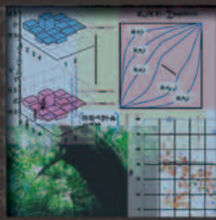
Research Organization of Information and Systems

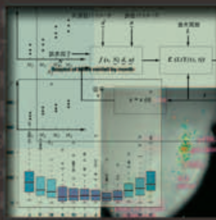
The Institute of Statistical Mathematics

2008-2009

Statistical Modeling

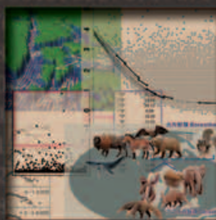


Data Science

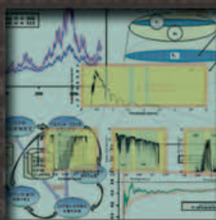


Mathematical Analysis and Statistical Inference

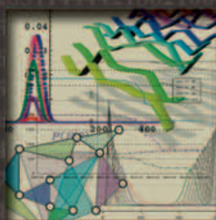
Prediction and Knowledge Discovery



Risk Analysis

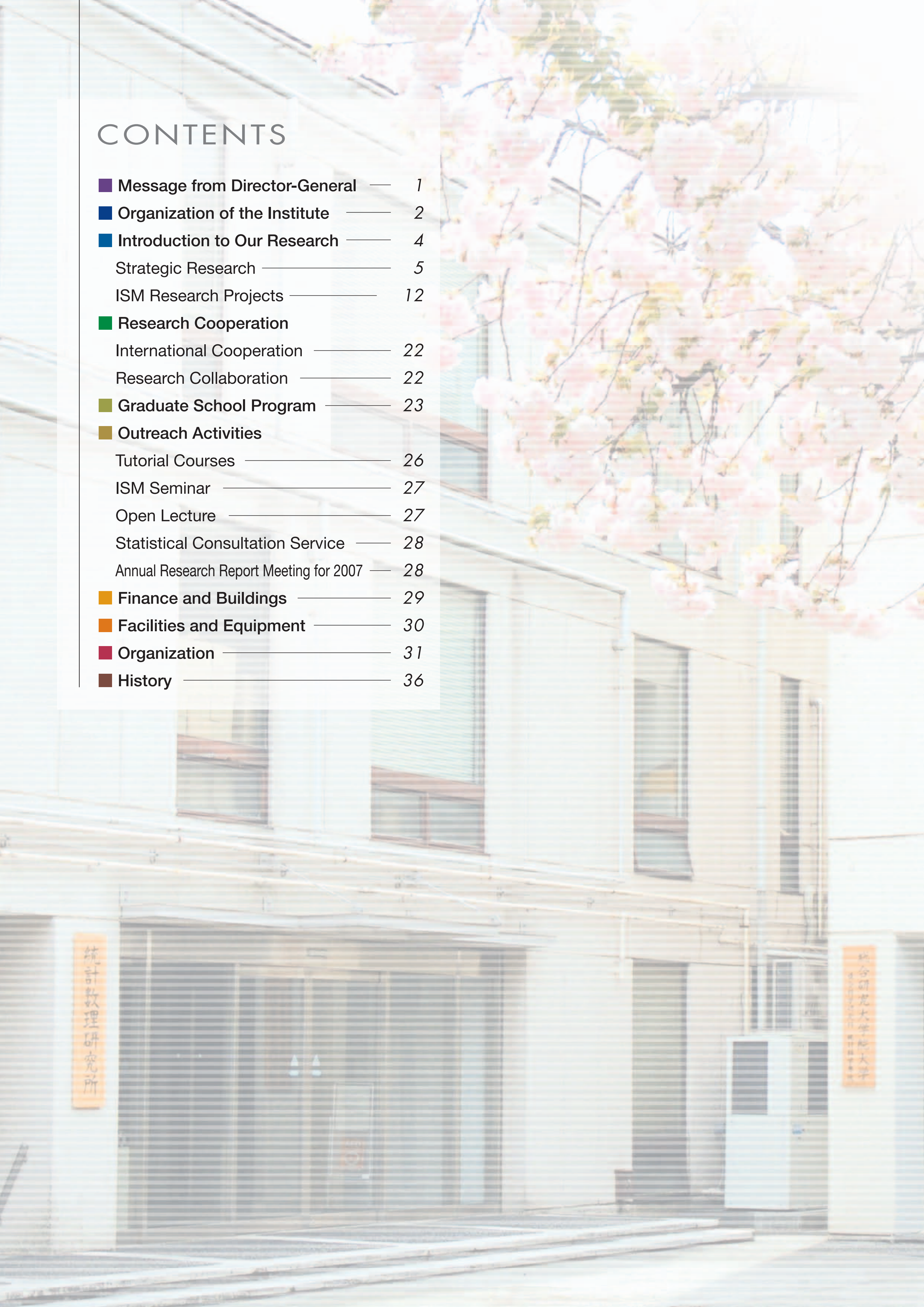


Research Innovation



Engineering and Technical Support

ISM



CONTENTS

Message from Director-General

As ICT (information and communication technologies) make rapid progress in areas such as measurement, telecommunication, information retrieval and computation, the human society and technological environment of academic research is changing significantly. In the late 20th century, progress in the fields of simulation-based computational science, as well as theoretical and experimental science, fostered advances in previously intractable fields such as nonlinear phenomena and complex systems, which were difficult to be treated by conventional analytic methods. With the coming of the 21st century, increased availability of a variety of large-scale data sets has facilitated progress in many fields of research, and a method for prediction and discovery called the 4th science, which is based on large-scale data analysis, is making significant progress. Only with the establishment of the 4th science can we use both the deductive approach (computational science) and the inductive approach (large-scale data science), which are based on computers, in addition to conventional theoretical and experimental sciences that depend on researchers' intuition, and establish the scientific methodology essential for research in the information age.

With this in view, the Institute of Statistical Mathematics has promoted research aimed at establishing methods for prediction and knowledge discovery based on large-scale data sets by setting up the Prediction and Knowledge Discovery Research Center and the Risk Analysis Research Center. In the current fiscal year, we also set up a Research Group for Reliability and Quality Assurance of Service and Product in the Risk Analysis Research Center and started research on the reliability of products and quality of services, which have become an important social problem in recent years. In addition, we have set up a Research Innovation Center in order to systematically support research activities aimed at innovation in statistical mathematics, and started research on social survey information, statistical inference based on function analysis, Monte Carlo algorithm and random numbers.

The Institute of Statistical Mathematics, which serves as an inter-university research organization, has promoted joint research projects with various organizations in academic research. Establishment of such joint research coordination and research networks is increasingly important in a society that is increasingly complex and diverse. In particular, our effort to establish an NOE (Network of Excellence), which the Risk Analysis Research Center has promoted, is attracting attention in many areas. While focusing on these activities, the Institute of Statistical Mathematics intends to contribute to society by fostering the progress and development of statistical modeling, large-scale data analysis and mathematical approaches.

We would appreciate your continued understanding and support for the activities of the Institute of Statistical Mathematics.

Genshiro Kitagawa
Director-General
The Institute of Statistical Mathematics

Organization of the Institute

Basic Research

Department of Statistical Modeling

The Department of Statistical Modeling conducts research on the modeling of causally, temporally and/or spatially interrelated complex phenomena, including intelligent information processing systems. It also conducts on model-based statistical inference methodologies.

■ Spatial and Time Series Modeling Group

The Spatial and Time Series Modeling Group works on modeling and inference for the statistical analysis of time series, spatial and space-time data, and their applications to prediction and control.

■ Intelligent Information Processing Group

The Intelligent Information Processing Group works on concepts and methods for the extraction, processing and transformation of information in intelligent systems, motivated by an active interest in practical problems in engineering and science.

■ Graph Modeling Group

The Graph Modeling Group works on analyses of the data generated by systems with a graph structure and on the modeling required in order to reconstruct the original system.

Department of Data Science

The Department of Data Science aims to develop research methods for surveys, multidimensional data analyses, and computational statistics.

■ Survey Research Group

The Survey Research Group focuses on research related to statistical data collection and data analyses.

■ Multidimensional Data Analysis Group

The Multidimensional Data Analysis Group studies methods for analyzing phenomena grasped on multidimensional space and ways for collecting multidimensional data.

■ Computational Statistics Group

The Computational Statistics Group studies sophisticated uses of computers in statistical methodology such as computer-intensive data analyses, computational scientific methods and statistical systems.

Department of Mathematical Analysis and Statistical Inference

The Department of Mathematical Analysis and Statistical Inference carries out research into general statistical theory, statistical learning theory, the theory of optimization, and the practice of statistics in science.

■ Mathematical Statistics Group

The Mathematical Statistics Group is concerned with aspects of statistical theory and probability theory that have statistical applications.

■ Learning and Inference Group

The Mathematical Statistics Group develops statistical methodologies that enable researchers to learn from data sets and to properly extract information through appropriate inference procedures.

■ Computational Mathematics Group

The Computational Mathematics Group studies computational algorithms together with mathematical methodologies used for statistical modeling in the sciences.

Strategic Research

Prediction and Knowledge Discovery Research Center

The Prediction and Knowledge Discovery Research Center studies the statistical modeling and inference algorithms that can be used to extract useful information from the huge amount of data which complex systems produce, and thus attempts to solve real-world problems in many different scientific domains, especially genomics, earth and space sciences.

■ Molecular Evolution Research Group

The Molecular Evolution Research Group researches the area of molecular phylogenetics, and seeks to develop statistical methods for inferring evolutionary trees of life using DNA and protein sequences.

■ Data Assimilation Research Group

The Data Assimilation Research Group aims at developing new, advanced data assimilation techniques to combine different information from dynamical simulation and observation data.

Statistical Seismology Research Group

The Statistical Seismology Research Group is concerned with the evaluation of seismicity anomalies, detection of crustal stress changes, their modeling, and the probability forecasting of large aftershocks and earthquakes.

Statistical Genome Diversity Research Group

The Statistical Genome Diversity Research Group aims to construct novel methodologies for learning and inference from a variety of data sets in the rapidly growing area of bioinformatics.

Risk Analysis Research Center

The Risk Analysis Research Center is pursuing a scientific approach to the study of the increased uncertainty and risk associated with the increasing globalization of society and the economy. The center is also constructing a network for risk analysis in order to contribute to the creation of a reliable and safe society.

Food and Drug Safety Research Group

The Food and Drug Safety Research Group aims to develop the statistical framework and methodology of quantitative risk evaluation for substances ingested by the human body.

Environmental Risk Research Group

The Environmental Risk Research Group studies the statistical methodologies related to environmental risk and environmental monitoring.

Financial Risk and Insurance Research Group

The Financial Risk and Insurance Research Group explores the use of statistical modeling methods to quantify the risks involved with financial instruments and insurance products.

Research Group for Reliability and Quality Assurance of Service and Product

The Research Group for Reliability and Quality Assurance of Service and Product aims to increase the safety of products and services by developing statistical methods that contribute to quality assurance and reliability and by promoting the adoption of these methods in the industrial world.

Research Innovation Center

The objective of this center is to establish innovative research in statistical mathematics to keep up with new trends in the academic and real worlds. The center carries out original research projects, ranging over both pure and applied frontiers.

Social Survey Information Research Group

The Social Survey Information Research Group collects several social survey results with a view to assembling them to develop a statistical methodology to describe the social world.

Functional Analytic Inference Research Group

This group aims to develop nonparametric methodology for statistical inference using reproducing kernel Hilbert spaces given by positive definite kernels, and to apply these techniques to causal inference problems.

Advanced Monte Carlo Algorithm Research Group

The Advanced Monte Carlo Algorithm Research Group aims to develop Markov Chain Monte Carlo and Sequential Monte Carlo algorithms and study their applications.

Random Number Research Group

This group carries out research into random number generation, including hardware random number generators, and testing randomness with methods such as time series analysis.

Research Support

Center for Engineering and Technical Support

The Center for Engineering and Technical Support assists the development of statistical science by managing the computer systems used for statistical computing, facilitating public outreach, and supporting the research activities of both staff and collaborators.

Computing Facilities Unit

The Computing Facilities Unit is in charge of managing computer facilities, software for research, networking infrastructure and network security.

Information Resources Unit

The Information Resources Unit is responsible for maintaining an extensive library and an electronic repository, and is in charge of planning statistical education courses to popularize research results.

Media Development Unit

The Media Development Unit is in charge of the publication and editing of research results and is responsible for public relations.

Introduction to Our Research

Strategic Research

The ISM is divided into 3 categories; Basic Research, Research Support, and Strategic Research. The Basic Research section provides a framework for fundamental research, accepting original contributions from all staff members, with an emphasis on developing statistical methods that can address current societal issues. Research Support acts precisely as its name implies, as a part of the Inter-University Research Organization. Strategic Research also combines the strengths of Basic Research, Research Support, and other universities and research institutes, in order to address important societal issues.

The Strategic Research section has established the Prediction and Knowledge Discovery Research Center and the Risk Analysis Research Center. These two centers are organized as follows:

- **Prediction and Knowledge Discovery Research Center (PKDRC)**
  - Molecular Evolution Research Group
  - Data Assimilation Research Group
  - Statistical Seismology Research Group
  - Statistical Genome Diversity Research Group
- **Risk Analysis Research Center (RARC)**
  - Food and Drug Safety Research Group
  - Environmental Risk Research Group
  - Financial Risk and Insurance Research Group
  - Research Group for Reliability and Quality Assurance of Service and Product

ISM Research Projects

The institute also emphasizes the importance of its own and the exploratory research. Every year we invite researchers from within the institute to propose new research projects. A relatively sizable amount of funding is allocated to approved projects, and strategic research leading to future development is encouraged. In 2007, we have approved the following ten projects. No limitation was placed on theme or category.

Research Theme	Project Leader
Reconstructing functional neuronal circuits from optical imaging data	Yoshiyasu Tamura
Mathematical molecular ecology – Long-term forest research by genetic spatial point process	Shimatani Kenichiro
Rainfall data analysis using a Bayesian generalized linear model	Toshio Ohnishi
Speaker verification with non-audible murmur	Tomoko Matsui
Control and modeling of high-order vibrating systems and compensation of input nonlinearity	Yoshihiko Miyasato
Study on computational algorithms for state-space modeling	Takashi Tsuchiya
Estimation of quantum states from observation	Yukito Iba
Developing statistical models for bycatch in tuna fisheries	Mihoko Minami
Sampling selections after the great merger of the Heisei era	Wataru Matsumoto
Construction of statistical methodologies in quality engineering	Toshihiko Kawamura

Establishment of a New Center

In April 2008, our institute established a Research Innovation Center in order to establish innovation in statistical mathematics by promoting research projects based on freewheeling thinking, including exploratory projects, while emphasizing new trends in society and academic research.

- **Research Innovation Center**
  - Social Survey Information Research Group
  - Functional Analytic Inference Research Group
  - Advanced Monte Carlo Algorithm Research Group
  - Random Number Research Group

The following introduces research activities and projects conducted in the last fiscal year by the several groups at the Strategic Research Center.

