

STATISTICS MAJOR OFFERED BY THE DEPARTMENT OF MATHEMATICS AT KEIO UNIVERSITY: CURRICULUM AND STUDENT-LED EVENTS

Mihoko Minami and Tomoshige Nakamura

Department of Mathematics
Keio University, Yokohama, Kanagawa 223-8522 JAPAN
mminami@math.keio.ac.jp

The Department of Mathematics at Keio University has offered two major programs for undergraduates since it was restructured in 1981. The graduate program of the Department of Mathematics also offers major and minor programs in mathematics and statistics that provide the students the opportunity to work towards masters or doctoral degrees under close supervision. As one of a limited number of such universities, our statistics major program consists of not only standard statistics courses but also courses and seminars that introduce advanced theories and applications in various areas of statistics and machine learning. In addition to these courses and seminars, the events led by graduate students, Keio Data Fest and Summer Research Camp, also offer good opportunities for students to learn to analyze real-world data.

INTRODUCTION

Mathematics originated from the natural human desire to unearth the wonders of numbers and diagrams. Mathematical Science is the collective term for the many fields of study that have been established at the boundary between mathematics and various other sciences. In mathematics, we learn various fields of pure mathematics such as geometry, algebra, and applied mathematics. In statistics, we place an emphasis on building statistical expertise that can help solve problems in engineering, medical science, and many other fields. The Department of Mathematics of Keio University was established in 1974 as the Department of Mathematical Engineering and was renamed and restructured in 1981 when the Faculty of Engineering was reorganized and expanded as the Faculty of Science and Technology. Our “Department of Mathematics” is in fact a department of “Mathematical Science.” Our department has two majors: mathematics and statistics. Currently, there are 26 faculty members in the department consisting of 2-4 specialists in each field of pure mathematics, applied mathematics, and statistics. Graduate programs for master’s and doctoral degrees are also offered. Students can work for master’s or doctorates in either Science or Engineering, according to their degree objectives.

UNDERGRADUATE STUDY AT THE FACULTY OF SCIENCE AND TECHNOLOGY AND DEPARTMENT OF MATHEMATICS

An academic year for undergraduates at Keio University commences in April similar to most Japanese universities. The semester system is used as the academic calendar system at Keio University with some changes to accommodate quarter system courses. Classes meet once a week, one 90-minute session for a two unit class, two consecutive 90-minute sessions for a four unit class. At the Faculty of Science and Technology, approximately 1000 first-year students are categorized into five groups based on areas of interest and take substantial basic courses in mathematics such as introduction to linear algebra and calculus, physics, chemistry, biology, laboratory experiments, and programming.

From the second year onwards, each undergraduate student selects a department. There are eleven departments in the Faculty of Science and Engineering, but the departments students can select depend on the groups they belong to. Only the students in Group 2 can select the Department of Mathematics, and on average, approximately 60 students select the Department of Mathematics each year. Curriculum for second-year students at the Department of Mathematics consists of 4 compulsory courses: “Foundation for Mathematical Science I,” “Foundation for Mathematical Science II,” “Computer Science & Laboratory,” and “Introduction to Statistical Science,” and 16 elective courses including linear algebra, algebra, topology, probability, and computer science. “Introduction to Statistical Science” is a new course offered from the 2018 academic year. This is

described later in more detail. Second-year elective courses teach basic concepts of mathematics, and most students take many of them.

In their third year, students study the foundations of the field of their interest and related fields. Compulsory courses for third-year students are “Mathematical Analysis and Its Exercise” for the mathematics major and “Statistical Science and Its Exercise” for the statistics major. Although students are recommended to take both courses, currently, many students do not take both courses. This is why we have changed the system from the current year as described later. Elective courses in fields other than statistics include real analysis, finite mathematics, algorithms, probability theory, mathematical planning, information processing, actuarial mathematics, and risk theory as well as courses in the fields of pure mathematics. Elective courses in statistics are described in more detail in the next section.

Currently, students have to decide their major at the beginning of their third year although they can change their major later if they fulfill requirements. However, this system will also change from the next year allowing students to choose their major in December, that is, one month before the end of the fall semester and the academic year. At the same time, each student chooses his/her supervisor for graduation study.

