

# STATISTICS



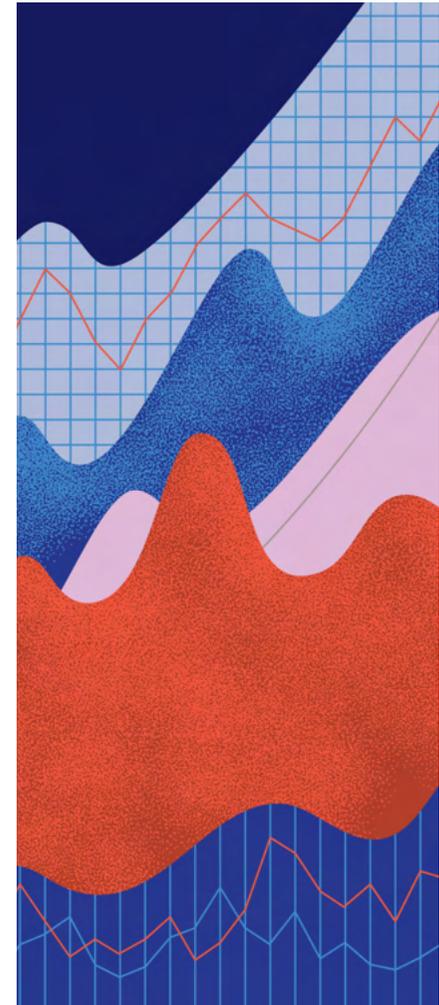


---

The Department of Statistics seriously adopts strategies such as increasing the use of statistics, ensuring that it takes a strong place in all decision-making processes, enabling graduates to participate effectively in data-based analysis, helping students develop a scientific, critical, creative, and innovative perspective, making sure and encouraging society to use statistical science in the most accurate, reliable, and ethical manner as their missions.

## Content

- |   |                        |    |                          |
|---|------------------------|----|--------------------------|
| 3 | Message from the Chair | 8  | Undergraduate Curriculum |
| 4 | About Us               | 13 | Graduate Curriculum      |
| 6 | Facts and Figures      | 18 | Faculty                  |
| 7 | Minor Program          |    |                          |







## Message from the Chair

---



Welcome to the Department of Statistics at Middle East Technical University.

The department, founded in 1975, has been contributing ever since to the research and training undergraduate and graduate students in the field of theoretical, computational and applied statistics.

At present, we have 11 full-time faculty members, many of whom have received their doctorate degree from universities in US or UK. The department has major and minor undergraduate programs, Master of Science (M.S.) program, and a Doctor of Philosophy (PhD) program.

Most of our students are employed even before they graduate, and they pursue their careers in well-known universities, companies, government agencies worldwide. Statisticians and data scientists are ranked among the top jobs all over the world.

We are working in the exciting fields of data science, computing, making decisions under uncertainty, and developing new methods for solving real life problems. Statistics is the science and art of developing and studying methods for collecting, analyzing, interpreting and presenting data. In today's age of data-driven world, we, statisticians, have become indispensable part of all sorts of research, and we work with scientists from many other fields, including genetics, medicine, policy, economics, and engineering.

I invite you to the fascinating world of statistics. Please review our website, and contact me if you have any questions that are not addressed on this website.

With my best wishes of success!  
Sincerely,  
Prof. Dr. Özlem İlk Dağ

# About Us

---

The education in statistics at Middle East Technical University dates back to the 1960s. It was first established under the Faculty of Economics and Administrative Sciences as the Department of Economics and Statistics. In 1975 it was separated into two departments, namely Economics and Applied Statistics. In the same year, the Department of Operations Research and Statistics was established and became a part of the Faculty of Arts and Science. In 1982, the Department of Operations Research and Statistics was combined with the Department of Applied Statistics and renamed as the Department of Statistics, contributing to the research and training of students.

The Department of Statistics, at the moment, has 11 full-time faculty members. They publish their research in top-tier academic journals, and several have national and international research projects. Starting in 1982, many distinguished academicians joined our department as visiting faculty and contributed to developing our department in different scientific areas.

Our department accepts approximately 40 undergraduate and 15 graduate students each year. Students who graduate with a Bachelor of Science in Statistics are prepared to be decision-makers in almost every facet of business and various levels of government, climate, healthcare, and other industries where large and complicated data sets need to be analyzed, understood, and explained. Our students are also trained to collaborate effectively with people in different fields. This program





engages students in data analysis, statistical methodology, and computer languages such as Python and R.

Besides a solid foundation in the theory of statistics at the graduate level, substantial knowledge in statistical methodology and statistical computing is also gained. Our department also has an interdisciplinary M.Sc. program for students who have graduated from fields other than statistics. Doctoral students build a body of research that should substantially contribute to the science of statistics. The scientific areas in which high-level research is conducted include machine learning, big data, data science, time series analysis, panel data analysis, robust regression, computational statistics, biostatistics, and system reliability.

The curriculum is updated frequently to stay updated with the world's developments. The most recent changes in the undergraduate and graduate curriculums have been processed in 2022 in the direction of data science.

Almost all of our graduates are employed just after their graduation and work in academia (METU, Hacettepe University, University of Michigan, etc.), in government (TCMB, TUIK, TUBITAK, Aselsan, etc.), or in the private sector (all banks, Roche, IBM, etc.).

# Facts and Figures



Founded in 1975

## Duration of Study

Bachelor 4 years

Master of Science 2 years

PhD 4 years



## FULL-TIME FACULTY

9 Professors

2 Associate Professors

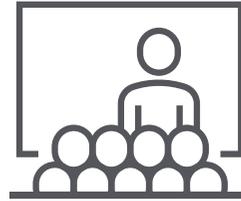
## STUDENTS

423 Undergraduate

46 Graduate

47 International

2 Minor



## COURSE CREDITS

143 Bachelor of Science

21 Master of Science

29 PhD



## PROJECTS & PUBLICATIONS\*

Projects:

5 TUBITAK

5 International

7 National

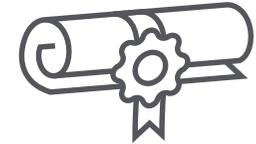
7 BAP

Publications:

52 International

2 National

*\*Last 5 years*



## DEGREES

1875 B.Sc.

140 M.Sc.

34 Ph.D.

İSTATİSTİK BÖLÜMÜ  
DEPARTMENT OF STATISTICS  
1975

# Minor Program

---

This program is designed for students of non-statistical background.

## **Preliminary courses required for admission:**

---

MATH119 Calculus I or equivalent

---

MATH120 Calculus II or equivalent

---

MATH260 Basic Linear Algebra or equivalent

---

STAT156 Statistical Methods or equivalent

---

## **Compulsory Courses**

---

STAT 203 Probability I

---

STAT 204 Probability II

---

STAT 303 Mathematical Statistics I

---

STAT 304 Mathematical Statistics II

---

STAT 363 Linear Models I

---

Plus one elective course offered by the Department of Statistics.