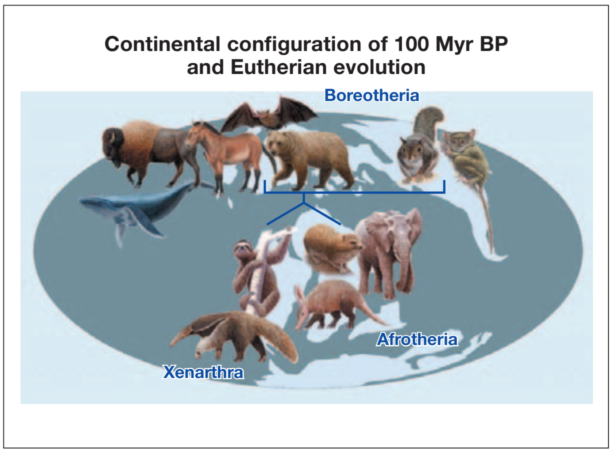
Molecular Evolution Research Group

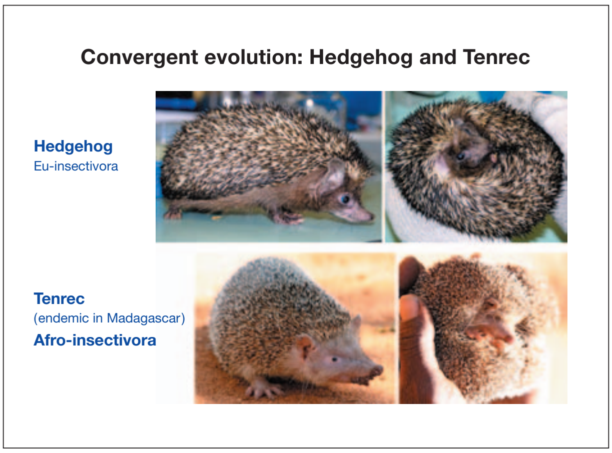
Toward Integrative Understanding of Biodiversity

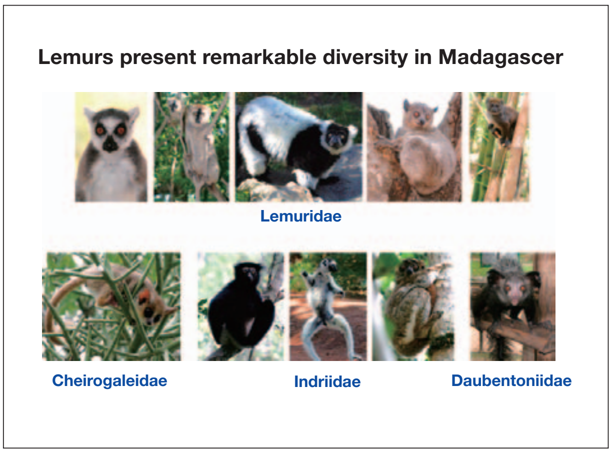
The aim of this project is to understand the biodiversity on Earth from various view points encompassing from molecular level to ecological level. Evolutionary view is indispensable in the integrative understanding of the biodiversity, and the methods for inferring molecular phylogeny are essentially important. In this project, we are developing models of nucleotide substitutions during evolution and methods for phylogenetic inference based on maximum likelihood. In developing the statistical methods, we simultaneously try to solve real problems of biological importance.

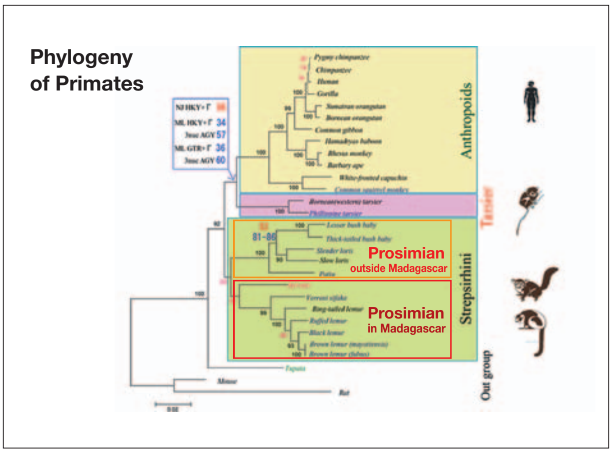
The biological problems we are working on include evolution of vertebrates such as mammals and birds, evolution of land plants, and the biodiversity of Malagasy fauna. Madagascar has been isolated from any continents for a long geological time, and has developed a unique fauna with high level of combined species richness and endemism. In Madagascar, we are working on the biogeographic problems of tenrecs and lemurs (mammals) and baobabs (plants). Spiny tenrec is morphologically very similar to hedgehog, and had long been classified in Insectivora together with hedgehogs, moles, and shrews, but molecular phylogenetic analyses clarified that tenrecs belong to Afrotheria which includes elephants, hyraxes, and aardvarks, and that the similarity between spiny tenrec and hedgehog is due to convergent evolution. Estimation of the time-scale of lemurs and baobabs gave important clues in clarifying the origin of these groups of organisms in Madagascar.

Jun Adachi









Data Assimilation Research Group

Data Assimilation: Time-dependent Information Fusion from Numerical Simulation and Large-scale Observation

Data Assimilation (DA) is a technique for a synthesis of information from a dynamic (numerical) model and observation data. It is an emerging area in earth sciences, particularly oceanography, stimulated by recent improvements in computational and modeling capabilities and the increase in the amount of available observations. In statistical methodology, DA can be formulated in the generalized state space model, where the system and observation model correspond to large-scale numerical model-based simulations and large-scale satellite- and/or ground-based measurement systems, respectively. Past studies for DA employed a linear Gaussian state space model and applied Kalman filter. The Kalman filter based methods, however, do not allow for the strong nonlinear and/or non-Gaussian disturbance behaviors. Many phenomena in earth sciences tend to be discussed in terms of a complex system in which the nonlinear non-Gaussian fluctuations (disturbances) play an important role. The nonlinear non-Gaussian DA method needs to be developed in an attempt to improve a performance of prediction ability of our environment. We are therefore constructing new computation methods based on the sequential DA methods and conducting five DA projects. One is done with the ensemble Kalman filter that assimilates the TOPEX/Poseidon altimetry to the coupled ocean-atmosphere simulation model. The second DA is done with a particle filter for Tsunami simulation model to correct bottom topography. The third DA project is aimed at finding an uncertainty in sea bottom topography which plays an important role in conducting an ocean tide simulation. The fourth is the DA project to estimate distributions of ring current ions and electric potential in the inner magnetosphere by assimilating the series of the ENA data obtained by the HENA imager on board of IMAGE satellite into a kinetic ring current model (CRCM). The last project is to apply the DA methodology to combine a simulation model with observed data like microarray gene expression data for understanding biological pathways.

Tomoyuki Higuchi

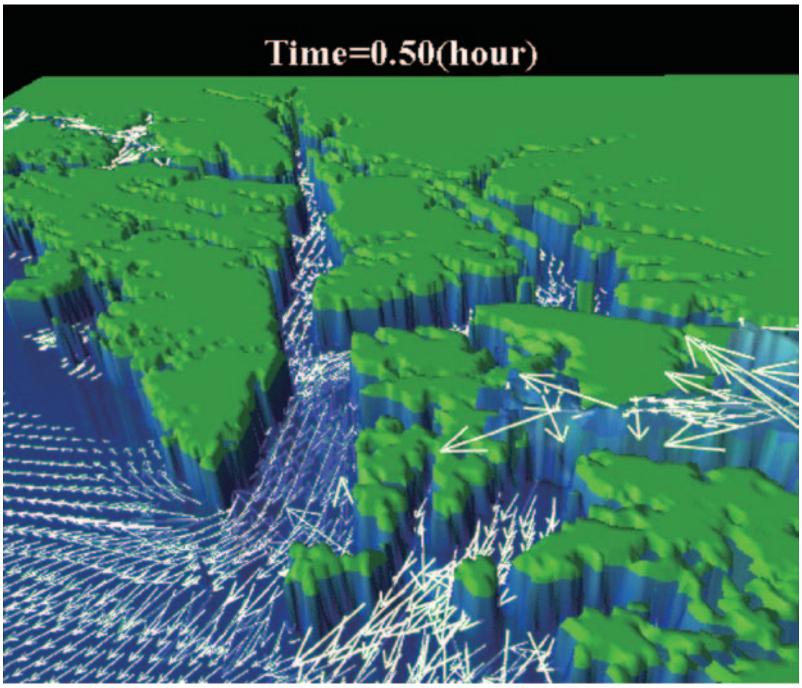


Figure: Visualization of a data assimilation experiment for high resolution ocean tide modeling

Statistical Seismology Research Group

Immediate and Updated Forecasting of Aftershock Hazard

Aftershock probability

Typically, aftershock occurrences are modeled as a non-stationary Poisson process in time with intensity rate obeying the Omori-Utsu law. The magnitude distribution of earthquakes is modeled by the Gutenberg-Richter's (G-R) law when all aftershocks were detected throughout the observation period. A method of aftershock probability forecasting was proposed based on the joint intensity rate of aftershocks, where the parameters are estimated by the maximum likelihood method in the respective empirical laws. Then, the probability of observing one or more aftershocks in an interval of magnitude range and time range is calculated. Public forecasts have been implemented by the responsible agencies in California and Japan.

Difficulty of the forecasting

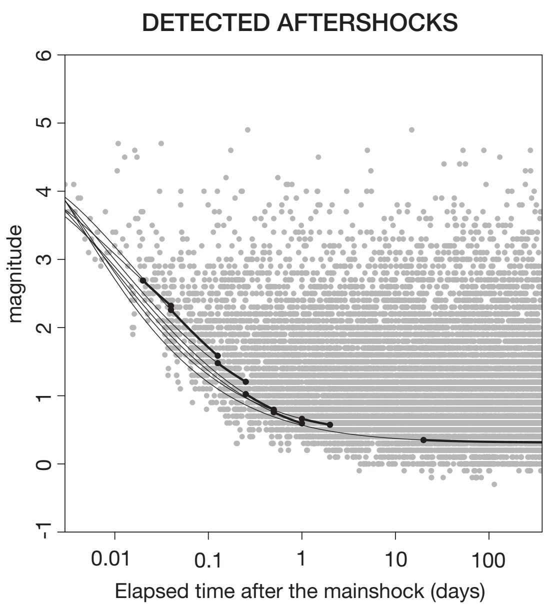
However, the considerable difficulty in the above procedure is that, due to the contamination of arriving seismic waves, detection rate of aftershocks is extremely low during a period immediately after the main shock when the forecasting is most critical for public in the affected area. Gray disk in the left diagram shows magnitude of the detected aftershock at each time. Therefore, for the forecasting of a probability during such a period, from the one day though, they adopt a generic model with a set of the standard parameter values except for a particular parameter.

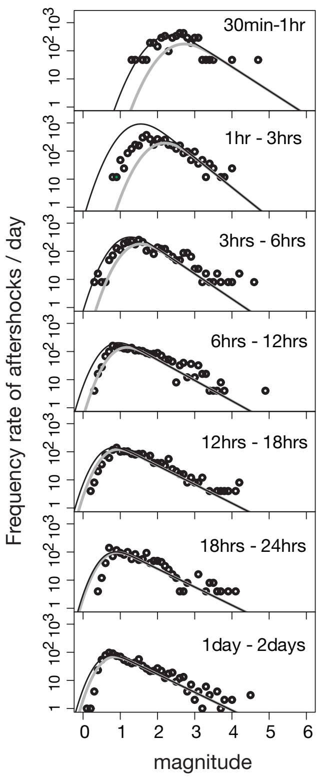
Immediate forecasting taking the detection rate into account

For more effective and urgent forecasting, we propose an estimation method for predicting the underlying occurrence rate of aftershocks of any magnitude range, based on a magnitude frequency model that combines the G-R law with the detection rate function. This procedure enables quasi-real-time probability forecasting of aftershocks immediately after the main shock (cf., right diagram), when the majority of large aftershocks are likely to occur. The presently proposed prediction procedure depends on the capacity for determining aftershock magnitude. Thin and thick smooth curves in the left panel indicate the estimated and extrapolated (predicted) magnitude of 50% detection at each updated time, respectively. The solid curves in the right panel represent the predicted magnitude frequency distribution, and the circles show the corresponding empirical frequencies of the aftershocks.

Furthermore, like the G-R relationship, the Ishimoto-Iida formula provides an empirical relationship of the maximum amplitude size recorded at a seismograph, so that the similar prediction of large seismic intensity at a particular site can be implemented more directly and quickly.

Yosihiko Ogata





Statistical Genome Diversity Research Group

Pattern Recognition for Association Study to Phenotypes

Objective

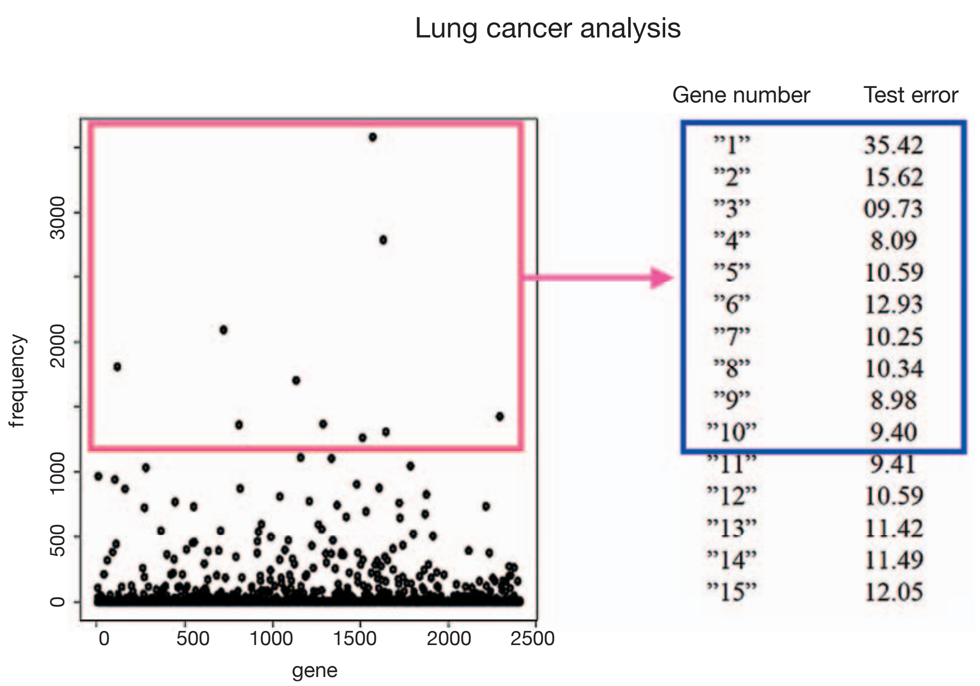
Our research group aims at proposing and developing statistical methodology leading to knowledge discovery for huge genome and omics datasets in genomic sciences and technology.

Data analyses often face to double-edged errors of false discovery and missing discovery, called 'p>>n problem'. We challenge establishment for novel statistical methods by discussing various possible approaches in reexamination of conventional methods.

Association study for phenotype

In medical science treatment effect is typical as an important phenotype. It is reported that an anti-cancer drug has different effects to patients, and sometimes leads to fatal adverse effect to specific individuals. If proper prediction for an individual to the effect is given before treatment, then it greatly improves the medical procedure. We study the prediction method based on genome and omics datasets such as microarray, proteome and SNP typing to discover several genes associated with the phenotype. For example, we do pattern recognition by learning a dataset of gene expressions for subject groups with different degree of effect. In this case the data dimension p is about 30 thousands and the sample size is usually several tens. We consider statistical machine learning methods for the problem. We discuss 'Exhaustive Boost' learning algorithm for solving a difficult problem in which any single prediction machine is not able to attain good performance. The method repeats randomly K-fold validation for boost learning and we select genes by the frequency of genes trained with good score. Figure 1 shows the gene frequency in 2000 repetitions of 30-combined machine by AdaBoost and we confirm that the best ten-gene prediction selected according to the frequency table gets a good performance.