In their final year of undergraduate study, students take advanced courses in their field and conduct graduation research. One noteworthy aspect is that students who choose a field in probability and applied mathematics such as mathematical planning and numerical computation can choose either mathematics or statistics as their major. The main difference between the two majors other than their compulsory courses is that a bachelor’s degree in science is awarded for a mathematics major and a bachelor’s degree in engineering is awarded for a statistics major on completing the requirements. There are four faculty members in statistics, three in probability, and four in applied mathematics in our department. Roughly speaking, among approximately 60 students, half choose statistics as their major and half of the statistics major students choose statistics for their graduation study.

UNDERGRADUATE STATISTICS COURSES

The Department of Mathematics currently offers ten undergraduate statistics courses represented in Table 1. Until the current third-year students, their first statistics courses by the Department of Mathematics were offered in the third year of undergraduate study, although there is a standard basic statistics course by a professor at the Department of Administration Engineering as a second-year elective course.

Table 1. Undergraduate statistics courses offered at the Department of Mathematics

Course No.	Year	Semester	Course name	Units	Notes
#1	Second	Fall	Introduction to Statistical Science	2	1)
#2	Third	Spring	Statistical Science and Its Exercise	4	2)
#3			Mathematical Statistics 1 and Its Exercise	4	
#4		Fall	Data Analysis and Its Exercise	4	
#5			Mathematical Statistics 2	2	
#6			Data Design	2	3)
#7	Fourth	Spring	Nonlinear Models	2	4), 5)
#8			Time Series Models	2	
#9		Fall	Special Lecture in Statistical Mathematics 1	2	5)
#10			Special Lecture in Statistical Mathematics 2	2	5)

1) A compulsory course for all students in Department of Mathematics from the 2018 academic year.

2) A compulsory course for the statistics major.

- 3) This course will be abolished from the 2019 academic year.
- 4) This course name does not reflect the course contents.
- 5) These courses are taught by adjunct professors.

There are two main sequences of statistics courses: Mathematical Statistics and Statistical Modeling. “Statistical Science and Its Exercise” (#2) in spring semester of the third year is the compulsory course for statistics major. This course introduces how to explore, interpret, and analyze data and basic of statistical modeling using R. Subsequently, “Data Analysis and Its Exercise” (#4) in the fall semester introduces various statistical models and procedures. These courses meet one day every week in two consecutive class periods. A lecture is provided in the first period, and the second period is an exercise session. In each exercise session, a few theoretical problems and several exercises on data analysis are assigned. Six teaching assistants help students to solve problems and perform data analysis with statistical package R.

Mathematical Statistics courses are “Mathematical Statistics 1 and Its Exercise” (#3) and “Mathematical Statistics 2” (#5). These courses are not compulsory because not all students in statistics major choose statistics for their graduation study. However, students who choose statistics as their graduation study are strongly recommended to take these courses.

In the final year of undergraduate study, four elective classes are offered. “Time Series Models” (#8) is taught by a faculty member of the department, but the other three courses are taught by young adjunct professors. In the 2018 academic year, these three courses deal with the following topics:

- Statistical inferences with random combinatorial models
- Asymptotic theory
- Statistical theory of multivariate probability distributions

Table 2. Topics of the “Introduction to Statistical Science” course

Topic No.	Contents
#1	Introduction of Statistical Science, Explore data with R
	Qualitative/quantitative variables, histogram, box and whisker plot, five-number summary, distribution function, density function, normal distribution, mean and variance, outliers
#2	Explore the relationship between variables
	Correlation and regression, Pearson’s/Spearman’s/Kendall’s correlation coefficients, correlation and stratification, partial correlation
#3	Linear regression model (statistical inference, confidence interval, hypothesis testing)
	Method of least squares, residuals, adjusted R-squares, variable selection
#4	Discriminant analysis
	Discrimination rule, Fisher’s discrimination method, linear discrimination
#5	Principle component analysis
	Constrained maximization of quadratic forms, principal score, contribution ratio
#6	Contingency table analysis
	Statistical independence, chi-square test, Fisher’s exact test
#7	Data design
	survey sampling, random sampling, stratified sampling

Additionally, “Actuarial Mathematics and Risk Theory” as a course for third-year students and “Actuarial Mathematics” for fourth-year students are also offered. From the 2018 academic year, a new course “Introduction to Statistical Science” will commence in the Fall semester. This is a

compulsory course for all second-year students in the Department of Mathematics. Table 2 presents the topics discussed in this course.