# Course Descriptions

## **STAT156 Statistical Methods (3-2)4**

Sampling distributions, estimation, confidence intervals, hypothesis testing distribution fitting, analysis of variance for one factor design, linear regression, association between two categorical variables, basic nonparametric procedures.

## **STAT203 Probability I (3-2)4**

Sample space, events, basic combinatorial probability, conditional probability, Bayes theorem, independence, random variables, distributions, expectation.

## **STAT204 Probability II (3-2)4**

Transformations of random variables, generating functions, conditional expectation, limit theorems, central limit theorem, limiting distributions.

## **STAT303 Mathematical Statistics I (3-2)4**

Common theoretical distributions. Sampling distributions. Principles of point estimation. Techniques of estimation. Properties of point estimators. Optimality criteria in estimation. Selected topics from robust inference. Bayesian inference.

## **STAT304 Mathematical Statistics II (3-2)4**

Region (interval) estimation. Hypothesis testing. Optimality properties for hypothesis testing. Likelihood ratio tests. Sequential tests.

## **STAT363 Linear Models I (3-2)4**

Simple and Multiple Linear Regression Models. Estimation, interval estimation and test of hypothesis on the parameters of the models. Model adequacy checking. Multicollinearity. Transformation. Prerequisites: MATH 260, STAT 156

# Undergraduate Curriculum

## First Semester

STAT101	Intro to Statistics and Data Science I	(3-2)4
MATH119	Calculus with Analytic Geometry	(4-2)5
ENG101	English for Academic Purposes I	(4-0)4
CEIT101	Introduction to Computer Programming	(2-2)3
STAT112	Int. to Data Processing and Visualization	(3-2)4

## Second Semester

STAT102	Intro to Statistics and Data Science II	(3-2)4
MATH120	Calculus for Functions of Several Variables	(4-2)5
ENG102	English for Academic Purposes II	(4-0)4
CENG240	Programming with Python for Engineers	(2-2)3
BA100	Career Planning	NC
IS100	Introduction to Information Technologies and Application	NC

# Course Descriptions

## **STAT101 Introduction to Statistics and Data Science I (3-2)4**

Understand the fundamentals of statistics and data science, understanding the difference between a population and a sample in a given dataset. Basic statistical definitions and learn how to work with different sorts of data. How to visualize different data types. Descriptive statistics: measures of central tendency, asymmetry, variability, correlation and covariance. What are the random variables in a dataset? Distinguish and work with different distributions that describe distinct data sources.

## **STAT102 Introduction to Statistics and Data Science II (3-2)4**

Basic statistical analyses in different types of data. Sampling distributions of distinct data sources. Inferential statistics in the data science. Estimation, confidence intervals and hypothesis testing under various data types. Distribution fitting and analysis of variance for one factor design in a given dataset. Linear regression and association between two categorical variables. Basic nonparametric procedures under various data types.

Prerequisite: STAT 101

## **STAT112 Introduction to Data Processing and Visualization (3-2)4**

Basic definitions and managing different types of data. Introduction to manipulation (indexing, subsetting, reshaping, transforming etc.), visualization, mapping and analysis of data. Dealing with common problems like missing or inconsistent values in datasets. Use of related R and/or Python programming packages. Merging multiple data tables (equivalent to an SQL JOIN).

## **STAT203 Probability I (3-2)4**

Sample space, events, basic combinatorial probability, conditional probability, Bayes' theorem, independence, random variables, distributions, expectation

# Undergraduate Curriculum

## Third Semester

STAT203	Probability I	(3-2)4
STAT291	Statistical Computing	(3-2)4
MATH219	Intro. to Differential Eqns.	(4-0)4
MATH260	Basic Linear Algebra	(3-0)3
HIST2201	Principles of Kemal Atatürk I	NC
NDE	Non Department Elective	(3-0)3

## Fourth Semester

STAT204	Probability II	(3-2)4
STAT250	Applied Statistics	(4-2)5
STAT295	Object Oriented Programming	(3-2)4
MATH250	Advanced Calculus in Statistics	(4-2)5
ENG211	Academic Oral Presentation Skills	(3-0)3
HIST2202	Principles of Kemal Atatürk II	NC

## Fifth Semester

STAT303	Mathematical Statistics I	(3-2)4
STAT311	Modern Database Systems	(3-2)4
STAT363	Linear Models I	(3-2)4
STAT365	Survey Sampling Techniques	(4-2)5
TURK303	Turkish I	NC

# Course Descriptions

## STAT204 Probability II (3-2)4

Transformations of random variables, generating functions, conditional expectation, limit theorems, central limit theorem, limiting distributions. Prerequisite: STAT 203, MATH 119

## STAT250 Applied Statistics (4-2)5

Sampling distributions. Sample drawing techniques. Estimation and testing for one or two population characteristics. Maximum likelihood estimation of parameters. Measures of association. Simple and multiple regression. Introduction to design of experiments, analysis of variance; one-way, multiway classifications. Multiple comparisons. Basic nonparametric procedures. Elementary time series analysis; trends, seasonality, forecasting. Indexing. Some applications in medicine, science, engineering and social sciences. Prerequisite: STAT 156

## STAT291 Statistical Computing (3-2)4

Introduction to statistical techniques in statistical software. Managing and analyzing data using statistical database packages. Introduction to MATLAB with applications to matrix algebra. Prerequisites: CENG 240, STAT 156

## STAT295 Object Oriented Programming (3-2)4

Introduction to Object-Oriented Programming (OOP) with a language (e.g C++, Python). Programming elements. Functions. Classes and objects. Constructors and destructures. Operator overloading. Type conversion. Inheritance. Pointers. Polymorphism. Templates. Exception handling. String manipulation. File handling. Advanced Input/Output. OO system analysis, design and development. Prerequisite: CENG 230 or CENG 240

## STAT303 Mathematical Statistics I (3-2)4

Common theoretical distributions. Sampling distributions. Principles of point estimation. Techniques of estimation. Properties of point estimators. Optimality criteria in estimation. Selected topics from robust inference. Bayesian inference. Prerequisite: STAT 204 or CD, MATH 120

# Undergraduate Curriculum

## Sixth Semester

STAT304	Mathematical Statistics II	(3-2)4
STAT333	Data Structures and Algorithms	(3-2)4
STAT361	Computational Statistics	(3-2)4
STAT364	Linear Models II	(3-2)4
STAT376	Stochastic Processes	(4-2)5
TURK304	Turkish II	NC

## Seventh Semester

STAT411	Statistical Data Mining	(3-2)4
STAT457	Statistical Design of Experiments	(3-2)4
STAT467	Multivariate Analysis	(4-2)5
DE	Departmental Elective	(3-0)3
DE	Departmental Elective	(3-0)3

## Eighth Semester

STAT401	Int. to Big Data	(3-2)4
STAT433	Statistical Machine Learning	(3-2)4
DE	Departmental Elective	(3-0)3

# Course Descriptions

## STAT304 Mathematical Statistics II (3-2)4

Region (interval) estimation. Hypothesis testing. Optimality properties for hypothesis testing. Likelihood ratio tests. Sequential tests.