Shinto Eguchi



Kawakita et al. (2006)

Food and Drug Safety Research Group

Benefit-risk Balance Evaluation of Food and Drug

■ Mission

Social responsibility of the food and drug safety program in RARC is to design appropriate databases and statistical methods for the benefit-risk balance evaluation of food and drug. We are carrying out cooperative research projects with visiting researchers and members from the research group of data management and statistics in Japan Pharmaceutical Manufacturers Association.

■ Benefit/risk analysis of drugs based on large-scale database

Various large-scale databases have been available for the scientific benefit / risk evaluation and risk management of post-marketing drugs in Europe and North America (Figure 1). In Japan, however, such a large-scale database does not exist. As the first step to improve this situation, we are building original databases using data collected from pre-marketing clinical trials and post-marketing surveillance (PMS), and we conduct benefit / risk analyses of drug effect.

Antihypertensive drugs and Non-Steroidal Anti-Inflammatory Drugs (NSAIDs) are often used together because both of them are treatments of the common diseases of hypertension and arthritis. Rofecoxib, a new type of NSAID which blocks selectively the enzymatic activity of cyclooxygenase 2, was expected to improve gastrointestinal side effects, but was withdrawn in 2004 because of its cardiovascular risks. The interaction between antihypertensive drugs and NSAIDs was examined based on PMS database of antihypertensive drugs containing the data of 143,509 patients. It is indicated that the effectiveness of antihypertensive drugs was attenuated by co-administration of NSAIDs.

■ Ad hoc study on a specific drug safety issue

An issue of drug safety often becomes a social concern, and the immediate scientific elucidation is required. An appropriate study design, implementation and statistical analysis are necessary to get an accurate scientific quantitative answer (Figure 2). We participate in an ad hoc study on a specific drug safety issue as experts of statistical science and epidemiology.

Toshiharu Fujita

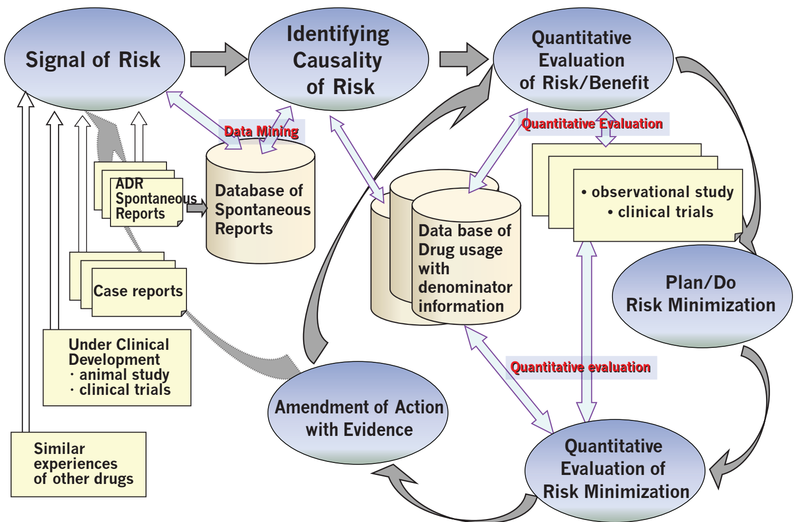
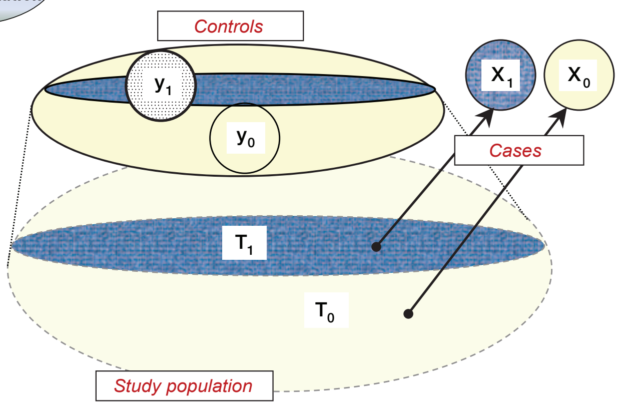


Figure 1: Risk management of Pharmaceuticals



T : person-time
X : incident cases
y : controls

Figure 2: Population based case-control study

Environmental Risk Research Group

The Effort of the Statistical Science towards Solution of an Environmental Problem

Mission

A risk analysis research center / environmental risk research section aims at performing the contribution towards solution of the environmental problem which is a modern subject by developing the optimal new statistical methodology for each environmental problem (for example, the problem of dioxin, the problem of global warming, the problem of continuous use of safe water). Moreover, in order to realize this purpose, in cooperation with the community of environmental science, research is carried out including a visiting teacher or a project researcher.

Source identification of dioxins

The toxicity of a certain dioxin (Fig.1) is stronger than that of the poison gas, sarin, and it is feared that dioxins threaten human health. The major sources of dioxins are activities of human beings including agricultural chemicals, bleaching and combustion processes. Dioxins emitted into environment are absorbed into the body of a human being through lungs and a food chain. Recently, the environmental contamination caused by dioxins is frequently detected in individual districts. To solve the problem, it is necessary to identify the sources. However, it is not easy to identify the sources, because there are many unknown sources of dioxins. Accordingly, we are developing statistical models which enable us to make inference on unknown sources. Additionally, to improve the accuracy of the inference, we are trying to complete the data of dioxins cooperating with National Institute of Environmental Studies and many environmental laboratories of local governments.

Error analysis of Retrieved CO_2 column density from the Earth Observation Satellite GOSAT

In recent year, Global warming is a growing concern. Greenhouse gas forcing has very likely caused most of the observed global warming. The Greenhouse Gases Observing Satellite (GOSAT) will be launched in early 2009. CO_2 and CH_4 which are greenhouse gases are observed by GOSAT. GOSAT will be the world's first satellite to observe the column abundance of carbon dioxide (CO_2) and methane (CH_4). The targets of the GOSAT mission are observation of CO_2 and CH_4 column abundance in 3-month average with 1% and 2% relative accuracy respectively from spectrum (Fig.2). We evaluate the retrieved CO_2 and CH_4 accuracy by numerical simulation. Further, the evaluation method of the retrieved CO_2 and CH_4 column abundance after launch will be developed.

Koji Kanefuji

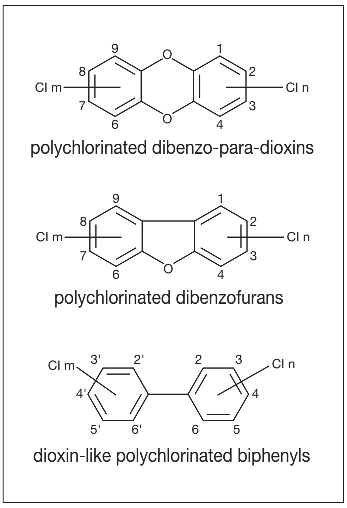


Figure 1 : Dioxins

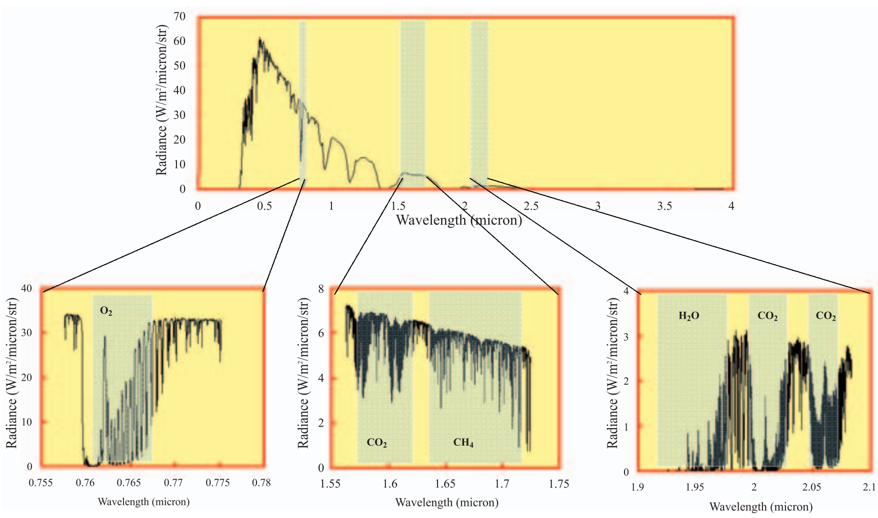


Figure 2 : Three-spectral bands of the GOSAT (bottom) and solar spectral radiance (top)

Financial Risk and Insurance Research Group

Game-theoretic Probability Approach to Mathematical Finance

Aim of our group

Financial Risk and Insurance Research Group mainly concerns the theory and practice in the quantitative analysis of risks involved with various financial instruments and insurance products from the view point of statistical modeling. Various research projects are under way not only by ISM researchers but by visiting and post doctoral researchers. This article introduces an ongoing project led by Dr. Masayuki Kumon, Research Associate of TRIC/ROIS.

Asset Trading as a Gambling Game

In 1950s statistical testing and estimation were unified as zero sum games played by human and nature; human wants to find the true characteristic from data, while nature tries to prevent such probes. This framework, well known as Statistical Decision Theory, has offered useful tools to analyze many statistical problems. Explaining by this analogy, it is possible to analyze some problems in probability and mathematical finance as zero sum games played by human (Skeptic) and nature or market (Reality). Skeptic tries to find any inclination of Reality through a gambling game, while Reality keeps Skeptic away from making profits. We call this framework the Game-Theoretic Probability Approach to finance. The most remarkable feature of this approach is that we do not assume certain probabilistic structures like Brownian motion but deduce them as natural consequences of characterization of betting games.

Optimal Strategy in Coin Tossing and SLLN

As an example, we see how we establish strong law of large numbers (SLLN) and its implication. Relate the head of a coin to the rise of stock market index, and the tail to other cases. Skeptic decides the strategy how he bets in advance, and Reality selects the outcome. In this game, Skeptic can asymptotically force Reality to obey the fair game by adopting a Bayesian model describing Reality's past behavior. The resulting rate and factor of convergence immediately give SLLN. Overlaying Reality's head selection rates (as deviations from the mean) on the bound derived from SLLN, we see it is when the Reality's path steps out of the bound that the Skeptic's asset process becomes more profitable. (See Figures 1 and 2.) This type of anomaly is almost always conceived in high frequent data, while it vanishes in daily observations implying the efficient market behavior of daily stock prices.

Yoshinori Kawasaki

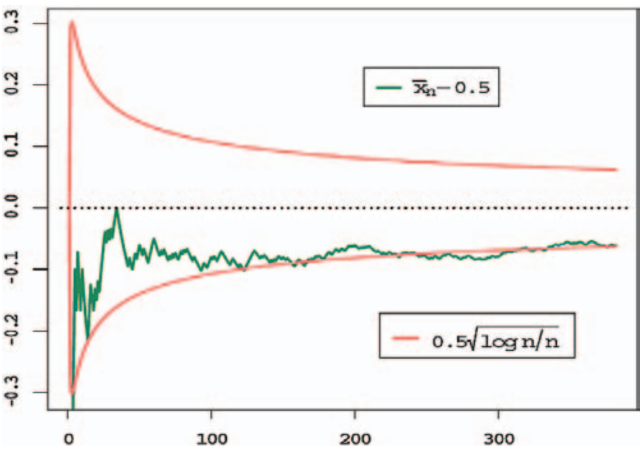


Figure 1: Reality's head selection rates (as deviations from the mean) and the bound derived from SLLN. The path steps out of the bound around 230th day.

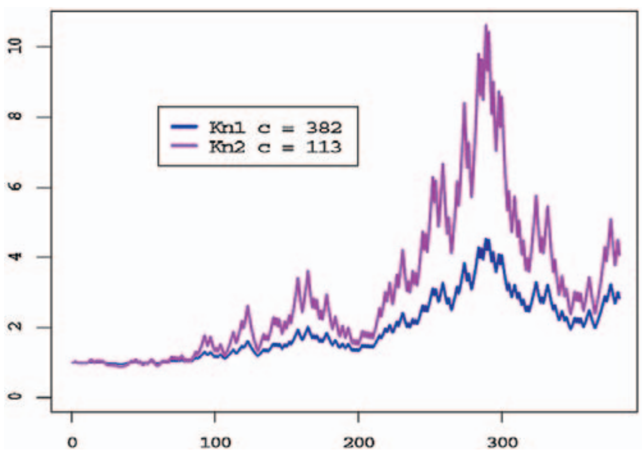


Figure 2: Asset process becomes profitable around 230th day when an "anomaly period" starts. Both figures appear by the courtesy of Dr. Kumon, TRIC/ROIS.

Reconstructing Functional Neuronal Circuits from Optical Imaging Data

Two putative respiratory rhythm generators (RRGs), pFRG and pre-BötC, exist in the brainstem. Signals originated from RRGs are dynamically processed in the neuronal network to form respiratory motor outputs. The popular approach has been employed correlation analysis between motor output and optical imaging data. However, the dynamic process of the respiratory motor pattern formation remains unanswered.

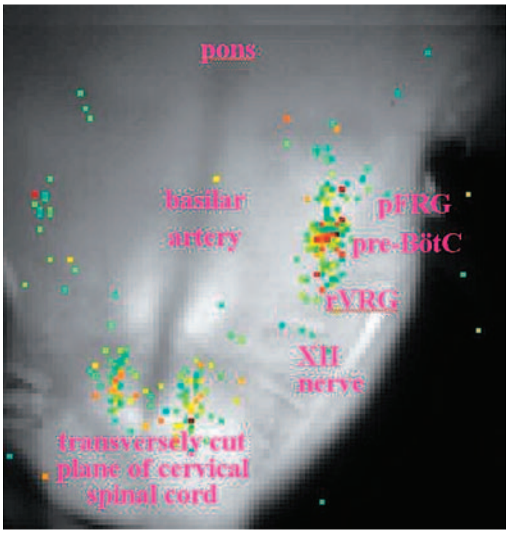


Figure1: Respiratory related regions on the rat ventral medullary surface extracted by the filter based on autoregressive model as innovation.

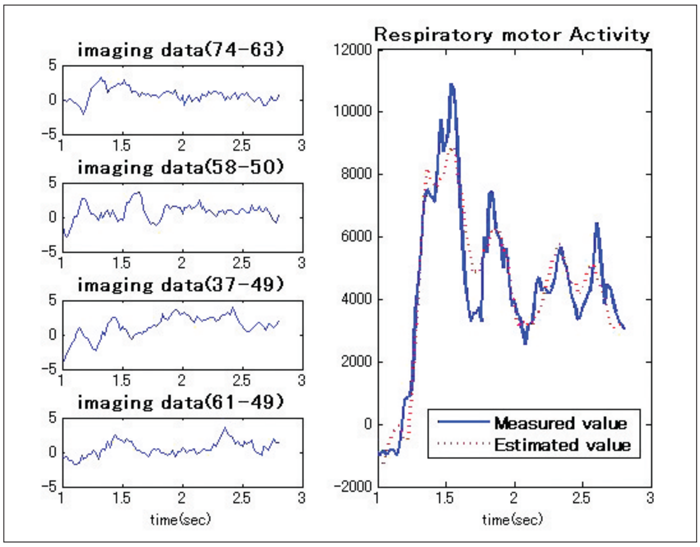


Figure 2: Comparison between the estimated and measured respiratory motor activities with multiple peaks.
(Left): Selected imaging data for estimation of the respiratory motor activity.
(Right): The measured value (the solid line) and the estimated value (the dotted line) of the respiratory motor activity.

