Professional Profile

Deniz Akdemir is a PhD statistician with advanced expertise in statistical methodology, experimental design, and complex data analysis. His highly-cited research (3,400+ citations) has established new statistical approaches for genomic prediction, mixed models, and high-dimensional data analysis. His career spans academic research and applied statistics across healthcare, agriculture, and clinical data analysis, where he leverages sophisticated quantitative methods to derive actionable insights from complex datasets. He is the creator of multiple statistical software packages including EMMREML and ml4t2e, which bridge traditional statistical theory with modern machine learning approaches.

Core Competencies

- Advanced Statistical Modeling
- Experimental Design & Data Analysis
- Predictive Analytics & Machine Learning
- Bayesian & Frequentist Methods
- Data Visualization & Communication

Professional Experience

- **Senior Clinical Data Scientist**, National Marrow Donor Program (2023 - Current)
  Lead complex statistical analyses of clinical transplant data, designing advanced statistical frameworks for stem cell transplant optimization, developing novel methodologies for time-to-event data analysis in clinical settings.
- **Statistical Methods Developer**, Independent Research (2010 - Present)
  Created the EMMREML package for fitting mixed models with known covariance structures, developed novel statistical approaches for high-dimensional data analysis, pioneered methodologies for training set optimization in predictive modeling.
- **Clinical Data Scientist**, National Marrow Donor Program (2021 - 2023)
  Applied rigorous statistical methodologies to patient outcome data, implemented Bayesian approaches to uncertainty quantification, developed predictive statistical models for clinical decision support.
- **Statistical Consultant**, Cornell Statistical Consulting Unit (2017 - 2019)
  Provided expert statistical guidance to academic researchers across disciplines, designed experimental studies, performed complex data analyses, and delivered statistical training workshops.
- **Statistical Methods Researcher**, University College Dublin & Cornell University (2011 - 2021)
  Conducted groundbreaking research in statistical methodologies for complex data structures, published extensively on novel statistical approaches, and developed computational tools for statistical analysis.

Education

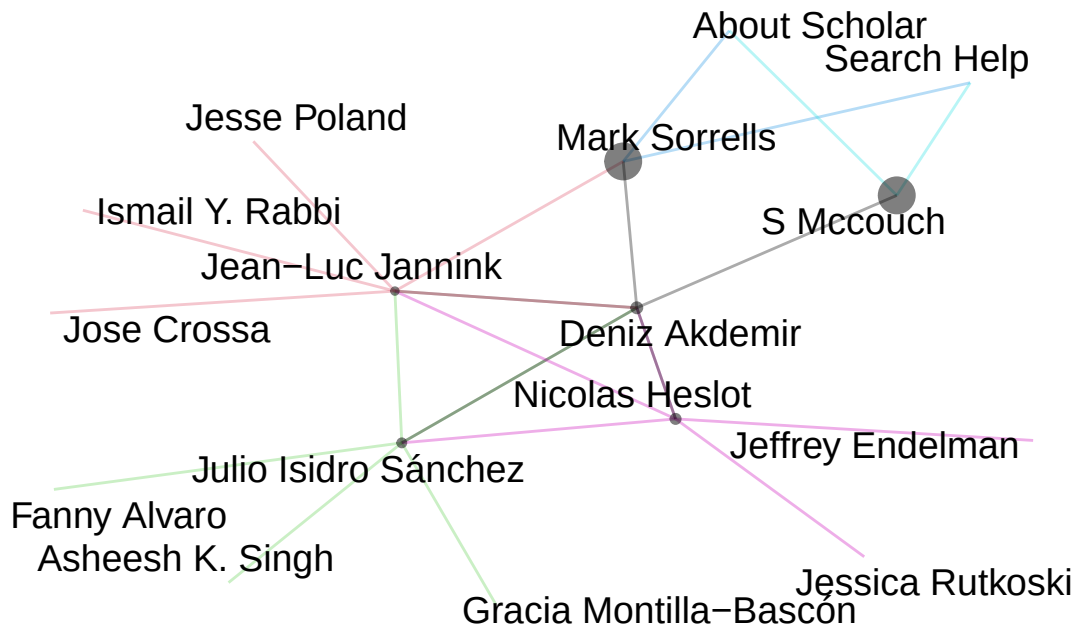
- **PhD. in Statistics**, Bowling Green State University (2009)
- **M.A. in Applied Statistics**, Bowling Green State University (2004)
- **M.S. in Statistics**, Middle East Technical University (2003)
- **B.A. in Business Administration**, Middle East Technical University (1999)