Prerequisite: STAT 303

## STAT311 Modern Database Systems (3-2)4

Introduction to database systems. Relational databases. Entity relationship (ER) model. Normalization. Structured Query Language (SQL). Designing databases. Introduction to distributed, parallel and object databases. Big data storage systems. Datawarehouses. Online Analytic Processing (OLAP). Big data analytics and NoSQL. Web data management. Cloud computing.

Prerequisite: STAT 291

## STAT333 Data Structures and Algorithms (3-2)4

Introduction data structures and algorithms with an object-oriented programming (e.g. C++, Python). Principles of algorithm design. Recursion. Backtracking. Complexity analysis. Big O notation. Arrays, lists, pointers. Stacks, queues, deques. Trees. Hash and symbol tables. Graphs. Searching, sorting, selection, string algorithms. Pattern-matching. Algorithm design. Greedy, divide and conquer algorithms. Dynamic programming. Classification of algorithms.

Prerequisite: STAT 295

## STAT361 Computational Statistics (3-2)4

Random number generation. Generating from other distributions. Monte Carlo methods for inferential statistics. Resampling. Data partitioning. Cross-validation. Bootstrapping. Jackknifing. Tools for exploratory and graphical data analysis. Nonparametric probability density estimation.

Prerequisite: STAT 291

## STAT363 Linear Models I (3-2)4

Simple and Multiple Linear Regression Models. Estimation, interval estimation and test of hypothesis on the parameters of the models. Model adequacy checking. Multicollinearity. Transformation.

Prerequisites: MATH 260, STAT 102 or STAT 156

## STAT364 Linear Models II (3-2)4

Simple nonlinear models, Less than full rank models : One-way , Two-way ANOVA models, Multiple comparison tests, Analysis of Covariance (ANCOVA) Models, Introduction to generalized linear models (GLM), Poisson regression, Logistic regression.

Prerequisite: STAT 363

## STAT365 Survey Sampling Techniques (4-2)5

Introduction to survey sampling. Probability sampling techniques. Simple random sampling. Stratified element sampling. Systematic sampling. Equal sized cluster sampling. Unequal sized cluster sampling. PPS selection techniques. Sampling errors. Survey research methods. Planning of sample surveys. Questionnaire design techniques. Survey research project.

Prerequisite: STAT 156 or STAT 102

# Course Descriptions

## **STAT376 Stochastic Processes (4-2)5**

Review of Probability. Theory Markov Chains. Discrete and Continuous time Markov Chains. Poisson Processes. Queuing Processes. Birth and Death Processes. Decision Analysis.

Prerequisite: MATH 260, STAT 204 or STAT 154

## **STAT401 Introduction to Big Data (3-2)4**

In general, turning massive data sets into information and information into value. The definition of big data and continues with programming tools to handle massive data sets. Parallelization used for big data. Ended with an environment which provides scalability to store large volumes of data on commodity hardware. Statistical tools used for big data. Visualization used for big data.

Prerequisite: STAT 295 and STAT 311

## **STAT411 Statistical Data Mining (3-2)4**

Descriptive and predictive mining. Data preprocessing: cleaning transformation. outlier detection, missing data imputation. Dimension reduction, Principal Component Analysis (PCA). Sampling, oversampling. Exploratory data analysis (EDA). Clustering methods: partitioning, hierarchical, density-based, model-based. Predictive modeling. Regression. Variable selection. Robust and nonlinear regression. Nonparametric regression. Classifiers. Logistic regression. Decision trees. Random Forest. Model evaluation and validation. Real-life applications using recent available software.

Prerequisite: STAT 363 and STAT 291

## **STAT 412 Statistical Data Analysis (3-2)4**

Types of data. Graphical and tabular representation of data. Approaches for finding unexpected in data. Exploratory data analyses for large and high-dimensional data. Analysis of categorical data. Elements of robust estimation. Handling missing data. Smoothing methods. Data mining.

Prerequisite: STAT 291 or STAT 292, STAT 363, or CD

## **STAT433 Statistical Machine Learning (3-2)4**

Regression and prediction, optimization, regularization (Ridge regression and LASSO), neural networks and deep learning, classification, kernel methods and support vector machines, decision trees, bagging and random forest, boosting algorithms, principal component regression, unsupervised deep learning. Applying the methods to real data.

Prerequisite: STAT 411

## **STAT440 Artificial Intelligence and Statistics (3-0) 3**

Foundations and history of Artificial Intelligence (AI). Logical programming. Problem solving. Searching. Game playing. Knowledge, reasoning, planning. Theorem proving. Uncertain knowledge and probabilistic reasoning. Hidden Markov Models. Kalman filters. Bayesian networks. Belief networks. Machine learning. Statistical learning. Reinforcement learning. Natural Language Processing (NLP). Pattern

recognition. Speech and image processing. Future of AI. Robotics.

Prerequisite: STAT 333

## **STAT 444 Advanced Statistical Computing (3-0)3**

Reading raw data files and Statistical Analysis Software (SAS) data sets, and writing the results to SAS data sets; subsetting data; combining multiple SAS files; creating SAS variables and recoding data values; creating listing and summary reports.

Prerequisite: STAT 156 or STAT 102

## **STAT 455 Statistical Business Analytics(3-2) 4**

General introduction to data structures; Statistical data collection and types of business data; Common business problems: customer analytics, segmentation, sales, demand, pricing, fraud, advertisement targeting; Introduction to marketing analytics, definition of marketing terms and definitions, statistical thinking for business problems; Exploratory data analysis and descriptive techniques for business data; Methods for acquiring and manipulating data; Applications of statistical methods for real business cases, machine learning algorithms; Data-centric decision support systems; Statistical applications & discussions.

## **STAT457 Statistical Design of Experiments (3-2)4**

Strategies for experimentation, randomized complete and balanced incomplete block designs, Latin squares. General, two-level and fractional factorials. Blocking and confounding in two-level factorials. Three and mixed level factorial and fractional factorials. Introduction to response surface methodology. Second-order experimental designs. Nonnormal responses. Unbalanced data in factorials. Split-plot designs, Nested designs, Random effect models. Repeated measures.