We recorded respiratory-related optical signals from the ventral surface of the neonatal rat brainstem in vitro using a voltage-sensitive dye and showed that two modeling approaches to optical time series data can characterize the dynamic process. The experiments were pursued mainly in Hyogo College of Medicine, and methodological studies were developed by the members of The Institute of Statistical Mathematics.

In the first approach, autoregressive (AR) models were fitted to time series data of each pixel for the range sufficiently before the onset of inspiration. Then, the remaining time series data were filtered using these AR parameters to obtain its innovation (filter output). We could extract the respiratory-related regions by the intensity of the innovation even without the information of respiratory motor activity (Fig.1). In the second approach, we considered a parametric model assuming the respiratory motor activity as the output and optical signals of each pixel as the input, and estimated parameter values for each pixel. The model consisted of a threshold function and a delay transfer function. Data from a single pixel were sufficient to estimate the respiratory motor activity pattern with a single peak, and data from a pair of pixels were sufficient to estimate the activity pattern with two peaks. For the activity with multiple peaks, it was necessary to divide the respiratory period into sections, compute the model parameters for each range, and calculate a weighted sum of estimated values of a specific set of pixels (Fig.2). From these results, close input output relation between the central neuronal activity and the respiratory motor activity was numerically supported.

We suggest that these results reflect the mechanisms of the central respiratory rhythm and pattern formation.

Yoshitaka Oku*1, ***Fumikazu Miwakeichi****2, ***Shigeharu Kawai****3, ***Yasumasa Okada****4, ***Makio Ishiguro****5, ***and Yoshiyasu Tamura****5

*1: Hyogo College of Medicine, *2: Graduate School of Engineering, Chiba University, *3: The Graduate University for Advanced Studies, *4: Keio University, *5: The Institute of Statistical Mathematics

Mathematical Molecular Ecology
— Long-term Forest Research by Genetic Spatial Point Process

■ Long-term forest research

Forest dynamics is working over hundreds of years (Picture 1), which is much longer than human life span. Even though, we need to predict forest structure in future from currently available information, otherwise, we could not prevent forest disturbance, global warming and other environmental issues. Statistical models should play a central role in clarifying the forest ecosystem and its future prediction. In particular, recent development of molecular genetics enables us to estimate forest dynamics process that had been carried out many years ago. This project focuses on seed dispersal and individual female reproductive success.

■ Molecular data and stochastic model

Denote by $P(j, G)$ the Mendelian transition probability that an offspring with genotype G can be born from mother j with her pollen pool. Suppose that seeds were dispersed according to the dispersal function $f(r)$ and that mother j at $\boldsymbol{Z}_j$ had individual female reproduction of U_j. Let $s(\boldsymbol{x})$ be the survival probability at $\boldsymbol{x}$ from the dispersal to now. Then the spatial distribution of genotyped offspring (Fig. 1) is expressed by the inhomogeneous Poisson process with the density function

$$\lambda_G(\boldsymbol{x}) = s(\boldsymbol{x})\left\{\sum_{j=1}^{M} P(j,G)U_j\ f(\|\boldsymbol{x}-\boldsymbol{Z}_j\|)\right\}$$

Applying the maximum likelihood method, we can obtain estimates of seed dispersal and individual female reproductive success (thus, the effective population size).

When a reproductive process was repeated over generations, resulting spatial patterns can also be modeled by a spatial point process. Let the density of parents be λ. Suppose that parents produce offspring according to the Poisson distribution with the expected value of 1, thus the density is invariant over generations. If the seed dispersal follows the probability density function of the two-dimensional normal distribution, the genetic relatedness between two offspring at the n-th generation is given by

$$I_n(r) = \frac{\frac{1}{2}\sum_{h=1}^{n}\left(\frac{1}{4}\right)^h \exp\left(-\frac{r^2}{4h\sigma^2}\right)/4\pi h\sigma^2\lambda}{1+\sum_{h=1}^{n}\exp\left(-\frac{r^2}{4h\sigma^2}\right)/4\pi h\sigma^2\lambda}$$

if two offspring have interdistance of r. Because we can estimate genetic relatedness between any pair of trees from genotyped data (Fig. 2), we can estimate seed dispersal and the density of the population (effective population size) by model fitting.

■ Mathematical Molecular Ecology

Mathematical molecular ecology allows us to obtain information about forest dynamics that had been done many years ago.

Shimatani Kenichiro

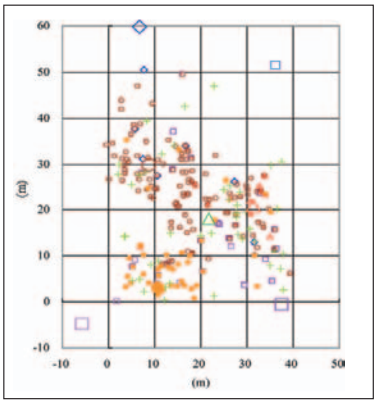


Figure 1:
Large symbols indicate mother trees and small ones offspring. The same symbol was used between offspring and its putative mother. Beech forest in northern Japan. The genetic data are microsatellite markers. The cooperative research with Tohoku University, Hiroshima University and Forestry and Forest Products Research Institute.

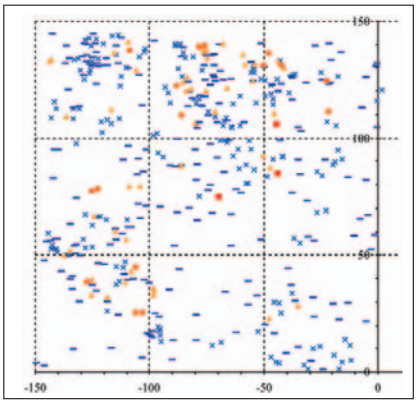


Figure 2:
Red to yellow trees are those who are surrounded by trees with high genetic relatedness whereas blue trees are surrounded with genetically unrelated trees. A fir population in Shiretoko Peninsula in Hokkaido, northern Japan. The cooperative research with Ryukyuu University, Kagoshima University, Tokyo University, Hiroshima University and Forestry and Forest Products Research Institute.

Picture 1:
Such a huge tree was probably born hundreds of years ago. This project aims at estimating the processes over long-term forest dynamics from genetic data and stochastic models.

Rainfall Data Analysis Using a Bayesian Generalized Linear Model

Non-negative data containing a lot of zeros

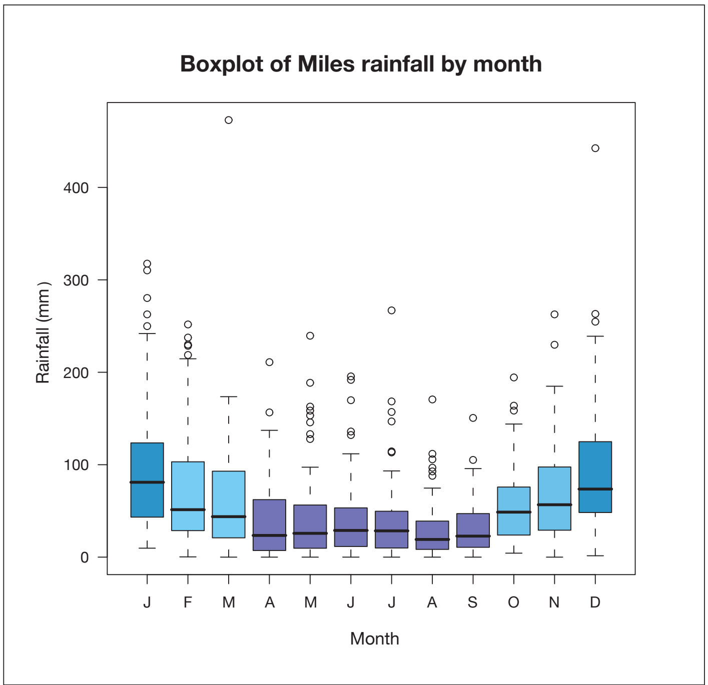
Rainfall data are peculiar in that they contain a lot of zero values. It sometimes happens that the total rainfall of a month is zero in Queensland, Australia, since it has been experiencing the worst draught in these several years. See the boxplot.

Southern Oscillation Index

SOI (Southern Oscillation Index) is known to affect the rainfall in Queensland. SOI is an index derived from the difference in air pressures between Darwin and Tahiti. Roughly speaking, if the air pressure in Darwin is higher, the amount of rainfall decreases due to dry wind from the desert; otherwise, wet wind from the Pacific Ocean brings rainfall.

Tweedie distribution

A Poisson mixture of gamma distributions, or Tweedie distribution, is useful in analyzing the rainfall data. It seems natural to the author that we assume the number of rainfall events follows a Poisson distribution and the amount of rainfall for each event is distributed according to a gamma distribution.



Data provided by the Queensland government

Tweedie generalized linear model

The generalized linear model, which is an extension of the normal linear model, is widely used in applied fields. Here we assume a generalized linear model based on the Tweedie distribution. We regard the total amount of rainfall of a month as a random variable, and postulate the following model:

Log (mean) = (Factor specific to observation point) + (SOI dependency).

Prior distribution

A prior distribution is assumed on the factor specific to the observation point. It is well known that Bayesian inference performs well when the parameter is high-dimensional. In the meantime, SOI dependency is assumed to be uniform. A Bayesian approach assuming a prior only on a subset of parameters was proposed in one of the "KYODO-RIYO-KENKYUs" of ISM.

Specialty of this study

A combination of the following two is the foremost characteristic of this study: (1) A co-researcher, Dr. Peter Dunn, University of Southern Queensland, has invented a program for evaluating with an excellent precision the density function of the Tweedie distribution, which has an extremely complicated form; (2) the author has revealed that the above generalized linear model has an analytical property that it admits a conjugate prior.

Result

We analyzed the total rainfall data in July from 1970 until 2007 for 45 observation points and estimated the SOI dependency in a quantitative way. We also proved its statistical significance.

Toshio Ohnishi

[Provision of data] Department of Primary Industry and Fishery of the Queensland government provided us with the rainfall data through another co-researcher, Mr. Daniel Burrell.

Speaker Verification with Non-audible Murmur

Non-audible murmur; NAM

NAM has been recently investigated in Shikano Laboratory of Nara Institute of Science and Technology(NAIST). NAM is produced in a voiceless utterance action and is uttered when one grumbles to oneself not intending to be heard by others, says prayers, or makes silent wishes. One only moves the speech organ while breathing, without vocal cord vibration or glottis narrowing. In NAM, the main information is below 4 kHz and information in higher frequency bands is not observed. Breath-induced vibration of the vocal tract is transmitted as NAM through the body directly to a condenser microphone worn on the surface of the skin below the mastoid bone.

Speaker verification using NAM

We are conducting collaborative research on speaker verification using NAM with Shikano Laboratory. In conventional speaker verification systems using normal speech, there is a problem that even though a text-dependent approach using a keyword phrase for each user is expected to provide high performance, this approach is not practical because of the opportunities for attacks involving interception and playback of live utterances. Using NAM instead of normal speech lets us safely take the text-dependent approach using keyword phrases, and it should provide effective and noise-robust authentication.

Effective use of keyword-specific acoustic features

We investigate two methods to make good use of keyword-specific acoustic features (Figure 1). First, as input data, we use NAM segments, which consist of several short-term feature vectors. Second, we introduce the global alignment (GA) kernel to better capture keyword-specific acoustic features. While standard kernels such as Gaussian and polynomial kernels are vector kernels, the GA kernel is a vector sequence kernel and constructed using similarities based on dynamic time warping (DTW) scores. The GA kernel can effectively handle time series with variable lengths and local dependencies between neighboring states of the time series. The definition of K_{GA} is shown at the upper right of Figure 1. $S(\pi_i)$ is the DTW score on the ith alignment π_i of all possible ones between two sequences **X** and **Y**. By using the two methods, we can achieve less than 1% speaker verification error with NAM data uttered by 55 speakers.

Tomoko Matsui

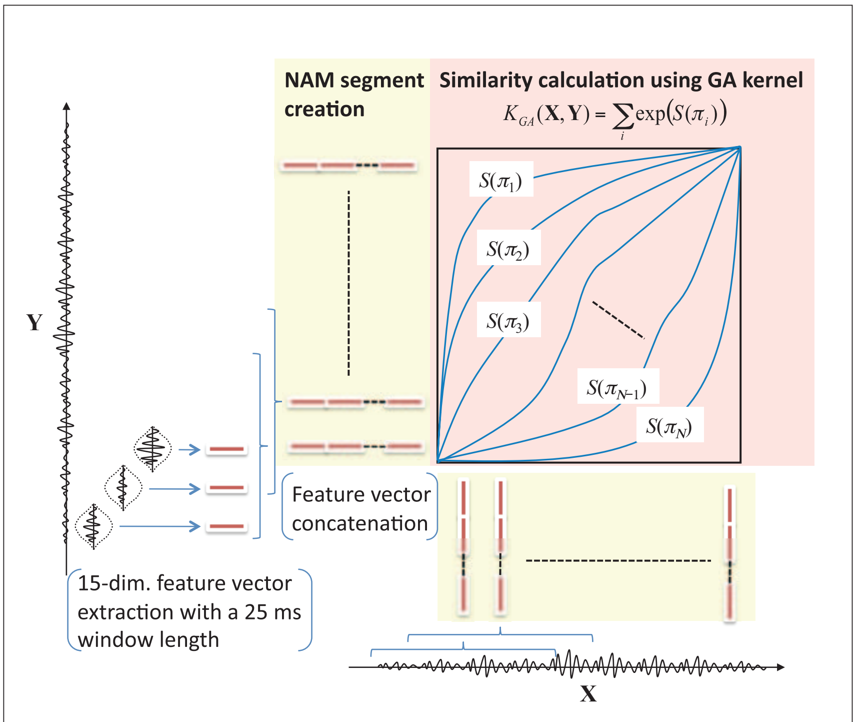


Figure 1. Effective use of keyword-specific acoustic features with NAM segments and GA kernel.

Members:
Hideki Okamoto, Mariko Kojima, Hiromichi Kawanami, Kiyohiro Shikano
(Nara Institute of Science and Technology)

Control and Modeling of High-order Vibrating Systems and Compensation of Input Nonlinearity

Control and Modeling of High-Order Vibrating Systems

Control of high-order vibrating systems, such as cantilever beams, flexible arms and large flexible structures, has been an important but difficult problem in the fields of robotics, space engineering and mechanical engineering. Since those high-order vibrating systems cannot be represented precisely via any models of fixed orders, it is hard to obtain appropriate controllers of finite dimensions which attain good attenuation of high-dimensional vibrating properties. Hence, to develop systematic procedures of modeling and control of high-order vibrating systems has been an important topic in statistical and control science.

