminar on Selective Inference, June 2, 2022, virtual event.

*"Randomization inference with high-dimensional data, powered by machine learning"*. USC Dornsife, Department of Economics, Feb. 25, 2022, in Los Angeles, CA.

*"Conformal inference: how to estimate the predictive uncertainty of black-box machine learning models"*. Merck & Co., Inc. Feb. 10, 2022, in Morristown, NJ (virtual event).

*Controlling the false discovery rate in GWAS with population structure.* Yale University, School of Public Health. Feb. 16, 2021, in New Haven, CT (virtual event).

*Negative control variables for testing conditional genetic associations.* University of Nottingham, School of Mathematical Sciences. Dec. 17, 2020, in Nottingham, UK (virtual event).

*Negative control variables for testing conditional genetic associations.* University of Milano - Bicocca, Department of Economics, Management and Statistics. Dec. 2, 2020, in Milano, Italy (virtual event).

*Gene hunting with knockoffs.*
Johns Hopkins University, Mathematical Institute for Data Science. Feb. 18, 2020, in Baltimore, MD.

*Developments in knockoff generation and genetic mapping.*
USC Marshall School of Business, DSO Department. Jan. 27, 2020, in Los Angeles, CA.

*Developments in knockoff generation and genetic mapping.*
UC Davis, Department of Statistics. Jan. 6, 2020, in Davis, CA.

*Reproducible localization of causal variants across the genome.*
Regeneron Pharmaceuticals, Inc. Sept. 17, 2019, in Eastview, NY.

*Multi-resolution localization of causal variants across the genome.*
23andMe, Inc. May 21, 2019, in Mountain View, CA.

*Reproducible localization of causal variants across the genome.*
Stanford University, Statistics Department seminar. July 16, 2019, in Stanford, CA.

*Deep knockoffs.*
Stanford University, Statistics Industrial Affiliates Meeting. Feb. 22, 2019, in Stanford, CA.

*New tools for reproducible variable selection with knockoffs.*
Collegio Carlo Alberto, Statistics Seminar. Dec. 19, 2018, in Torino, Italy.

*Gene hunting with knockoffs for hidden Markov models.*
Université Grenoble Alpes, Bayes in Grenoble Seminar. July 10, 2018, in Grenoble, France.

Contributed presentations

*Conformal Prediction using Conditional Histograms.*
NeurIPS, spotlight presentation. Dec. 9, 2021 (virtual event).

*Testing for outliers with conformal p-values.*
RMDS Lab, IM DATA Conference 2021, Oct. 28, 2021, in Pasadena, CA (virtual event).

*Efficient uncertainty estimation for machine learning classifiers.*
RMDS Lab, IM DATA Conference 2020, Nov. 2, 2020, in Pasadena, CA (virtual event).

*Classification with valid and adaptive coverage.*
NeurIPS, spotlight presentation. Dec. 8, 2020 (virtual event).

*Gene hunting with hidden Markov model knockoffs.*
Royal Statistics Society Conference, Sept. 3–6, 2018, in Cardiff, United Kingdom.

*Gene hunting with knockoffs for hidden Markov models.*
Workshop on Model Selection, Regularization and Inference, July. 12–14, 2018, in Vienna, Austria.

*Gene hunting with knockoffs for hidden Markov models.*
Computational and Methodological Statistics Conference, Dec. 16–18, 2017, in London, United Kingdom.

Poster presentations

*Searching for consistent genetic associations with a multi-environment knockoff filter.*
American Society for Human Genetics Annual Meeting, Oct. 18–22, 2021, virtual meeting.

*Conformal histogram regression.*
ICML Workshop on distribution-free uncertainty quantification, July 24, 2021, virtual meeting.

*Genetic mapping with knockoffs for diverse and related samples.*
American Society for Human Genetics Annual Meeting, Oct. 27–30, 2020, virtual meeting.

*Multi-resolution localization of causal variants across the genome.*
American Society for Human Genetics Annual Meeting, Oct. 15–19, 2019, in Houston, TX.

*Multi-resolution localization of causal variants across the genome.*
Higher-Order Asymptotics and Post-Selection Inference Workshop, Aug. 17–19, 2019, in St. Louis, MO.

*Gene hunting with hidden Markov model knockoffs.*
American Society for Human Genetics Annual Meeting, Oct. 16–20, 2018, in San Diego, CA.