We made this new course compulsory course for second-year students and changed the time that students make choices of their major from the beginning of the third year of their undergraduate study to the end of the third year for the current second-year students. This is partly because not only faculty members in statistics, but also other members in pure mathematics and applied mathematics are now aware that all students need to know more about statistics and data analysis, and it is better for students to learn about this before they decide their major.

GRADUATION RESEARCH

In the final year of undergraduate study, students belong to their supervisors' laboratory and conduct graduation research. At the Department of Mathematics, each laboratory has a seminar using two to three class periods once a week for undergraduates in general. Graduation thesis is not compulsory for a bachelor's degree at the Department of Mathematics, but all four professors in statistics require their students to write a graduation thesis, and the presentation meeting of graduation thesis is held jointly with four laboratories at the beginning of February.

GRADUATE STATISTICS MAJOR PROGRAM

After graduation, approximately half of the students progress to the master's course each year at Department of Mathematics at Keio University. In Japan, a student often goes to the graduate school of the same university that he/she graduates from. Many students at Keio University choose to do so and work with the same professor for a master's degree as for their undergraduate study. Additionally, a few to several graduate students join our master's program from other departments, other universities in Japan, or foreign countries. Progressing to the graduate school of the same university has some advantages such that the students and their supervisors can make study/research plans in the long term. Master's students take graduate course offered by Department of Mathematics and other departments and conduct independent study in the first year and master thesis study in the second year. Master thesis is mandatory for all graduate students at Keio University.

The Department of Mathematics offers various types of graduate courses. Graduate statistics courses and courses in related fields are presented in Table 3.

Table 3. Graduate statistics courses and courses in related fields

Course No.	Course title	Units
#1	Topics in Statistical Science A	2
#2	Topics in Statistical Science B	2
#3	Topics in Statistical Science C	2
#4	Mathematical Finance	2
#5	Special Course in Data Science	2
#6	Special Course in Information Mathematics	2
#7	Topics in Life Insurance Mathematics	2

Courses #1 to #4 are provided by faculty members of the Department of Mathematics. Topics introduced in Course #1 to #3 are not fixed and are chosen from current research topics and methods, and advanced statistical theory and methods that graduate students should know. Topics introduced in recent lectures are as follows:

- Statistical Machine Learning: Theory for the Lasso, SVM, Boosting, and others
- Bayesian Statistics: Theory and Methods
- Extensions of Linear Regression Model: GLM, GLMM, spline smoothing and GAMM
- Nonparametric models

“Special course in Data Science” is offered every other year. This course is designed to be omnibus, and five prominent researchers provide lectures on their research and recent hot topics using three class periods. Five topics introduced in recent year classes are as follows:

- An Introduction to Empirical Likelihood Methods
- Instrumental Variables Methods for Medical research
- A Frontier of Machine Learning: Applications in Material Science and Life Science
- Basic Theories and Applications of Kernel Method
- Modelling of Point Process Network and Statistical Analysis

Every two to three years, the Department of Mathematics offer an intensive course (#6) that covers basic to advanced theory on a specific topic on statistics, machine learning, and related fields. For example, the recent intensive courses discussed the following topics:

- Markov Chain Monte Carlo methods
- Statistical Causal Inference

Many graduate students also take courses offered by other departments/centers such as “Applied Statistical Analysis” and “Advanced Financial Engineering” by the Center for Open Systems Management, and “Artificial Intelligence” by the Center for Information and Computer Science.