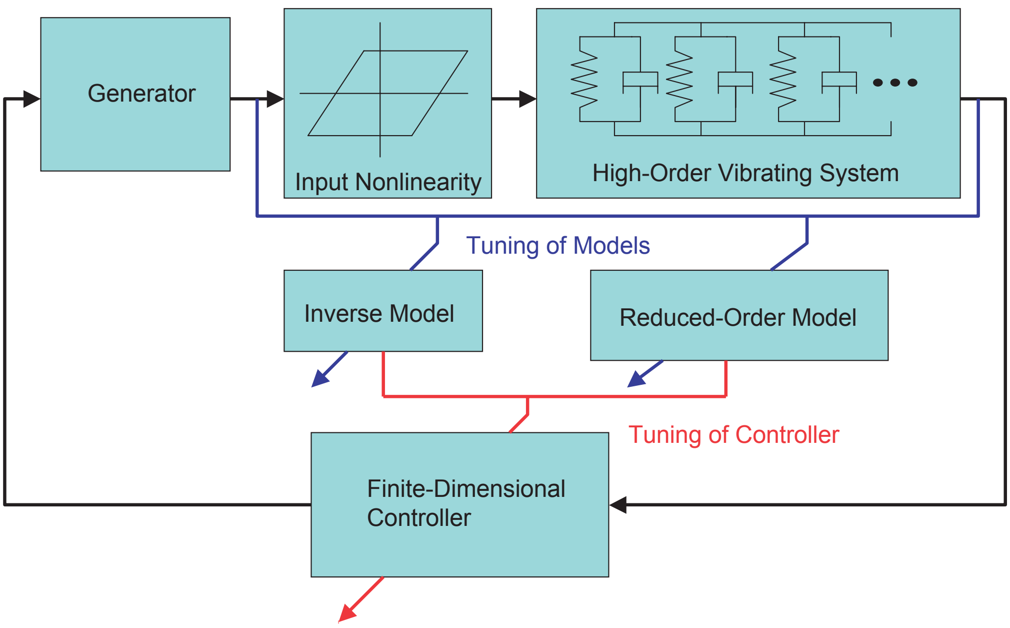
Input Nonlinearity

In order to suppress those vibrating systems, high-powered generators are necessary as control units. However, those controllers sometimes have uncertain nonlinear properties such as dead-zones and backlashes, and those cause additional undesirable vibrations to the total control systems.

Purpose and Result of Research Project

The purpose of the research project is to obtain integrated design procedures of control and modeling of high-order vibrating systems with input nonlinearities. First, reduced-order models of high-order vibrating systems are derived, and then, finite-dimensional compensators are designed for those reduced models. In order to regulate the effect of the neglected high-order modes, a stabilizer is introduced based on H_∞ control criterion where neglected modes (spill over modes) are regarded as external disturbances. Although there exist serious mismatches between the original vibrating systems and those reduced-order models, the stabilizer regulates those erroneous factors and makes overall systems bounded. Furthermore, control performance is prescribed explicitly by design parameters in the proposed H_∞ control schemes. For input nonlinearities, adaptive inverse models are introduced in order to compensate the effect of nonlinear characteristics by adjusting parameters in those inverse models. In this year, especially, design schemes of finite-dimensional adaptive H_∞ controllers for distributed parameter systems of hyperbolic type (infinite dimensional vibrating systems), and tuning schemes of compensations of input nonlinearities at tracking control of robotic manipulators, were investigated and several design tools were developed.

Yoshihiko Miyasato



Study on Computational Algorithms for State-space Modeling

The state-space model is one of the fundamental information processing models which has diverse applications such as time-series analysis, signal processing and image processing etc. In these applications, one is interested in the following three conditional distributions of the state at each time point, (i) predictive distribution: the distribution of the state $\boldsymbol{x}(t)$ at time t given the observations $\boldsymbol{Y}(0)$, ..., $\boldsymbol{Y}(t\text{-}1)$ at time 0, ..., t-1 ; (ii) filtering distribution: the distribution of $\boldsymbol{x}(t)$ given $\boldsymbol{Y}(0)$, ..., $\boldsymbol{Y}(t)$; (iii) smoothing distribution: the distribution of $\boldsymbol{x}_t$ given $\boldsymbol{Y}(0)$, ..., $\boldsymbol{Y}(T\text{-}1)$ (the whole data). Computation of predictive, filtering, smoothing distributions are referred to as prediction, filtering and smoothing, respectively. Recent drastic innovation of computation in both speed and memory stimulated development of strong and powerful numerical methods for prediction, filtering and smoothing where the state distribution is represented numerically rather than analytically, for instance, with histograms, with Gaussian sum densities, and with the set of particles. While these new approaches have the great advantage of flexibility in modeling, exhaustive usage of memory in smoothing is a serious problem in its practical use. In particle filters, required storage would be typically O(NT), where N is the space required to store the filtering distribution at each time point and T is the length of the time series. This is an unfortunate drawback which limits applicability of the new approaches in many situations.

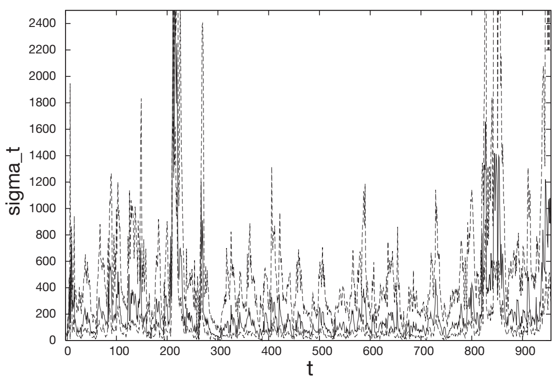


Figure1: The smoothed volatility, the ordinary method (the solid line is the median of the distribution. Thick broken, thin broken, dotted, and chain lines represent 2.3%, 15.9%, 84.1%, 97.7% points, respectively).

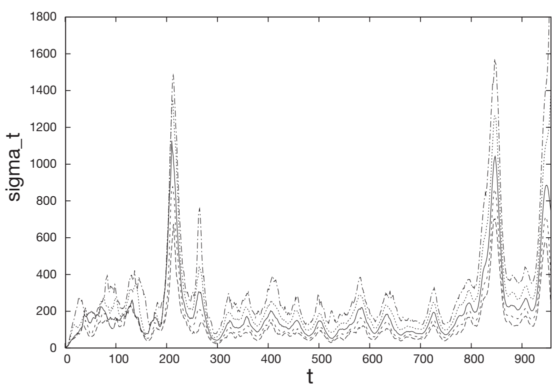


Figure2: The smoothed volatility, the proposed new method (the solid line is the median of the distribution. Thick broken, thin broken, dotted, and chain lines represent 2.3%, 15.9%, 84.1%, 97.7% points, respectively).

We have developed a new generic implementation of numerical smoothing which reduces the space complexity to O(N log T) from O(NT) by employing carefully designed recursive recomputation. The expense of the new scheme in time complexity is O(log T) times computation of filtering distributions. If one is ready to pay this cost, then it would be possible to increase drastically the number of particles in particle smoother. Indeed, we are able to compute smoothing distributions using 20 times more particles than before to obtain much higher resolution smoothing distributions. Figures 1 and 2 show smoothing distributions of volatility of Nikkei 225 index from 1987 January to 1990 March. It is seen that much higher resolution is achieved in Figure 2 where the new method is employed in computing the smoothing distribution. We are developing easy implementation of the method and parameter adjustment algorithm based on this approach. As an application, we plan to use the method to analyze motion picture (video-data) of two mice in a cage (Figure 3) to study genomes which govern social behavior of the mouse.

Takashi Tsuchiya

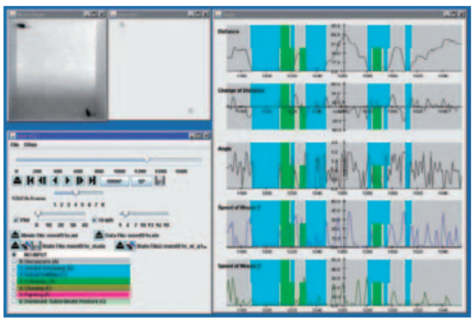


Figure3: Mice video-data analysis

Introduction to Our Research

Estimation of Quantum States from Observation

Inference on Quantum States

In these fifteen years, various types of "quantum information technology" arise, such as quantum computation, quantum cryptography. An art of the estimation of the quantum states realized by experiments is essential in any of these fields. This art is sometimes called "quantum tomography", because a quantum state is reconstructed from a set of observations each of which corresponds to the measurement of a "projection" of the quantum state. It has a close analogy with CT-scan used in hospitals.

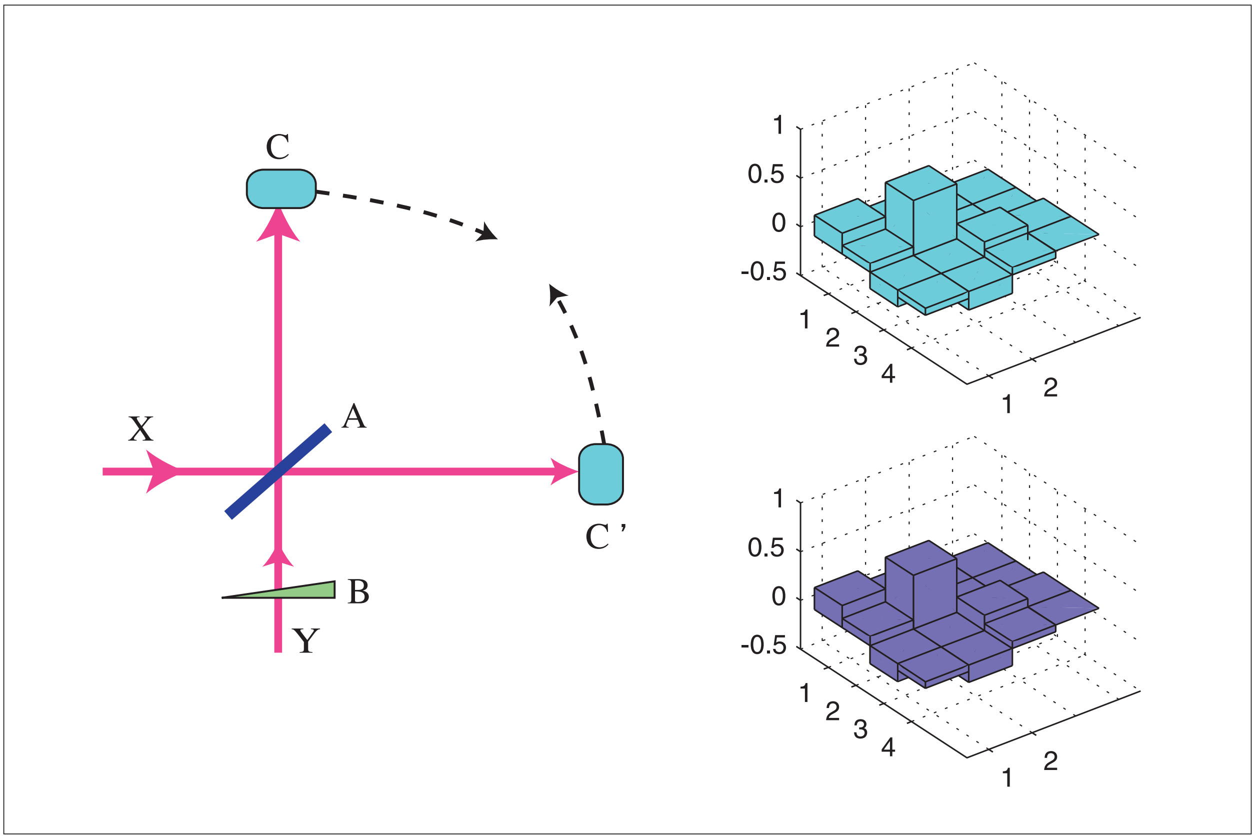
Background and Aim

The aim of the project is to develop algorithms for quantum state estimation based on modern numerical and statistical science. Since reconstruction of hidden states from noisy and incomplete observation is a central problem in modern statistical science, development of efficient methods of the estimation of quantum states with error assessment will be a great challenge to statistical scientists. A target of the project is to improve methods of the estimation of density matrices, which is a typical example of large-scale convex programming, a subject extensively studied by researchers in ISM.

Current Status

We are collaborating with Keiji Matsumoto (NII) and. Mituhiro Fukuda (Tokyo Tech) in this project. We implemented a convex programming algorithm for the maximum likelihood estimation of density matrices and successfully tested it with sets of real and artificial data. We also developed Bootstrap approaches for error estimation and showed that it provides error estimates with a moderate computational effort using the convex programming algorithm.

Yukito Iba



The left panel of the figure shows an experiment of quantum tomography. The right panels show a result of a numerical experiment with artificial data, where the upper panel (blue) shows the original density matrix and the lower panel (purple) shows the reconstructed density matrix from a set of artificial data. Accurate reconstruction is obtained in this case.

Developing Statistical Models for Bycatch in Tuna Fisheries

Developing statistical models for bycatch in tuna fisheries

Historically, when people have considered the impact of fisheries on marine resources, their concerns have largely focused on the population status of the target species. In the last decade, however, the issue of bycatch reduction has been getting more attention.

Bycatch problem in tuna fisheries

In the eastern Pacific Ocean, purse-seine nets are used to catch tunas. Tunas are detected by fishermen in three different ways: in association with floating objects (“floating object” sets), in association with herds of dolphins (“dolphin” sets), and as free-swimming schools visible at the surface. Incidental mortality of dolphins, sharks, sea turtles, and other species can occur during fishing operations for tunas. The incidental mortality of dolphins in this fishery was the first bycatch problem that attracted public attention. In the 1960’s, hundreds of thousands of dolphins are estimated to have been killed annually incidental to fishing operations in dolphin sets. With the development of dolphin-release techniques by fishermen, national legislation and international agreements establishing quotas on incidental mortalities, and the implementation of a seminar program designed to educate fishermen on methods for avoiding dolphin mortalities, incidental mortality of dolphins has declined to less than three thousand animals annually since 1998. Although dolphins are rarely killed incidental to fishing operations in other types of purse-seine sets, large amounts of bycatch of many other species can occur in these sets. We are working with Dr. Cleridy Lennert-Cody of the Inter-American Tropical Tuna Commission to develop new statistical methods to better model bycatch, and the complex relationships that exist between bycatch and catch species.

Analysis of shark bycatch counts

Annually, shark bycatch occurs in more than one third of the floating object sets. One notable characteristic of shark bycatch data is that there are many sets with zero bycatch, yet sets with large amounts of bycatch can also occur. We are developing a zero-inflated negative binomial regression model for the shark bycatch data. The zero-inflated negative binomial regression model assumes that there are two states: a “complete” state in which bycatch never happens, and an “incomplete” state in which bycatch might occur. In the incomplete state, bycatch counts are assumed to follow a negative binomial regression model. We have fit this model to shark bycatch data from floating object sets. The distribution of the predicted bycatch was found to be quite similar to that of the observed bycatch suggesting that the zero-inflated negative binomial is a reasonable model for these data. We are using this model to estimate temporal trends in shark bycatch rates, an important relative index of population status.

Analysis of species associations

As part of efforts to reduce bycatch in this fishery it is important to identify associations among catch and bycatch species. This is a difficult task because some species are caught only rarely and mostly in small amounts, whereas other species are caught much more frequently, sometimes in very large numbers. To this end, we are developing a new method for measuring association between variables in multivariate data that is based on a Tweedie distribution. The family of Tweedie distributions is a very flexible collection of statistical distributions that can accommodate many of the characteristics of fisheries data. We have applied this method to data from the purse-seine fishery for the year 2000. Preliminary results show that the indices of association obtained with this new method are related to different groupings of catch and bycatch species, and show clear large-scale spatial patterns. These indices also show relationships to environmental variables such as sea surface temperature, chlorophyll-a density and mixed layer depth. Information on relationships between indices of species association and environmental characteristics will help to guide the development of approaches for bycatch reduction.