Research Interests

Deniz's research focuses on the development and application of advanced statistical methodologies for complex data structures. His work spans Bayesian methods, mixed models, high-dimensional data analysis, and the integration of statistical theory with machine learning applications. He has made significant contributions to

statistical methodologies for genomic prediction and selection, with applications in both plant breeding and clinical data analysis. His recent interests include developing statistical frameworks for machine learning applied to time-to-event data, Bayesian optimization techniques, and methods for sparse data in longitudinal settings.

Collaboration Network



Core Skills

- **Statistical Methodology:** Advanced multivariate analysis, Bayesian and frequentist methods, time series analysis, high-dimensional data modeling, mixed models, nonparametric statistics, survival analysis.
- **Data Analysis & Modeling:** Experimental design, hypothesis testing, regression analysis, statistical learning, model validation, predictive analytics, data visualization techniques.
- **Machine Learning Integration:** Statistical foundations for machine learning, dimensionality reduction, ensemble methods, cross-validation strategies, Bayesian optimization.
- **Statistical Computing:** Expert in R (package development, performance optimization), proficient in SAS, Python, and C++ for statistical programming and high-performance computing.
- **Software Development:** Creator of the EMMREML package for mixed models with known covariance structures, developer of ml4t2e for machine learning in time-to-event data.
- **Applied Statistics:** Extensive experience applying statistical methods to real-world problems in genomics, clinical trials, agriculture, and financial modeling.

Career Summary

- **Senior Clinical Data Scientist, National Marrow Donor Program, Minneapolis, USA (2023 - Current):** Engaged in statistical and machine learning analysis of stem cell transplant data, focusing on research into donor optimization. Applied for grants and submitted manuscripts to peer-reviewed journals. Wrote patent applications.

- **Clinical Data Scientist, National Marrow Donor Program, Minneapolis, USA (2021 - 2023):** Engaged in statistical and machine learning analysis of stem cell transplant data, focusing on research into donor optimization.
- **Postdoctoral Research Associate, School of Agriculture and Food Science, University College Dublin, Dublin, Ireland (2019 - 2021):** Conducted research in innovative statistical methodologies and contributed to high-impact publications.
- **Statistical Consultant, Cornell Statistical Consulting Unit, Cornell University, Ithaca, NY, USA (2017 - 2019):** Provided statistical consulting services for researchers at Cornell University, including the preparation and presentation of statistics workshops.
- **Postdoctoral Research Associate, Department of Plant Breeding and Genetics, Cornell University, Ithaca, NY, USA (2011 - 2017):** Focused on research developing new methodologies in genomic selection and prediction, mixed models, and machine learning, advising graduate students and preparing statistical software.
- **Visiting Assistant Professor, Department of Statistics and Actuarial Science, University of Central Florida, Orlando, FL, USA (2010 - 2011):** Responsibilities included teaching Data Mining Methodology, Theoretical Statistics, Applied Time Series Analysis, and Nonparametric Statistics.
- **Visiting Assistant Professor, Department of Mathematics and Statistics, Ohio Northern University, Ada, OH, USA (2009 - 2010):** Taught Statistics for Professionals, Statistics for Engineers, and Statistical Computing, catering to various undergraduate levels.

Key Statistical Projects

- **EMMREML Statistical Package:** Developed a specialized R package for fitting mixed models with known covariance structures, implementing efficient algorithms for large-scale genomic data analysis that significantly outperform standard implementations in computation time.
- **Statistical Frameworks for High-dimensional Data:** Created novel statistical methodologies for analyzing high-dimensional data with complex correlation structures, with particular applications in genomics and multi-omics integration.
- **Bayesian Optimization Methods:** Implemented advanced Bayesian approaches for experimental design optimization and statistical model parameter tuning, enhancing efficiency in resource-limited research settings.
- **Time-to-Event Data Analysis Innovations:** Developed new statistical methods bridging survival analysis with machine learning techniques, resulting in the ml4t2e package that enables more accurate prediction of time-to-event outcomes in clinical and biological applications.