Prerequisite: STAT 363 or CD.

## **STAT460 Nonparametric Statistics (3-0)3**

Review of basic statistics. Distribution-free statistics, ranking statistics, U statistics. Large sample theory for U statistics. Tests based on runs. Asymptotic relative efficiency of tests. Hypothesis testing, point and interval estimation. Goodness of fit, rank-order (for location and scale), contingency table analysis and relevant models. Measures of association, analysis of variance.

Prerequisite: CD.

## **STAT461 System Simulation (3-2)4**

Introduction to discrete-event system simulation and simulation software. Statistical models in simulation. Queuing models. Input data modeling. Variance reduction techniques. Verification and validation of simulation models. Output analysis for a single model. Comparison and evaluation of alternative system design.

Prerequisite: STAT 156 and STAT 292

## **STAT462 Biostatistics (3-2)4**

Populations and samples. Types of biological data. Data transformations. Survival data analysis. Life tables. Sample size determination in clinical trials. Measures of association. The odds ratio and some properties. Application of generalized linear models and logistic regression to biological data. Analysis of data from matched samples.

Prerequisite: STAT 156

## **STAT463 Reliability (3-0)3**

Reliability studies. Statistical failure models. Censoring and truncation and their types. Useful limit theorems in reliability. Inference procedures for lifetime distributions. System reliability. Bayesian methods. Accelerated life testing.

Prerequisite: STAT 304

## **STAT464 Operations Research (2-2)3**

Basic operations research methodology. Basic models such as network flow models, project scheduling, dynamic programming, and production and inventory control. LP and game theory. Two person zero-sum games and mixed strategies.

Prerequisite: MATH 260

## **STAT467 Multivariate Analysis (4-2)5**

Sample mean vector and sample covariance matrix; matrix decomposition; multivariate normal and Wishart distributions; parameter estimation; hypothesis testing; MANOVA; principal components; factor analysis; multivariate classification and clustering; canonical correlation.

Prerequisites: MATH 260, STAT 156

## **STAT472 Statistical Decision Analysis (3-2)4**

Introduction to decision making and types of decision situations. Bayes theorem and Bayesian decision theory. Prior, posterior and conjugate prior distributions. Loss functions. Empirical Bayesian approach. Utility theory for decision making. Value of information. Sequential decision procedures. Multidecision problems.

Prerequisite: STAT 204

## **STAT477 Statistical Quality Control (2-2)3**

Introduction to concepts of quality and total quality management. Basic principles of teamwork and learning. Probability in Quality Control. Methods and Philosophy of Statistical Process. Control Charts for variables and attributes. Cumulative-Sum and Exponentially Weighted Moving-Average Control Charts. Process Capability Analysis. Introduction to Experimental Design and Factorial Experiments. Taguchi Method, Lot-by-Lot Acceptance Sampling for attributes and by variables.

Prerequisite: STAT 156

# Course Descriptions

## **STAT479 Linear Programming (2-2)3**

Introduction to Linear Programming (LP). The simplex method. Transportation, assignment and transshipment problems. Sensitivity testing, duality theory and its applications. Advanced methods in LP and revised simplex algorithm.

Prerequisite: MATH 260

## **STAT480 Application of Statistical Techniques in Socio-Economic Research (3-2)4**

Principals of empirical socio-economic research. Formulation of research problems, determination of research design, application of sampling design. Strategies of field work, collection of data, improving data quality, selecting appropriate statistical methods. Evaluation of test of hypothesis and interpretation of findings. Preparation and presentation of a research proposal and report. Prerequisite: STAT 412

## **STAT482 Categorical Data Analysis (3-2)4**

Probability distributions and measures of association for count data. Inferences for two-way contingency tables. Generalized linear models, logistic regression and loglinear models. Models with fixed and random effects for categorical data. Model selection and diagnostics when response is categorical. Classification trees.

Prerequisite: STAT 304

## **STAT486 APPLIED STATISTICS and ECONOMETRICS (3-2)4**

The course contains theory of regression analysis for econometrics, and explains applications of regression analysis to a variety of econometric problems. Hypothesis testing and confidence intervals in multiple regression and Diagnostic Checking. Stationary and nonstationary time series. Cointegration and error correction models. Panel data regression models (Fixed and random effects model). Dynamic econometric models: Autoregressive and distributed-lag models. Stochastic regressors and the method of instrumental variables

## **STAT487 Insurance and Actuarial Analysis (3-0)3**

Basic definition of insurance. Historical background. Insurance applications in government and private sector, regulations and legislation in insurance. Fundamentals of insurance. Types of insurance, disaster insurance and risk management applications around the world. Turkish catastrophe insurance pool. Definition of risk, probability aspect of risk. Utility theory, claim processes, distribution of claim processes.

Prerequisite: CD.

## **STAT493 New Horizons in Statistics (3-0)3**

New advances in the field of statistics.

Prerequisite: CD.

## **STAT495 Applications in Statistics (2-2)3**

Applications of different statistical methods in various disciplines such as medicine, science, engineering and social sciences. Presentation of projects involving these applications as group studies.

Prerequisite: STAT 156

## **STAT497 Applied Time Series Analysis (3-2)43**

Time series as a stochastic process. Means, covariances, correlations, stationarity. Moving averages and smoothing. Stationary and nonstationary parametric models. Model specification. Estimation and testing. Seasonality. Some forecasting procedures. Elementary spectral domain analysis. Exponential smoothing methods. Unit root tests.

Prerequisite: CD.

## **STAT499 Undergraduate Research (1-4)3**

This course is intended to improve the research capabilities of graduating students. Each student will be given a project and an academic advisor; lectures will be given on research design, data evaluation and report writing. A final report and/or seminar is required at the end of the semester.

Prerequisite: CD.



# Graduate Curriculum

M.S. degree in Statistics is offered as Statistics with and without thesis (not active). At the Doctorate level, there is only one program, Ph.D. in Statistics, applied in two tracks: admission by M.S. degree and admission by B.S. degree.

## Master of Science Curriculum in Statistics Scientific Preparation

STAT291	Statistical Programming	(3-2)4
STAT363	Linear Models I	(3-2)4
STAT295	Object Oriented Programming	(3-2)4
STAT433	Statistical Machine Learning	(3-2)4
or		
STAT412	Statistical Data Analysis	(3-2)4

## M.S. Program With Thesis

STAT500	M.S. Thesis	NC
STAT510	Research Methods and Ethics in Statistics and Data Science	NC
STAT542	Seminar I	NC
STAT543	Seminar II	NC
STAT571	Data Mining and Machine Learning	(2-2)3
STAT572	Probability and Statistics for Data Science I	(3-0)3
STAT573	Probability and Statistics for Data Science II	(3-0)3
STAT8XX	Special Studies	NC

## Elective Courses

Four elective courses. Two of them could be from out of department.  
Totally seven courses with at least 21 credit hours

# Course Descriptions

## STAT500 M.S. Thesis NC

Program of research leading to M.S. degree arranged between student and faculty member. Students register to this course in all semesters starting from the beginning of their second semester while the research program or write-up of thesis is in progress.