Mihoko Minami

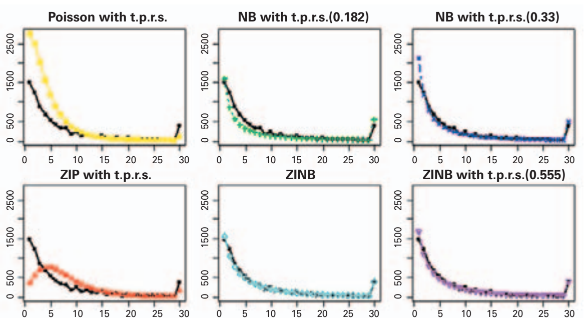


Figure 1: Observed and predicted distributions for shark bycatch counts.

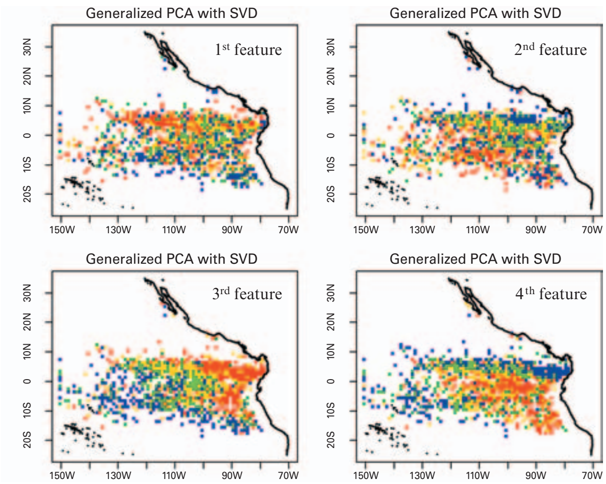


Figure 2: Spatial distributions of feature values.

Sampling Selections after the Great Merger of the Heisei Era

Survey Sampling for Social Research

In nationwide sample surveys, municipalities across the nation are stratified by regionality and population size (stratification). Sampling locations are randomly selected from each stratum and respondents are selected from the Voting Registers or the Basic Resident Registers in these locations. This procedure, stratified multistage random sampling, is popular in Japan.

However, the Great Merger of the Heisei Era (1999-2006), which was the nationwide merger of municipalities, dramatically changed Japan's administrative districts. The total number of Japanese municipalities, which was more than 3,000, decreased to less than 2,000 in March 2006 (see Figure 1). Around the same time, an increasing number of local governments began limiting access to Voting Registers or the Basic Resident Registers or complicating the order of the registers to make using registers difficult indirectly. In addition, since November 2006, owing to the amended Public Officers Election Law and Basic Resident Registration Law, we have been unable to access the registers for certain kinds of surveys. In particular, accessing the Voting Registers requires not only public interest but also the use of survey topics related to politics or elections. Therefore, the post-Great Merger survey sampling environment for social research differs from its predecessor.

Change of the Sampling Enviroment and its Influences

This study addresses the details and influences of the institutional changes that occurred after the Great Merger of the Heisei Era.

As for access to Voting Registers, the requirements of political or electoral topics leads to a high hurdle of the previous examination; however, as a result of our survey regarding access procedures for Voting Registers, we found that the actual cases of access after the examinations were mostly smooth and the registers with irregular orders (not address order) were rare.

On the other hand, as for the influences of mergers of municipality, the occurrence of transitions from rural districts to general cities and from general cities to specified cities as well as an increase in cities with a large population are evident. As the result of it, the population allocated to each stratum changed dramatically from the period 2000 to 2005. (This explanation assumes the case where we use three strata: specified cities, general cities, and rural districts.)(see Figure 2).

Originally, stratified sampling can improve the precision of the estimator by stratification based on the prior information. As is currently observable, the radical change of administrative districts, which correspond to components in strata, will influence the expected effects of stratification provided by the administrative districts as prior auxiliary information. Thus, this study examines the precisions of the stratification design influenced by the above-described change of districts.

Wataru Matsumoto

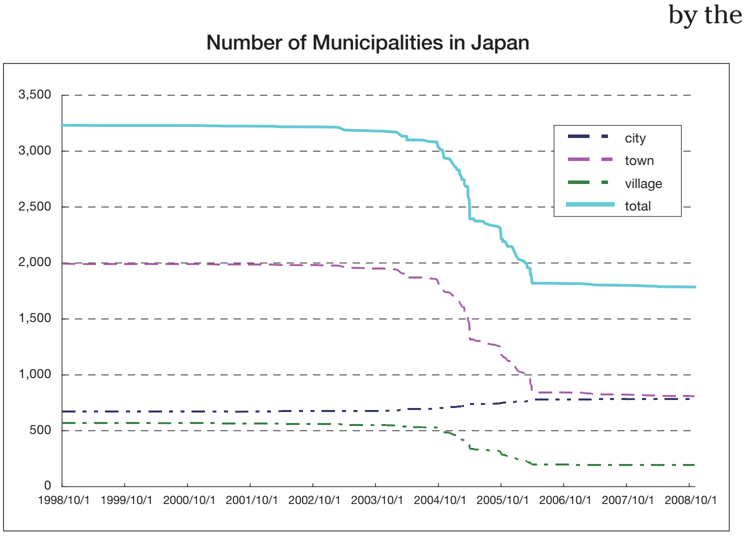


Figure 1: This figure was drawn on the data provided by Ministry of Internal affairs and Communications, "Transition Table of Municipality."

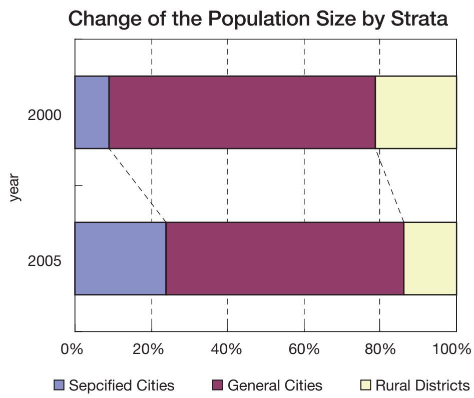


Figure 2: This figure was drawn on the data provided by Ministry of Internal Affairs and Communications, "Census (H12, H17)."

Construction of Statistical Methodologies in Quality Engineering

Outline of research

Quality engineering (Taguchi methods) is a technical methodology devised more or less independently by Genichi Taguchi over half a century for the purpose of making systematic improvements to quality. A key technique of Taguchi methods is parameter design, which is known to reduce the "variability" of characteristics at the product design and development stages, and there have been many reports of successes in manufacturing industries such as the automotive and electrical industries as part of the establishment of Quality-Japan.

But although the development of theoretical methodologies for parameter design in Taguchi methods is being actively pursued overseas as an extension of statistical quality control (SQC) or statistical science, in Japan it has only been researched to a limited extent. Therefore this study aims to devise a statistical methodology for quality engineering by clarifying the points of similarity between the design of experiments according to Fisher and Taguchi.

SQC and Taguchi methods

In SQC, based on experimental data on the quality characteristics of materials, mechanical equipment, products and the like, techniques such as the 7 QC tools, Fisher's design of experiments method and regression analysis are used to investigate the causes and analyze the important factors. On the other hand, Taguchi methods treat the causes of variation in the customer usage conditions, environmental conditions and the like as noise factors, and aim to discover the conditions that make the effects of these factors as small as possible. In parameter design, inevitable errors resulting from noise factors are normally larger than the random errors dealt with by SQC, so more effort is put into the descriptive statistics approach rather than into postulating probability distributions. However, it is possible to take a statistical approach by introducing noise factors into the model as parameter factors, and considering the probability distribution of SN ratios, thereby enabling a discussion of statistical inference.

Statistical aspects of Taguchi methods

We consider the comparison of the performance and the systems with dynamic characteristic (Fig. 1). The data obtained by comparative tests with dynamic systems in Taguchi methods is shown in Fig. 2. To judge the performance of these systems, we devised "a test of the equality of SN ratios" by considering the error distribution in the model. In this way, we were able to scientifically judge the performance comparison of the systems. In this study, we proposed a test method that is compatible with data sampled both actively and passively.

Toshihiko Kawamura

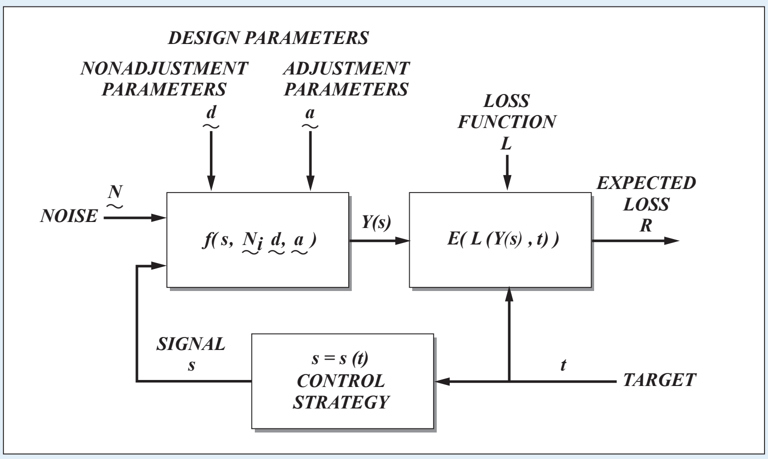


Figure 1: Systems with dynamic characteristics (Leon et al. (1987), p.256, Technometrics)

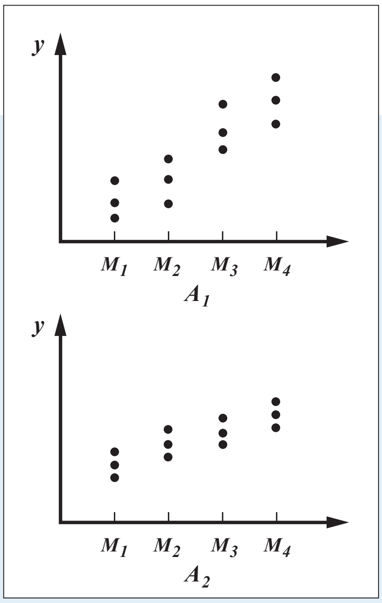


Figure 2: Data to compare the SN ratios

Research Cooperation

International Cooperation

Associated Foreign Research Institutes

Organization name	Address	Conclusion day
The Statistical Research Division of the U.S. Bureau of the Census	USA (Washington)	July 27, 1988 -
Stichting Mathematisch Centrum	The Kingdom of the Netherland (Amsterdam)	May 10, 1989 -
Statistical Research Center for Complex Systems, Seoul National University	The Republic of Korea (Seoul)	October 17, 2002 -
Institute for Statistics and Econometrics, Humboldt University of Berlin	Germany (Berlin)	December 8, 2004 -
Institute of Statistical Science, Academia Sinica	Taiwan (Taipei)	June 30, 2005 -
The Steklov Mathematical Institute	Russia (Moscow)	August 9, 2005 -
Central South University	China (Changsha)	November 18, 2005 -
Soongsil University	The Republic of Korea (Seoul)	April 27, 2006 -
Department of Statistics, University of Warwick	The United Kingdom (Coventry)	January 16, 2007 -
The Indian Statistical Institute	India (Kolkata)	October 11, 2007 -

Research Collaboration

This academic study program provides researchers from other academic institutes with access to the facilities of the Institute, and provides opportunities for researchers to conduct theoretical and applied studies on statistics.

Number of Activities

2002	2003	2004	2005	2006	2007
114	99	108	124	122	120

Fields of Research Collaboration

Research collaboration is classified by research field as follows. Applicants can use the table below to find the most appropriate type of project.

ISM Fields

Number	Fields	Number	Fields
a	Spatial and time series modeling	f	Computational statistics
b	Intelligent information processing	g	Mathematical statistics
c	Graph modeling	h	Learning and inference
d	Survey research	i	Computational mathematics
e	Multidimensional data analysis	j	Others

Major Research Fields

Number	Fields	Major Research Domains
1	Statistical mathematics	Mathematical theory of statistics, optimization, etc.
2	Information science	Algorithms, use of computer in statistics, etc.
3	Biological science	Medicine, pharmacy, epidemiology, genetics, etc.
4	Physical science	Space, planet, earth, polar region, materials, etc.
5	Engineering	Mechanics, electronics, control, chemistry, architecture, etc.
6	Human science	Philosophy, art, psychology, education, history, geography, culture, language, etc.
7	Social science	Economics, low, politics, society, management, official statistics, population, etc.
8	Others	Other research fields

Graduate School Program

Organization

The Institute of Statistical Mathematics is one of the platforms of the Graduate University for Advanced Studies (the headquarters in Hayama, Kanagawa), which was opened in October 1988 to offer graduate education. Since its opening, the Institute has created the Department of Statistical Science and, since April 1989, has accepted students for education and research in doctoral programs. In 2006, the Institute adopted a five-year system, offering either a five-year education and research program, or a three-year education and research program starting from the third year of study.

- President
  - School of Cultural and Social Studies
    - Regional Studies
    - Comparative Studies
    - Japanese Studies
    - Japanese History
    - Cyber Society and Culture
    - Japanese Literature
  - School of Physical Sciences
    - Structural Molecular Science
    - Functional Molecular Science
    - Astronomical Science
    - Fusion Science
    - Space and Astronautical Science
  - School of High Energy Accelerator Science
    - Accelerator Science
    - Materials Structure Science
    - Particle and Nuclear Physics
  - School of Multidisciplinary Sciences
    - Statistical Science
    - Polar Science
    - Informatics
  - School of Life Science
    - Genetics
    - Basic Biology
    - Physiological Sciences
  - School of Advanced Sciences
    - Biosystems Science
    - Photo Science
    - Evolutionary Studies of Biosystems

Outline of Education and Research

The course includes modeling, forecasting, inference and designing of data-gathering systems in order to extract information and knowledge from the real world based on empirical data, as well as basic, mathematical and applied education and research related to these fields. The course aims to provide the student with skills that help to contribute to solving important and connected issues and give the ability to perform original research.

Field of Education and Research	Contents
Statistical Modeling	Education and research focuses on dynamic modeling such as spatial and space-time modeling, graphical modeling of temporally and/or spatially inter-related complex phenomena, and intelligent information processing. We also provide education and research on statistical inference based on various models, methods of calculation for inference, and evaluation of models based on data.
Data Science	We provide education and research on data design, investigation and analytical methods to cope with the uncertainty and incompleteness of information, as well as on computational statistics.
Mathematical and Statistical Inference	We provide education and research on the theory of statistics and related basic mathematics, statistical learning theory to extract information from data by automated learning and inference techniques; as well as theory and applications of optimization and computation algorithms which serve as the basis for computational inference.

Features of Education and Research

- The course is the only integrated doctoral program on statistical science in Japan. It has received students from a wide variety of disciplines and has offered education and research on statistical science by professors specialized in many different fields, from theory through to practical applications.
- The Institute of Statistical Mathematics, the platform for the course, is equipped with a world-class super computer, high-speed 3D graphic computers and simulators to generate physical random numbers, as well as a variety of software, including original statistical software developed by the Institute.
- The academic publications and books on statistical and mathematical sciences produced are some of the best in the world.
- The library holds an extensive collection of books and journals, being one of the best in the world.
- In its role as an inter-university research institute, the Institute holds frequent workshops and seminars by visiting professors and researchers from both Japan and abroad. Students are free to attend and participate.
- It is possible to collaborate with researchers from other universities and institutions. It is also possible for students to develop their own projects by participating in research projects with other institutions through the Transdisciplinary Research Integration Center, Research Organization of Information and Systems.