INDEPENDENT STUDY AND MASTER THESIS STUDY

Generally, master’s students meet their supervisors every week and discuss their study and research. Research areas of four faculty members in statistics in our department include biomedical statistics, statistical machine learning, computational finance, actuarial science, environmental, and biological risk assessment and Bayesian prediction in addition to theoretical statistics and general statistical science that are common interests. Thus, students choose their research topic from a wide range of fields in statistics and machine learning based on their interest.

WEEKLY SEMINAR AND SEMINARS BY GUEST SPEAKERS

A weekly statistics seminar is held every Wednesday where all students and faculty members from the statistics laboratory gather. Speakers are graduate students, and they introduce their results or provide a summary of independent study and master thesis study once in every semester. Making a presentation is a good opportunity for students to gain perspective on their research and to learn how to communicate their research effectively. Seminars by guest speakers are also held occasionally.

STUDENT-LED EVENTS: KEIO DATA FEST AND SUMMER RESEARCH CAMP

Students learn not only by taking courses and attending seminars but also through their active involvement in teaching each other by participating or organizing the following events.

Keio Data Fest

Inspired by DataFest at UCLA (Gould, Baumer, Çetinkaya-Rundel, and Bray 2014, Gould 2014, Bray 2014), our graduate students started Keio Data Fest in 2017. Keio Data Fest is a competition but with tutorials and is held in March. During this three-day event, teams of up to three undergraduate students work to extract insights from a large data set. This provides a good opportunity for students to engage in solving real-life big data problems that are more complex than the ones they work with in class. A unique feature of Keio Data Fest is that graduate students provide tutorials about software package R, handling of big data, statistical methods, and data analysis during the event so that students who are interested in data analysis but have not had an opportunity to work with it can also join the event and make a good start for learning data analysis.

At the 2017 Keio Data Fest, 16 students participated in and analyzed the Pronto cycle share data obtained from a website that is now closed. Pronto Cycle Share was a public bicycle sharing system in Seattle, Washington, that operated from 2014 to 2017. Two datasets were provided to participants: the station dataset that contained information of 60 docking stations and the trip data of approximately 260,000 trips that was far larger than any data set the students had encountered in the classroom. Participants tackled extract insights from the data combined with external data.

The 2018 Keio Data Fest is held between March 5-7. The datasets are provided by *Net Protections, Inc.*, which is a Japanese company that specializes in post payment service. These datasets contain a large number of transactions of purchases and customer information. Approximately 20 to 30 students register for this event this year. It is noteworthy that they are not only from the Department of Mathematics, but also from other departments in the Faculty of Science and Engineering, the Faculty of Economics, the Faculty of Law, the Faculty of Environment and Information Studies, and the Faculty of Policy Management. The event also provides an effective opportunity for students to interact with people from different backgrounds.

Summer research camp

During the school breaks for summer, in many cases in September, statistics professors and their undergraduate and graduate students lodge together for three to four days in a research camp. Students make teams randomly at the camp and work together to solve problems and analyze the data that a group of a few graduate students prepare. The tasks are often about theories, models, and algorithms that can provide a basis for students' graduation research and project research. Summer research camp provides a good opportunity to students to learn statistics intensively and understand deep insights. More importantly, we have a lot of fun!

SUMMARY AND CONCLUSIONS

We have introduced the curriculum of the statistics major and activities for the students and by the students who study statistics at the Department of Mathematics, Keio University. We believe that the strength of the graduates from the Department of Mathematics is their ability to engage in logical mathematical thinking acquired by solid course work and graduation/thesis study. Statistical courses offered by the Department of Mathematics tend to emphasize more on theoretical aspects of statistical and machine learning models and procedures rather than applications. However, students have many options and opportunities to deepen and expand their knowledge and experiences in various different ways cooperatively with other students.

REFERENCE

- Bray, A. (2014) A Festival of Data: Student Perspectives, *Amstat News*, 447, 3-4.
- Gould, R. (2014) DataFest: Celebrating Data in the Data Deluge. In K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9, July, 2014), Flagstaff, Arizona, USA*. Voorburg, The Netherlands: International Statistical Institute. *The Proceedings of ICOTS9*.
- Gould, R., Baumer, B., Çetinkaya-Rundel, M., & Bray, A. (2014) Big Data Goes to College. *Amstat News*, 444, 17-19.