## STAT510 Research Methods and Ethics in Statistics NC

Research design in the field of statistics following ethical standards, ethical issues in scientific research, how to write a thesis with ethics, journal types, publication types, citations, plagiarism, how to be a graduate student, how to be a researcher.

## STAT542 Seminar I NC

Seminar course for M.S. students in Statistics.

## STAT543 Seminar II NC

M.S. students prepare and present a seminar in their thesis topic.

## STAT571 Data Mining and Machine Learning (2-2)3

Unsupervised learning. Principal component analysis (PCA), clustering methods. Rule learning, association rules. Supervised learning. Multiple linear regression. K-nearest neighbors. Logistic regression. Linear discriminant analysis. Linear model selection. Regularization techniques. Ridge regression, LASSO. Splines. Generalized additive models (GAMs). Tree-based methods. Ensemble learning. Bagging, random forest, boosting. Support vector machines. Neural networks and deep learning. Evaluating the performance of machine learning algorithms. No Free Lunch theorems. Bias-variance decomposition. Bagging. Boosting. Generative adversarial networks (GANs). Autoencoders. Variational Autoencoders.

## STAT572 Probability and Statistics for Data Science I (3-0)3

Introduction to history and concepts of probability, statistics and data science, artificial intelligence, data mining, machine learning, deep learning, comparison of these topics, probability, combinatorics, random variables, some common discrete and continuous distributions and their properties, expectations, joint distribution functions, conditional distributions, distribution functions, moment generating functions, transformations of variables, multivariate normal distribution, limit theorems.

## STAT573 Probability and Statistics for Data Science II (3-0)3

Order statistics. Likelihood functions and likelihood theory. Exponential families. Point estimation. Some properties of estimators. Interval estimation. Hypothesis testing.

## STAT598 Term Project in Statistics NC

A project is carried out under the supervision of a faculty member in a specified area of Statistics. Students are required to write a report about their work.

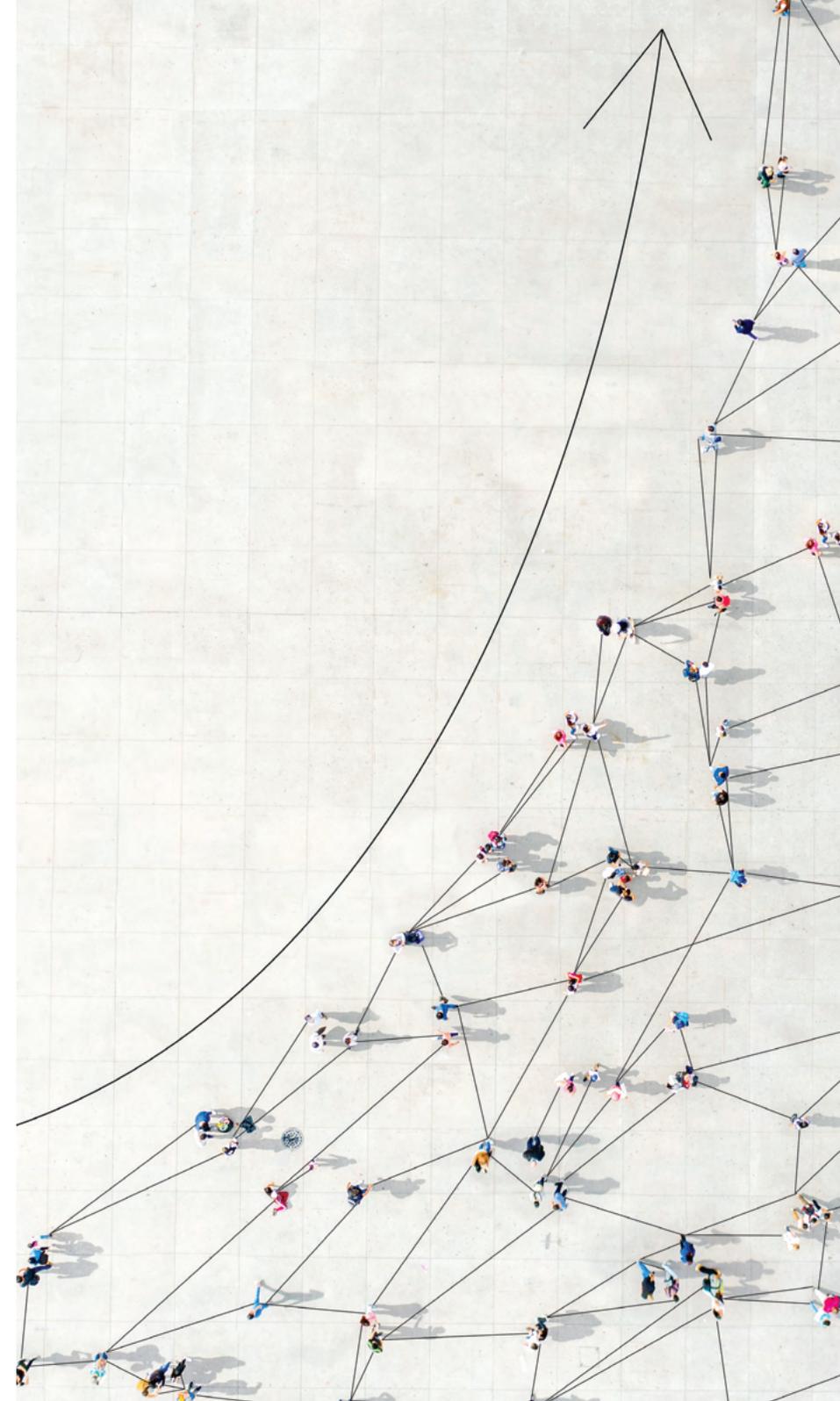
# Graduate Curriculum

## M.S. Program in Statistics without Thesis (not active)

STAT500	M.S. Thesis	NC
STAT510	Research Methods and Ethics in Statistics and Data Science	NC
STAT542	Seminar I	NC
STAT543	Seminar II	NC
STAT571	Data Mining and Machine Learning	(2-2)3
STAT572	Probability and Statistics for Data Science I	(3-0)3
STAT573	Probability and Statistics for Data Science II	(3-0)3
STAT598	Term Projects in Statistics	NC
STAT8XX	Special Studies	NC

## Elective Courses

Seven elective courses. Two of them could be from out of department.  
Totally ten courses with at least 30 credit hours



# Graduate Curriculum

## Doctor Of Philosophy Curriculum in Statistics Scientific Preparation

STAT570	Data Handling and Visualization	(3-0)3
STAT571	Data Mining and Machine Learning	(2-2)3
STAT572	Probability and Statistics for Data Science I	(3-0)3
STAT573	Probability and Statistics for Data Science II	(3-0)3

Total: 16 credits

## Ph.D. Program

STAT510	Research Methods and Ethics in Statistics and Data Science*	NC
STAT600	PhD Thesis	NC
STAT635	Advanced Computational Statistics	(3-0)3
STAT636	Advanced Generalized Linear Models	(3-0)3
STAT642	Seminar in Statistics and Data Science I	NC
STAT643	Seminar in Statistics and Data Science II	NC
STAT647	Probability Theory	(4-0)4
STAT648	Advanced Statistical Inference	(4-0)4
STAT8XX	Special Studies	NC

Five elective course(s) approved by the Department of Statistics.