Course Requirements and Type of Degree Granted

- Requirements to complete the course are as follows:
  Completion of at least 40 credits while a graduate student of five years, or completion of at least 10 credits while a doctorate student of three years who graduated Master's course meeting all the criteria set by the thesis committee of the Institute and successfully completing the final examination.
- On completion of the course, either a Doctorate in Statistical Science or, if the thesis deals mainly with an inter-disciplinary field related to statistical science, a Doctorate of Philosophy is awarded.
- The required number of years of study will be flexible if a student demonstrates outstanding research results.

Number of Students (As of April 1, 2008)

Doctor's course three years

Year of enrollment	2002	2003	2004	2005	2006	2007	2008
Number of students	1 ①	1 ①	4 ③	2 ②	3 (1)	6 ③	4 ②

* The figures in parentheses indicate the number of foreign students being supported at government expense.
* The figures in circles indicate those who are employed by other organizations.

Doctor's course five years

Year of enrollment	2006	2007	2008
Number of students	1	2	1

University Background of Students

National and public universities

• Hokkaido University • Tohoku University (2) • Fukushima University • University of Tsukuba (5) • Saitama University
• Ochanomizu University • Hitotsubashi University • Chiba University • The University of Tokyo (13) • Tokyo Gakugei University
• Tokyo University of Agriculture and Technology • Tokyo Institute of Technology (2) • Tokyo University of Marine Science and Technology
• Nagoya University (2) • Toyohashi University of Technology • Kyoto University (3) • Osaka University • Okayama University
• Shimane University (2) • Kyushu University (2) • Oita University • Japan Advanced Institute of Science and Technology, Hokuriku
• Osaka City University

Private universities

• Keio University (4) • Chuo University (5) • Tokyo University of Science (5) • Toyo University • Nihon University (2),
• Japan Women's University • Hosei University (7) • Waseda University (5) • Kyoto Sangyo University • Okayama University of Science
• Kurume University

Foreign universities

• Aston University • Center for Analysis and Prediction, China Seismological Bureau
• Chinese Academy of Sciences, Institute of Applied Mathematics • Jahangirnagar University (2) • Northeastern University, China
• Ohio University • Stanford University • The Hong Kong University of Science and Technology • Universidade Eatadual of Campians
• University of Colorado at Boulder (2) • University of Dhaka (2) • University of Hawaii • University of Malaya • University of Rahshahi
• University of Science and Technology of China

Degrees Awarded

Year	Doctor of Philosophy
1996	3 [1]
1997	1
1998	4 [1]
1999	6
2000	5
2001	5

Year	Doctor of Philosophy
2002	4
2003	8 [1]
2004	4
2005	4
2006	8 [1]
2007	7 [1]

* [ ] Ph.D. on the basis of the dissertation only (included in the total)

Current Position of Alumni (As of April 1, 2008)

National and public universities, and public organizations

• Obihiro University of Agriculture and Veterinary Medicine, Professor • University of Tsukuba, Professor (2)
• University of Hyogo, Professor • Saitama University, Associate Professor • The University of Electro-Communications, Associate Professor
• The University of Tokyo, Associate Professor • Kyushu University, Associate Professor • Kyushu Institute of Technology, Associate Professor
• The Institute of Statistical Mathematics, Associate Professor • University of Tsukuba, Lecturer • Hokkaido University, Assistant
• Chiba University, Assistant • Tokyo Institute of Technology, Assistant • Hiroshima University, Assistant
• Kyushu University, Assistant Professor • University of the Ryukyus, Assistant • The Institute of Statistical Mathematics, Assistant Professor (5)
• The University of Tokyo, Project Researcher • Tokyo Institute of Technology, Research Fellow
• Nara Institute of Science and Technology, Research Fellow • The Institute of Statistical Mathematics, Project Researcher (5)
• Bank of Japan, Project Post • Japan Broadcasting Corporation • Financial Services Agency, Financial Research and Training Center, Researcher
• The Institute of Statistical Mathematics, JST CREST Research Fellow • JST Basic Research Programs Doctoral Research Fellow
• Railway Technical Research Institute, Senior Researcher • Statistical Information Institute for Consulting and Analysis
• Government Pension Investment Fund • Public School, Teacher • The Institute of Statistical Mathematics, Research Fellow (2)

Private universities

• Sapporo Gakuin University, Professor • Meiji University, Professor • Doshisha University, Professor
• Tokyo Health Care University, Associate Professor • Nihon University, Associate Professor • Tokyo University of Information Science, Lecturer
• Josai University, Lecturer • Sapporo Gakuin University, Full-Time Lecturer • Tokyo Women's Medical University, Postdoctoral Fellow

Foreign universities

• Asia-Pacific Center for Security Studies Department, Associate Professor • Central South University, Professor
• Hong Kong Baptist University, Lecturer • Jahangirnagar University, Professor • Jahangirnagar University, Associate Professor (2)
• Massey University, Research Fellow • The University of Warwick, Research Fellow • University of Otaga, Research Fellow
• University of Rajshahi, Associate Professor • University of South Carolina, Research Fellow • Victoria University, Senior Lecturer

Private companies, etc.

• Hitachi, Ltd. Central Research Laboratory, Research Fellow • NTT Communication Science Laboratories, Research Fellow
• Seiwa Kikaku • Toyota Motor Corporation, Higashi-Fuji Technical Center, Research Fellow • NLI Research Institute, Chief Researcher
• Sankyo Co., LTD. • Mizuho Trust and Banking, Senior Researcher (2) • JP Morgan Trust Bank Limited, Vice President (Hosei University, Part-Time Lecturer)
• ATR Computational Neuroscience Laboratories, Research Fellow • Schlumberger Limited • Macquarie Securities, Japan, Quantitative Analyst
• Non-Life Insurance Rating Organization of Japan, Staff Member • Open Technologies Corporation

Outreach Activities

Tutorial Courses

History

The statistical education program started in 1944, the year that the Institute of Statistical Mathematics was founded, as an education program at the Numerical Computation Training Center of the Science Research Technical Assistant Training Center of the Ministry of Education, located in the Institute. In 1947, the affiliated Statistical Technician Training Center was opened as an educational organization for statistical technicians and instructors, in order to improve the levels of staff training within the statistical organizations of the government and to supply statisticians.

As social needs have changed, the purpose of the education program has gradually shifted away from the initial aim of supplying well qualified statistical technicians for the government, towards statistical education for working people. Tutorial courses were therefore initiated. Later, statistical methods became more widely used across a broad range of fields. Consequently, more extensive and sophisticated statistical education was required. To meet this demand, tutorial courses were expanded to cover a wider range of statistical topics. From 1965 to 1985, six to eight courses were held annually, not only in Tokyo, but also in provincial cities such as Osaka, Okayama and Fukuoka.

In 1985, the Institute was reorganized as a member of the Inter-University Research Institute Corporation, and the affiliated Statistical Technician Training Center was abolished. However, in response to consistent public demand for tutorial courses, the Center for Engineering and Technical Support, together with other departments, ran three to four courses annually. In 2005, the number of courses rose to 13. In 2007, the number of courses was 13, two of which were held at night.

Courses

The total number of courses held from 1969 to March, 2008 was 233, with a total of 17,750 participants. These courses covered a wide range of fields from basic to applied statistics. The following table lists the courses held in the past 5 years:

Year	Category	Title	Month	Number of participants
2005	Basic course	Introduction to Statistics	September	85
	Advanced course	A Junction of Informatics – Chordal Graph and its Applications –	September	13
	Standard course	Analysis of Qualitative Data by Quantification Methods	October	50
	Basic course	Nonlinear Time Series Analysis of Financial Data	October	52
	Basic course	Introductory Data Analysis with R	November	98
	Advanced course	Data Processing and LSI Design for Information and Telecommunications with the Latest Technologies	November - December	5
	Basic course	Introduction to Sampling Methods and Sampling Surveys	December	75
	Basic course	Introduction to Reliability Theory and Survival Data Analysis with R	December	39
2006	Standard course	Theory and Practice Inferring Molecular Phylogenies	January	72
	Basic course	Introduction to Probabilistic Evaluation of Risk	January	45
	Standard course	Non-Poisson Regression Models for Count Data	February	39
	Advanced course	Packing and Random Packing	February	7
	Basic course	Introduction to Time Series Analysis	March	72
	Standard course	A Course on Time Series Analysis for Economics and Finance	June	40
	Standard course	Advances in Kernel Methods: SVM, Nonlinear Data Analysis, and Structured Data	July	73
	Basic course	Basic Medical Statistics Using R	July	20
	Basic course	Introduction to Statistics	July	69
	Standard course	Lectures on Information Theory and Mobile Telecommunication Technologies – Systems and Hardwares for Large-Scale Data Processing –	August-September	13
	Basic course	International Standardization of Statistical Methods – Precision and Trueness of Measurement Methods and Results – Capability of Detection	September	22
	Advanced course	A New Trend of Adaptive and Learning Control Theory	September	14
	Basic course	A Game Theoretic Approach to Mathematical Finance	November	21
	Basic course	Introduction to Quantitative Methods for Social Sciences	November - January	43
	Standard course	Statistical Pattern Recognition	November	65

Year	Category	Title	Month	Number of participants
	Basic course	Introduction to Statistical Data Analysis	November - March	13
	Standard course	Statistical Mathematics of Rock-Scissors-Paper Game	November - December	7
	Standard course	An Introduction to Statistical Analysis Based on the Theory of Martingales	December	38
2007	Basic course	Introduction to Risk Analysis with R – Application of Tree-based and Nonparametric Modelling –	January	49
	Basic course	Introduction to Survey Data Analysis Using R	February	40
	Basic course	Introduction to Quantitative Methods for Social Sciences	May-July	54
	Basic course	Introduction to Sampling Methods and Sample Surveys	June	50
	Basic course	Elementary Course on Time Series Analysis	July	61
	Basic course	Introduction to Statistics	July	70
	Standard course	Statistics of Extremes	September	61
	Basic course	Introduction to Statistical Data Analysis focused on Multivariate Analysis	October	47
	Basic course	Quantification Methods for Qualitative Data	October-December	46
	Basic course	Recent Topics on International Standardization of Statistical Methods - Standardization of Sampling Inspection and Statistical Process Management -	November	11
	Basic course	Statistical Quality Control	November	16
	Standard course	Taguchi Methods for Robust Design	November	28
	Standard course	Geometrical Structures Underlying Information Quantities: Mathematics of Kullback-Leibler Information	December	52
2008	Basic course	Statistical Causal Analysis by Structural Equation Modelling	February	58
	Standard course	Regression Models for Count Data and Their Extension	February	63

The schedule of tutorial courses can be found on the website of the Institute of Statistical Mathematics.
http://www.ism.ac.jp/

ISM Seminar

The Institute holds a one-hour seminar every Wednesday to showcase the latest studies of our academic staff and guest researchers. The seminar is open to the public, starting at 1:30 PM. For details, please visit the website of the Institute.
http://www.ism.ac.jp/

Open Lecture

We hold an open lecture during Education and Culture Week every year (November 1 through 7), to introduce the Institute's activities and to promote statistical science. We invite lecturers to speak on a timely topic relating to statistical science. The topics in 2006 and 2007 were "Discovery and analysis of stochastic models" and "Contributions of statistical science to health sciences". The lecture is open to the general public. For further information, please visit the website of the Institute of Statistical Mathematics.
http://www.ism.ac.jp/

Statistical Consultation Service

The Institute provides a consultation service on statistical science for the general public and researchers as a means of actively sharing the benefits of our research with society. This service operates throughout the year. Please contact the Center for Engineering and Technical Support. The consultation service covers a variety of topic ranging from the basics of statistics to more specialized issues. Half of the consultants are from the private sector, and the rest are staff from public organizations, university teachers, and students. Each teacher at the Institute directly takes charge of about 20 specialized cases annually, about 40% of which are publicized at academic conferences, as being of benefit to society.

Annual Research Report Meeting for 2007

The Annual Research Report Meeting of the Institute was held on March 13 and 14, 2008. This annual meeting is an opportunity for the Institute's teachers and visiting teachers to present the results of their studies from the previous year. The meeting has been held every year since the foundation of the Institute in 1944. In the early years, as there were few staff members and participants, they were able to have enthusiastic discussions all day long. However, as the number of staff members has increased, the meeting is now held over two days and each presenter is given 15 minutes to report on the results of his/her own research. A wide variety of topics is covered, ranging from statistical contributions to issues facing today’s society to more fundamental studies. This year’s meeting started with an opening address by Junji Nakano, and greetings from Genshiro Kitagawa. A total of 46 research education staff and 6 visiting teachers then gave their presentations. In addition, a poster session was held by 18 research fellows and this year’s research education staff. From 2004, we have begun to compile and distribute proceedings in advance. We also host visitors from other organizations. The program for this meeting is available from the website.

Finance and Buildings

Administration Subsidy and Others (2007)

Type	Personnel expenses	Non-personnel expenses	Total
Expenditure	923,527	748,737	1,672,264

Unit: ¥1,000

Accepted External Funds (2007)

Type	Subcontracted research	Joint research	Contribution for scholarship	Total
Items	4	6	3	13
Income	61,479	7,537	4,700	73,716

Unit: ¥1,000

Grant-in-Aid for Scientific Research (2007)

Research Category	Items	Amount Granted
Grant-in-Aid for Scientific Research on Priority Areas	1	2,700
Grant-in-Aid for Scientific Research (A)	4	46,280
Grant-in-Aid for Scientific Research (B)	4	20,800
Grant-in-Aid for Scientific Research (C)	12	15,080
Grant-in-Aid for Exploratory Research	2	2,000
Grant-in-Aid for Young Scientists (A)	1	7,800
Grant-in-Aid for Young Scientists (B)	9	8,700
Grant-in-Aid for JSPS Fellows	3	3,300
Total	36	106,660

Unit: ¥1,000

Site and Buildings (As of April 1, 2007)

Site Area	5,033m^2
Area for Buildings (total)	6,305m^2

Name of Building	Floors	Total Area
Office	R3	4,855m^2
Information Statistics Research Laboratory	R3	1,024m^2
Computation Laboratory	R2	368m^2
Gym, etc.	S1	58m^2

Facilities and Equipment

Computation Resources (As of April 1, 2008)

Since January 2004, "Supercomputer System for Statistical Science" has been in operation and has analyzed a large volume of statistical data. The main components of this system comprise a SGI Altix3700 Super Cluster (a parallel computer system with 256 Itanium2 processors and about 2 TB main memory), a NEC SX-6 (a vector-type computer system with 12 vector processors and 128 GB main memory), and a HITACHI SR11000 (a parallel computer sub-system with 64 Power4+ processors and 128 GB main memory). In January 2006, "System for Computational Statistics" was renewed. The main components of this system consist of a HP XC4000 Cluster system with 256 Opteron processors for computing node and 640 GB main memory, a SGI Prism visualization system with 16 Itanium2 processors and 32 GB main memory, and a large display system (Multi Opt View).

In December 1998, an Ethernet network using 1000Base-SX as a main trunk and 100Base-TX as branches was laid out as a Local Area Network. Workstations, personal computers in researchers' offices, "Supercomputer System for Statistical Science", and "System for Computational Statistics" were all connected to the network. This Local Area Network enables distributed processing and computation resources and statistical data to be used effectively. The development of programs, which run on "Supercomputer System for Statistical Science" and on "System for Computational Statistics" by operating from workstations and personal computers in researchers' offices, is also underway. To encourage joint research with researchers both in Japan and abroad, as well as the exchange of e-mails, the network is connected to the internet through SINET. The connection speed of 1.5 Mbps during 1999 has risen to 100 Mbps since July 2002. Since April 2007, the network has also been connected to the SINET3 with 2.4 Gbps bandwidth. Some machines are able to communicate at a rate of 1 Gbps. Through terminal servers, the network is also accessible from a public line. In addition, comprehensive network security measures have been implemented such as the adoption of anti-virus software as well as a network monitoring system.