Statistical Workshops and Training

- **Applied Statistical Methods:** Conducted comprehensive workshops at Cornell University on advanced statistical techniques for researchers across disciplines, focusing on practical implementation and interpretation.
- **Statistical Computing:** Delivered specialized training in R programming for statistical analysis, covering package development and efficient computation for large datasets.

Selected Publications (Statistics Focus)

1. **Akdemir, D.**, Isidro-Sánchez, J., & Jannink, J. L. (2015). Training set optimization under population structure in genomic selection. *Theoretical and Applied Genetics*, 128(1), 145-158. (391 citations)
2. **Akdemir, D.**, Isidro-Sánchez, J., & Jannink, J. L. (2016). Genome-wide prediction models that incorporate de novo GWAS results. *PLoS ONE*, 11(8), e0161054. (346 citations)

3. **Akdemir, D.**, Jannink, J. L., & Isidro-Sánchez, J. (2015). Optimization of genomic selection training populations with a genetic algorithm. *Genetics Selection Evolution*, 47(1), 38. (196 citations)
4. **Akdemir, D.** & Jannink, J. L. (2016). Efficient breeding by genomic mating. *Frontiers in Genetics*, 7(5), 210. (112 citations)
5. **Akdemir, D.** & Isidro-Sánchez, J. (2019). Multi-objective optimized genomic breeding strategies for sustainable food improvement. *Heredity*, 122(5), 672-683. (103 citations)

Statistical Methods & Software:

1. **Akdemir, D.** (2015). EMMREML: Fitting Mixed Models with Known Covariance Structures. *R package version 3.1.* https://CRAN.R-project.org/package=EMMREML (87 citations)
2. **Akdemir, D.** (2011). Array variate random variables with multiway kronecker delta covariance matrix structure. *Journal of Multivariate Analysis*, 102(5), 798-806. (52 citations)
3. **Akdemir, D.** & Gupta, A. K. (2011). Array variate random variables with multiway kronecker delta covariance matrix structure. *Journal of Algebraic Statistics*, 2(1), 98-113.
4. **Akdemir, D.** & Hames, K. (2016). Locally epistatic genomic relationship matrices for genomic association and prediction. *Journal of Quantitative Genetics and Genomics*, 1(2), 45-55.
5. **Akdemir, D.** & Beavis, W. D. (2021). TrainSel: an R package for selection of training populations. *BMC Bioinformatics*, 22(1), 1-10. (33 citations)

Education

- **PhD. in Statistics**, Bowling Green State University, Bowling Green, OH, USA, 2009.
- **M.A. in Applied Statistics**, Bowling Green State University, Bowling Green, OH, USA, 2004.
- **M.S. in Statistics**, Middle East Technical University, Ankara, Turkey, 2003.
- **B.A. in Business Administration**, Middle East Technical University, Ankara, Turkey, 1999.

Professional References

- **Dr. Yung-Tsi Bolon**
  - **Affiliation:** Director, Immunobiology & Bioinformatics Research, NMDP, Minneapolis, Minnesota, United States
  - **Relationship:** Supervisor at the National Marrow Donor Program
  - **Contact:** ybolon@nmdp.org
- **Dr. Julio Isidro-Sanchez**
  - **Affiliation:** Associate Professor: Centro de Biotecnologia y Genomica de Plantas, Universidad Politecnica de Madrid, Instituto Nacional de Investigacion y Tecnologia Agraria y Alimentaria, Campus de Montegancedo - UPM, 28223-Pozuelo de Alarcon, Madrid, Spain
  - **Relationship:** Expert in plant breeding and genetics, collaborator on various projects
  - **Contact:** j.isidro@upm.es
- **Dr. Jhonathan Pedroso**
  - **Affiliation:** Research Scientist at Corteva Agriscience, Corteva, Johnston, Iowa, USA
  - **Relationship:** Industry partner in genomic tool development, contributed to software enhancements
  - **Contact:** jhowpd@gmail.com
- **Dr. Lynn Johnson**
  - **Affiliation:** Interim Director and Statistical Consultant, Cornell Statistical Consulting Unit, Cornell University, Ithaca, NY, USA
  - **Relationship:** Coworker at the Cornell Statistical Consulting Unit

  - **Contact:** lms86@cornell.edu
- **Dr. Roberto Fritsche Neto**
  - **Affiliation:** Assistant Professor, Department of Plant, Environmental Management & Soil Sciences, LSU
  - **Relationship:** Collaborator on various projects
  - **Contact:** rneto1@lsu.edu