Totally nine courses with at least 29 credit hours.

\* If not taken during M.S.

# Course Descriptions

## STAT600 Ph.D. Thesis NC

Program of research leading to Ph.D. degree arranged between student and a faculty member. Students register to this course in all semesters starting from the beginning of their second semester while the research program or write-up of thesis is in progress.

## STAT 635 Advanced Computational Statistics (3-0)3

Exploring multidimensional data. Discovering structure in data. Bootstrapping basics and dependent data. Data partitioning. Statistical pattern recognition: classifiers and clustering. Bivariate and multivariate smoothing. Nonparametric regression models. Advanced topics in Markov Chain Monte Carlo (MCMC).

## STAT 636 Advanced Generalized Linear Models (3-0)3

Review of matrix algebra. A theoretical development of generalized linear models. Estimation, interpretation and inferences for generalized linear models for responses from different distributions, such as Gaussian, Binomial, Poisson. Loglinear models. Penalized estimation.

## STAT642 Seminar in Statistics I NC

Seminar course for Ph.D. students in Statistics.

## STAT643 Seminar in Statistics II NC

Ph.D. students prepare and present a seminar in their thesis topic.

## STAT 647 Probability Theory (4-0)4

This is an introductory measure-theoretic probability theory course. We will be covering the fundamental concepts in Statistics to the full extent, such as Law of Large Numbers, Central Limit Theorem, Convergence, Dependence, Independence, Conditional Expectation. The course covers all the content that appear on the Probability Qualification Examination.

## STAT 648 Advanced Statistical Inference (4-0)4

Theory of likelihood based estimation, likelihood construction for advance models, theory of robust estimation, theory of hypothesis testing

# Graduate Curriculum

## Integrated Ph.D. Program

### Required Courses

STAT510	Research Methods and Ethics in Statistics and Data Science*	NC
STAT542	Seminar I	NC
STAT543	Seminar II	NC
STAT571	Data Mining and Machine Learning	(2-2)3
STAT572	Probability and Statistics For Data Science I	(3-0)3
STAT573	Probability and Statistics For Data Science II	(3-0)3
STAT600	PhD Thesis	NC
STAT635	Advanced Computational Statistics	(3-0)3
STAT636	Advanced Generalized Linear Models	(3-0)3
STAT642	Seminar in Statistics and Data Science I	NC
STAT643	Seminar in Statistics and Data Science II	NC
STAT647	Probability Theory	(4-0)4
STAT648	Advanced Statistical Inference	(4-0)4
STAT8XX	Special Studies	NC

8 elective course(s) approved by the Department of Statistics.

Totally fifteen courses with at least 47 credit hours.

\* If not taken during M.S.



# Graduate Courses

---

## Elective Courses in Statistics Department

STAT504	Non-Parametric Statistical Inference and Methods	(3-0) 3
STAT505	Sampling Theory and Methods	(3-0) 3
STAT509	Applied Stochastic Processes	(3-0) 3
STAT518	Statistical Analysis of Designed Experiments	(3-0) 3
STAT529	Statistical Bioinformatics	(3-0) 3
STAT545	Longitudinal Data Analysis	(3-0) 3
STAT553	Actuarial Analysis and Risk Theory	(3-0) 3
STAT554	Computational Statistics	(3-0) 3
STAT557	Statistical Modeling I	(3-0) 3
STAT558	Statistical Modeling II	(3-0) 3
STAT559	Applied Multivariate Analysis	(3-0) 3
STAT560	Logistic Regression Analysis	(3-0) 3
STAT562	Univariate Time Series Analysis	(3-0) 3
STAT563	Multivariate Time Series Analysis	(3-0) 3
STAT564	Advanced Statistical Data Analysis	(3-0) 3
STAT565	Decision Theory and Bayesian Analysis	(3-0) 3
STAT566	Reliability Theory and Methods	(3-0) 3
STAT567	Biostatistics and Statistical Genetics	(3-0) 3
STAT568	Statistical Consulting	(3-0) 3
STAT570	Data Handling and Visualization	(3-0) 3
STAT574	Statistics and Data Science Computing	(3-0) 3
STAT575	Computational Tools for Data Science	(2-2) 3
STAT576	Neural Networks for Data Science	(3-0) 3
STAT577	Big Data Analytics	(3-0) 3
STAT578	Artificial Intelligence and Data Science	(3-0) 3
STAT579	Statistical Pattern Recognition	(3-0) 3
STAT580	Stochastic Processes in Machine Learning	(3-0) 3

# Faculty



**Ayşen Dener Akkaya**

*Professor of Statistics*

**Research Fellow:** Stanford University, USA, Feb.2022-April 2022

**Post-Doctoral Fellow:** Princeton University, USA, 1997

**Ph.D:** Statistics Middle East Technical University, 1995

**M.Sc.:** Statistics Middle East Technical University, 1987

**B.Sc.:** Statistics Middle East Technical University, 1985

**RESEARCH AREA:** Random functions and applications, Time series analysis, Stochastic processes, Risk and decision analysis, Stochastic modeling in earthquake engineering, Seismic hazard analysis and earthquake insurance, Reliability, Robust inference, Autoregression

#### **PUBLICATIONS:**

- Bayrak, Ö. T., & Akkaya, A. D. (2018). Inference of autoregressive model with stochastic exogenous variable under short-tailed symmetric distributions. Iranian Journal of Science and Technology, Transactions A: Science, 42, 2105-2116.
- Kalafat, E., Oztürk, E., Kalaylıoğlu, Z., Akkaya (Dener) and A., Khalil, A. (2020). Correspondence: The Ratio of Umbilical and Cerebral Artery Pulsatility Index in the Assessment of Fetal Risk: Numerator and Denominator Matter. Ultrasound Obstet Gynecol, 56, 290-295.
- Akkaya, A. D., Bayrak, Ö.T, Yentur, B. (2022). Adaptive Estimation of Autoregression Model under Long-tailed Symmetric Distribution. Communications in Statistics - Simulation and Computation.



**Barış Sürücü**

*Professor of Statistics*

**Research Fellow:** McMaster University, Canada, 2004

**Ph.D:** Statistics, Middle East Technical University, 2003

**M.Sc:** Statistics, Middle East Technical University, 1999

**B.Sc:** Statistics, Middle East Technical University, 1996

**RESEARCH AREA:** Robust Statistics and Inference, Statistical Distributions and Their Applications, Goodness-of-Fit Tests, Outlier Detection, Decision Support Systems, Big Data Analytics, Data Science, Reliability and Survival Models, Survey Design and Sampling, Environmental Statistics, Financial Statistics, Quality Control, Circular Statistics, Sports Statistics

**Teaching Interests:** Statistical Theory, Reliability, Generalized Linear Models, Stochastic Processes, Multivariate Statistics

#### **PUBLICATIONS:**

- Sürücü, B. (2015). Testing for censored bivariate distributions with applications to environmental data. Environmental and ecological statistics, 22, 637-649.
- Surucu, B., & Sazak, H. S. (2011). Graphical methods for reliability data. Wiley Encyclopedia of Operations Research and Management Science, 1-11.
- Tiku, M. L., & Sürücü, B. (2009). MMLEs are as good as M-estimators or better. Statistics & probability letters, 79(7), 984-989.



**Burçak Başbuğ Erkan**

*Professor of Statistics*

**Ph.D:** Statistics, London School of Economics and Political Science, 2007

**M.Sc:** Statistics, University of Warwick, 2001

**B.Sc.:** Statistics, Middle East Technical University, 1999

**RESEARCH AREA:** Disaster risk management, Disaster risk reduction, Resilience, Disaster and/or development, M&E, Accountability, Insurance, Statistics

#### **PUBLICATIONS:**

- Comfort, L. K., Çelik, S., & Erkan, B. B. B. (2023). Learning from stress: Transforming trauma into sustainable risk reduction. Journal of Design for Resilience in Architecture and Planning, 4(Special Issue), 25-38.
- Erkan, B. B., Crawley, F. P., & Murray, V. (2023). The Need for Data Policy in Times of Crisis An IDPC CODATA Report Following a Scientific Workshop Held on 22 October 2022 in Leiden, The Netherlands. Data Science Journal, 22(1).
- Basbug Erkan B. B., Disaster Risk Management, "in: A Companion to Modern Turkey's Centennial Political, Sociological, Economic and Institutional Transformations since 1923", Ozerdem Alparlan, Ozturk Ahmet Erdi, Editor, Edinburgh University Press, Edinburgh, pp.553-569, 2023.

# Faculty



**Ceren Vardar Acar**

*Assoc. Professor of Statistics*

**Ph.D:** Department of Mathematics and Statistics, Bowling Green State University (BGSU), Bowling Green, Ohio, USA, 2002-2008.

**M.Sc:** Department of Statistics, Middle East Technical University (METU), 2000-2002

**B.Sc:** Department of Statistics, Middle East Technical University (METU), 1995-1999

#### RESEARCH AREA:

Stochastic processes and their applications mainly in finance, insurance and physics, Path properties of stochastic processes, Risk analysis, Probability theory, Financial engineering, Asymptotical Distributions, Extreme Values and Risk Management, Financial Mathematics/Statistics

#### PUBLICATIONS:

- Çağlar, M., Kyprianou, A., & Vardar-Acar, C. (2022). An optimal stopping problem for spectrally negative Markov additive processes. *Stochastic Processes and their Applications*, 150, 1109-1138.
- Vardar-Acar, C., Çağlar, M., & Avram, F. (2021). Maximum drawdown and drawdown duration of spectrally negative Lévy processes decomposed at extremes. *Journal of Theoretical Probability*, 34, 1486-1505.
- Avram, F., Grahovac, D., & Vardar-Acar, C. (2020). The W, Z scale functions kit for first passage problems of spectrally negative Lévy processes, and applications to control problems. *ESAIM: Probability and Statistics*, 24, 454-525.



**Ceylan Talu Yozgatligil**

*Professor of Statistics*

**Ph.D:** Statistics, Temple University, Philadelphia, USA, Ph.D., Statistics, January 2007

**M.Sc:** Statistics, Middle East Technical University, Ankara, TÜRKİYE.

**B.Sc:** Statistics, Middle East Technical University, Ankara, TÜRKİYE.

#### RESEARCH AREA:

Time Series Analysis: Univariate and Multivariate Methods, Machine/Deep Learning Applications in Time Series, Data Science, Hidden Markov Models, Statistical Inference, Spatio-temporal Analysis, Survival Analysis and Competing Risks

#### PUBLICATIONS:

- Oflaz Z., Yozgatligil C., Selcuk-Kestel A. S. (2023) Modeling comorbidity of chronic diseases using coupled hidden Markov model with bivariate discrete copula. *Statistical Methods in Medical Research*. 32 (4), 829-849.
- Oflaz Z., Yozgatligil C., Selcuk-Kestel A. S. (2022) Estimation of disease progression for ischemic heart disease using latent Markov with covariates. *Statistical Analysis and Data Mining*, 16(1), 16-28.
- Neslihanoglu S., Ünal E., Yozgatligil C. (2021). Performance comparison of filtering methods on modelling and forecasting the total precipitation amount: A case study for Muğla in Turkey, *Journal of Water and Climate Change*, 12(4), 1071-1085.



**Fulya Gökalp Yavuz**

*Assoc. Professor of Statistics*

**Post-doc:** Statistics, Purdue University, College of Science, West Lafayette, IN, July 2016-July 2017

**Ph.D:** Statistics, Yıldız Technical University, College of Art and Science, Istanbul, Türkiye, Jan 2011-Jun 2015

**M.Sc:** Statistics, Purdue University, College of Science, West Lafayette, IN, Jan 2008-May 2010

**B.Sc:** Statistics, Hacettepe University, College of Science, Ankara, Türkiye, Sep 2001-Jun 2006

#### RESEARCH AREA:

Statistical Theory and Applications, Linear Mixed Models, Parallel Programming, Robust Distributions, Shrinkage Methods for Model/Variable Selection, Machine Learning Methods, Statistical Methods in Big Data Applications, R Programming

#### PUBLICATIONS:

- Çakar, S. & Gokalp Yavuz, F. (2023) Nested and robust modeling techniques for fNIRS data with demographics and experiment related factors in n-back task, *Neuroscience Research*. vol.186, pp.59-72
- Oltulu, O. & Gokalp Yavuz, F. (2023). Examining parallelization in kernel regression. *Soft computing*.
- Guney, Y., Arslan, O., & Gokalp Yavuz, F., (2022). Robust estimation in multivariate heteroscedastic regression models with autoregressive covariance structures using EM algorithm. *Journal of Multivariate Analysis*, 191, 105026.

# Faculty



**İnci Batmaz**  
*Professor of Statistics*

**Research Fellow:** Imperial College London, UK, 2017.

**Ph.D:** Computer Engineering, Ege University, İzmir, Turkey, 1987-1993.

**Ph.D:** Dissertation (Fulbright Scholar), Carnegie-Mellon University (CMU), Pittsburgh, PA, USA, 1988-1991.

**M.Sc:** Computer Engineering, Ege University, İzmir, Türkiye, 1984-1986.

**B.Sc.:** Department of Statistics, Middle East Technical University (METU), Ankara, Türkiye, 1979-1984.

#### RESEARCH AREA:

Data Science & Analytics (DS&A), Streaming & Big Data, Data Mining (DM), Computational Statistics, Machine Learning, Statistical Computing, Statistical Software Development, Early Warning System Development, Response Surface Methodology (RSM), Metamodeling, Design of Experiments (DOE), Statistical and Computational Modeling, Quality Improvement, Environmental Statistics, System Simulation

#### PUBLICATIONS:

- Yerlikaya-Özkurt, F., Yazıcı, C., & Batmaz, İ. (2023). cmaRs: A powerful predictive data mining package in R. *SoftwareX*, 24, 101553.
- Batmaz, İ. (2022). Data Mining Approaches to Meteorological Data: A Review of NINLIL Climate Research Group Studies. *Operations Research: New Paradigms and Emerging Applications*, 107-135.



**Mohammad Qamarul Islam**  
*Professor of Statistics*

**Ph.D:** Statistics, 1989, Middle East Technical University, Türkiye

**M.Sc:** Statistics, 1972, University of Karachi, Pakistan

#### RESEARCH AREA:

Statistical Inference, Robust Statistical Methods, Nonparametric Methods, Multivariate Analysis, Regression Analysis, Design and Analysis of Experiments, Econometrics

#### PUBLICATIONS:

- Burbano Moreno, Á. A., Melo-Martinez, O. O., & Qamarul Islam, M. (2021). Inference in Multiple Linear Regression Model with Generalized Secant Hyperbolic Distribution Errors. *Ingeniería y Ciencia*, 17(33), 45-70.
- Dogan, E., Islam, M. Q., & Yazici, M. (2018). Real exchange rates and job flows: evidence from Turkey. *Applied Economics*, 50(42), 4489-4499.
- Dogan, E., Islam, M. Q., & Yazici, M. (2017). Firm size and job creation: evidence from Turkey. *Economic research-Ekonomska istraživanja*, 30(1), 349-367



**Özlem İlk Dağ**  
*Professor of Statistics*

**Ph.D:** Department of Statistics, Iowa State University, USA, 2000-2004

**M.Sc:** Department of Statistics, Iowa State University, USA, 1998-2000

**B.Sc:** Department of Statistics, Middle East Technical University, Türkiye, 1994-1997

#### RESEARCH AREA:

Statistical data analysis, Longitudinal data, Multilevel modeling, Interactive (dynamic) graphics, Statistical computing, Bayesian data analysis, Education, Biostatistics

#### PUBLICATIONS:

- Dağ, O., Kasıkcı, M., İlk, O., & Yesiltepe, M. (2023). GeneSelectML: a comprehensive way of gene selection for RNA-Seq data via machine learning algorithms. *Medical & biological engineering & computing*, 61(1), 229-241.
- Cinar, O., Iyigun, C., & İlk, O. (2022). An evaluation of a novel approach for clustering genes with dissimilar replicates. *Communications in Statistics-Simulation and Computation*, 51(12), 7458-7471.
- Inan, G., & İlk, O. (2019). A marginalized multilevel model for bivariate longitudinal binary data. *Statistical Papers*, 60, 601-628.

# Faculty



**Vilda Purutçuoğlu**

*Professor of Statistics*

**B.Sc.:** Department of Philosophy, Anadolu University, Türkiye, 2017-2020.

**Ph.D:** Department of Mathematics and Statistics, Lancaster University, U.K. 2004-2007.

**M.Sc.:** Department of Statistics, Middle East Technical University, Türkiye, 2002-2004.

**Minor:** Department of Economics, Middle East Technical University, Türkiye, 2000-2002.

**B.Sc.:** Department of Statistics, Middle East Technical University, Türkiye, 1998-2002.

#### **RESEARCH AREA:**

Bioinformatics, Systems biology, Network analyses  
Computational biology, High dimensional modeling and inference, Neuroscience, Neurofinance, Social statistics

#### **PUBLICATIONS:**

- Erkuş, E. C., & Purutçuoğlu, V. (2023). A new collective anomaly detection approach using pitch frequency and dissimilarity: Pitchy anomaly detection (PAD). *Journal of Computational Science*, 72, 102084, 1-31.
- Purutçuoğlu, V., Weber, G.W. & Farnoudkia, H. (2023). *Operations Research: New Paradigms and Emerging Applications*, CRC/Taylor and Francis Press.
- Purutçuoğlu, V. & Ayyıldız, E. (2023). *Statistics in the Field of Bioinformatics*, Nobel Press, 2nd Edition.



**Zeynep Işıl Kalaylıoğlu**

*Professor of Statistics*

**Ph.D:** Department of Statistics, North Carolina State University, USA, 2002.

**M.Sc.:** Department of Statistics, North Carolina State University, USA, 1999, with minor degree in Computational Engineering

**B.Sc.:** Department of Statistics, Middle East Technical University, Türkiye, 1995

#### **RESEARCH AREA:**

Bayesian Statistics, Circular Statistics, Model Selection

#### **PUBLICATIONS:**

- Camli, O., Kalaylıoğlu, Z., & SenGupta, A. (2022). Variable selection in linear-circular regression models. *Journal of Applied Statistics*, 1-25.
- Kalaylıoğlu, Z. (2022). Analysis of correlated circular and extremal data with a flexible cylindrical distribution. *Environmental and Ecological Statistics*, 29(1), 207-222.
- Camli, O., & Kalaylıoğlu, Z. (2021). Bayesian predictive model selection in circular random effects models with applications in ecological and environmental studies. *Environmental and Ecological Statistics*, 28, 21-34.

# STATISTICS

METU Dept. of Statistics    [stat.metu.edu.tr](http://stat.metu.edu.tr)  
Dumlupinar Bulvari, 06800    Phone: (0312) 210 2960  
Ankara-Turkiye                [stat@metu.edu.tr](mailto:stat@metu.edu.tr)