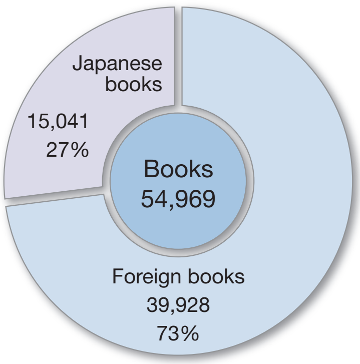

SGI Altix3700 Super Cluster, a parallel computer

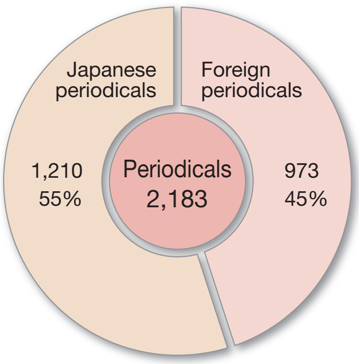
Library and Materials (As of April 1, 2008)

We have a large number of major Japanese/foreign journals covering a wide variety of fields including statistics, mathematics, computer science and informatics. In addition, we also have a large library consisting of books on humanities, social science, biology, medical science, science and engineering.

Besides contributed to Japanese and foreign publications, we also have a collection of journals that we publish ourselves: Annals of the Institute of Statistical Mathematics (English; Springer), Proceedings of the Institute of Statistical Mathematics (Japanese), Research Report (Statistical Researches mainly related to the Japanese National Character), Computer Science Monographs, Cooperative Research Reports (for collaborative research projects), Research Memorandum, ISM Reports on Statistical Computing, and ISM Report on Research and Education.

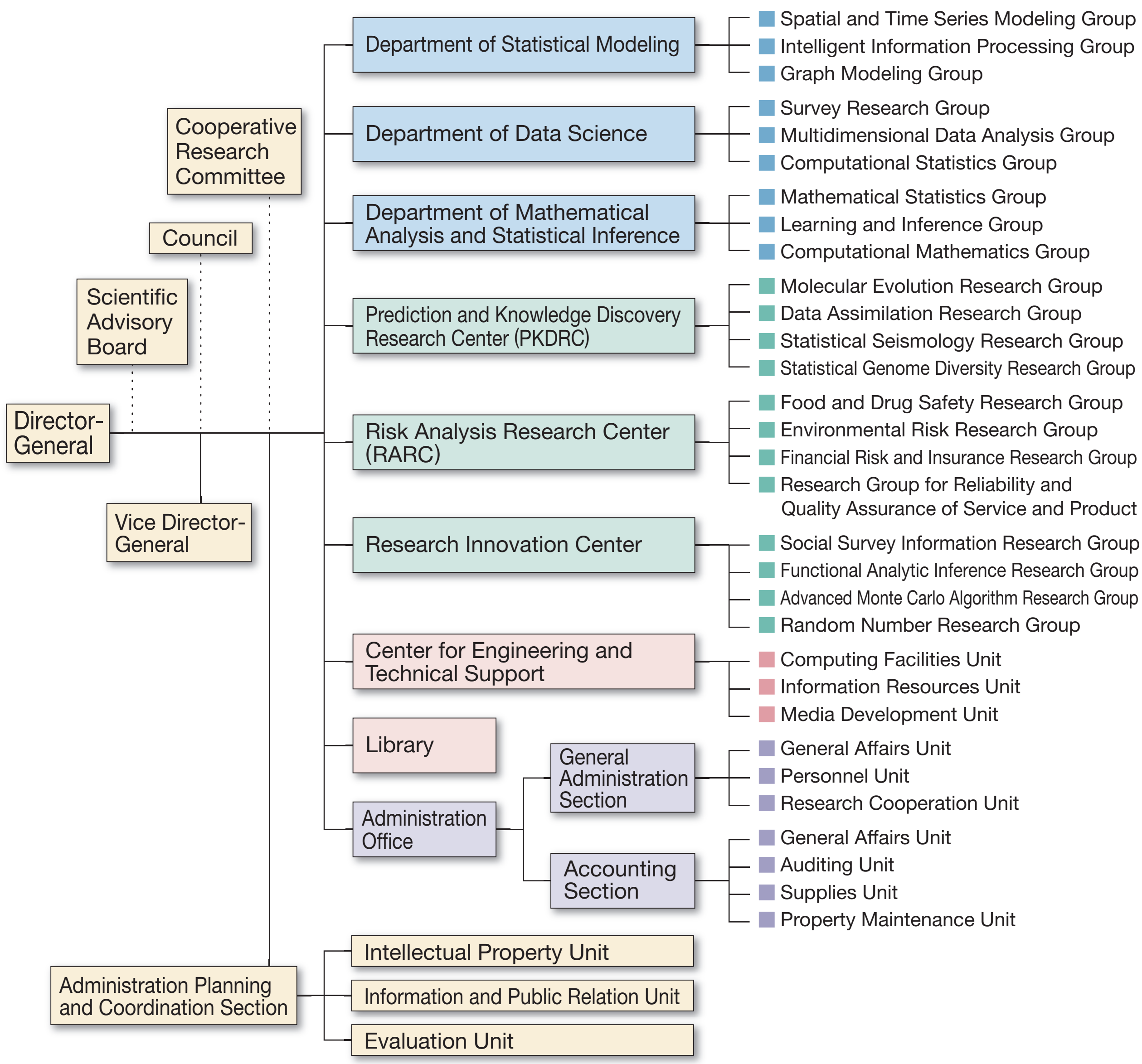
All materials are properly catalogued and can be searched from the web in order to meet the needs of researchers working in a wide of fields. We also accept photocopy requests.





Organization

Organization Diagram (As of April 1, 2008)



Number of Staff (As of July 1, 2008)

Type	Director-General	Professor	Associate Professor	Assistant Professor	Administrative Staff	Technical Staff	Total
Director-General	1						1
Department of Statistical Modeling		5	5	6			16
Department of Data Science		7	7	5			19
Department of Mathematical Analysis and Statistical Inference		5	5	4			14
Center for Engineering and Technical Support						9 (1)	9 (1)
Administration Office					14	2	16
Total	1	17	17	15	14	11 (1)	75 (1)

( ) Staff of reemployment

Staff (As of July 1, 2008)

Director-General	Genshiro KITAGAWA				
Vice Director-General	Masaharu TANEMURA	Vice Director-General	Yoshiyasu TAMURA	Vice Director-General	Tomoyuki HIGUCHI

Department of Statistical Modeling

Director Makio ISHIGURO

Spatial and Time Series Modeling Group

Prof.	Masaharu TANEMURA	Assoc. Prof.	Yoshinori KAWASAKI	Assist. Prof.	Ryo YOSHIDA
Prof.	Yosihiko OGATA	Assist. Prof.	Kenichiro SHIMATANI	Assist. Prof.	Jiancang ZHUANG
Prof.	Tomoyuki HIGUCHI	Assist. Prof.	Genta UENO		

Intelligent Information Processing Group

Prof.	Makio ISHIGURO	Assoc. Prof.	Yukito IBA	Visiting Assoc. Prof.	Masataka GOTO
Prof.	Tomoko MATSUI	Assoc. Prof.	Yumi TAKIZAWA	Assist. Prof.	Hiroshi SOMEYA
Visiting Prof.	Jean-Philippe VERT	Assoc. Prof.	Kenji FUKUMIZU		

Graph Modeling Group

Assoc. Prof.	Jun ADACHI	Assist. Prof.	Ying CAO

Department of Data Science

Director Takashi NAKAMURA

Survey Research Group

Prof.	Takashi NAKAMURA	Visiting Prof.	Nicolaos Emmanuel SYNODINOS	Visiting Assoc. Prof.	Hajime IHARA
Prof.	Ryozo YOSHINO	Assoc. Prof.	Tadahiko MAEDA	Visiting Assoc. Prof.	Takahiro HOSHINO
Visiting Prof.	Ikuo NASU	Assoc. Prof.	Takahiro TSUCHIYA	Assist. Prof.	Wataru MATSUMOTO

Multidimensional Data Analysis Group

Prof.	Toshiharu FUJITA	Visiting Prof.	Kunio SHIMIZU	Assist. Prof.	Sumie UEDA
Prof.	Hiroe TSUBAKI	Assoc. Prof.	Satoshi YAMASHITA	Assist. Prof.	Toshio OHNISHI
Prof.	Nobuhisa KASHIWAGI	Assoc. Prof.	Shigeyuki MATSUI	Assist. Prof.	Toshihiko KAWAMURA

Computational Statistics Group

Prof.	Yoshiyasu TAMURA	Visiting Prof.	Yuichi MORI	Assoc. Prof.	Seisho SATO
Prof.	Junji NAKANO	Visiting Prof.	Makoto MATSUMOTO	Visiting Assoc. Prof.	Takeshi KOSHIBA
Visiting Prof.	Yoshinari FUKUI	Assoc. Prof.	Naomasa MARUYAMA	Assist. Prof.	Nobuo SHIMIZU
Visiting Prof.	Michiko WATANABE	Assoc. Prof.	Koji KANEFUJI		

Department of Mathematical Analysis and Statistical Inference

Director Shinto EGUCHI

Mathematical Statistics Group

Prof.	Satoshi KURIKI	Assist. Prof.	Takaaki SHIMURA
Assoc. Prof.	Yoichi NISHIYAMA	Assist. Prof.	Kei KOBAYASHI

Learning and Inference Group

Prof.	Shinto EGUCHI	Assoc. Prof.	Shiro IKEDA	Assist. Prof.	Tadayoshi FUSHIKI
Assoc. Prof.	Mihoko MINAMI	Assoc. Prof.	Hironori FUJISAWA	Assist. Prof.	Masayuki HENMI

Computational Mathematics Group

Prof.	Takashi TSUCHIYA	Prof.	Atsushi YOSHIMOTO
Prof.	Yoshihiko MIYASATO	Assoc. Prof.	Satoshi ITO

Prediction and Knowledge Discovery Research Center (PKDRC)

Director Shinto EGUCHI

Molecular Evolution Research Group

Assoc. Prof.	Jun ADACHI	Adjunct Prof.	Masami HASEGAWA		
Assist. Prof.	Ying CAO	Research Fellow upon JSPS Program	Takeshi SASAKI		

Data Assimilation Research Group

Prof.	Tomoyuki HIGUCHI	Project Researcher	Kazuyuki NAKAMURA	Researcher, JST CREST	Tsukasa ISHIGAKI
Assist. Prof.	Genta UENO	Researcher, JST CREST	Shinya NAKANO		
Assist. Prof.	Ryo YOSHIDA	Researcher, JST CREST	Daisuke INAZU		

Statistical Seismology Research Group

Prof.	Yosihiko OGATA	Visiting Prof.	David VERE-JONES	Project Researcher	Ushio TANAKA
Visiting Prof.	Shinji TOHDA	Assist. Prof.	Jiancang ZHUANG		

Statistical Genome Diversity Research Group

Prof.	Shinto EGUCHI	Assoc. Prof.	Hironori FUJISAWA	Project Researcher	Shogo KATO
Prof.	Satoshi KURIKI	Visiting Assoc. Prof.	Kanta NAITO	Project Researcher	Takayuki SAKAGUCHI
Assoc. Prof.	Mihoko MINAMI	Assist. Prof.	Tadayoshi FUSHIKI		
Assoc. Prof.	Shiro IKEDA	Project Researcher	Md Nurul Haque Mollah		

Risk Analysis Research Center (RARC)

Director Hiroe TSUBAKI
Vice Director / Coordinator Toshiharu FUJITA
Coordinator Koji KANEFUJI
Coordinator Yoshinori KAWASAKI
Coordinator Toshihiko KAWAMURA

Food and Drug Safety Research Group

Prof.	Toshiharu FUJITA	Assoc. Prof.	Shigeyuki MATSUI	Assist. Prof.	Masayuki HENMI
Prof.	Hiroe TSUBAKI	Visiting Assoc. Prof.	Yoshimitsu HIEJIMA	Project Researcher	Yosuke FUJII
Visiting Prof.	Manabu IWASAKI	Visiting Assoc. Prof.	Toshimitsu HAMASAKI	Project Researcher	Kazuhiro SHIBAYAMA
Visiting Prof.	Kunihiko HAYASHI	Visiting Assoc. Prof.	Satoshi AOKI		
Visiting Prof.	Toshiya SATO	Assist. Prof.	Takaaki SHIMURA		

Environmental Risk Research Group

Prof.	Nobuhisa KASHIWAGI	Visiting Prof.	Hideshige TAKADA	Project Researcher	Masayuki KAGEYAMA
Prof.	Atsushi YOSHIMOTO	Assoc. Prof.	Koji KANEFUJI	Project Researcher	Takayuki FUJII
Visiting Prof.	Yukio MATSUMOTO	Visiting Assoc. Prof.	Hirokazu TAKANASHI		
Visiting Prof.	Kazuo YAMAMOTO	Visiting Assoc. Prof.	Tomohiro TASAKI		
Visiting Prof.	Yoshiro ONO	Project Researcher	Mitsuhiro TOMOSADA		

Financial Risk and Insurance Research Group

Visiting Prof.	Michimori INORI	Assoc. Prof.	Seisho SATO	Project Researcher	Masao UEKI
Visiting Prof.	Naoto KUNITOMO	Assoc. Prof.	Yoshinori KAWASAKI	Project Researcher	Michiko MIYAMOTO
Visiting Prof.	Hiroshi TSUDA	Visiting Assoc. Prof.	Toshinao YOSHIBA		
Assoc. Prof.	Satoshi YAMASHITA	Project Researcher	Masakazu ANDO		

Research Group for Reliability and Quality Assurance of Service and Product

Prof.	Hiroe TSUBAKI	Visiting Prof.	Kosei IWASE	Visiting Prof.	Shusaku TSUMOTO
Visiting Prof.	Kakuro AMASAKA	Visiting Prof.	Kazuo TATEBAYASHI	Assist. Prof.	Toshihiko KAWAMURA

Research Innovation Center

Director Kenji FUKUMIZU

Social Survey Information Research Group

Prof. Takashi NAKAMURA	Assoc. Prof. Tadahiko MAEDA	Assist. Prof. Wataru MATSUMOTO
Prof. Ryozo YOSHINO	Assoc. Prof. Takahiro TSUCHIYA	

Functional Analytic Inference Research Group

Assoc. Prof. Kenji FUKUMIZU	Assist. Prof. Kei KOBAYASHI

Advanced Monte Carlo Algorithm Research Group

Assoc. Prof. Yukito IBA

Random Number Research Group

Prof. Yoshiyasu TAMURA	Visiting Prof. Makoto MATSUMOTO	Assist. Prof. Sumie UEDA
Visiting Prof. Yoshinari FUKUI	Visiting Assoc. Prof. Takeshi KOSHIBA	

Project Researchers

Project Professor Atsushi FUKASAWA	Project Researcher Masashi KONOSHIMA	Project Researcher Xiaoling DOU
Project Lecturer Ajay JASRA	Project Researcher Hai Yen SIEW	

Center for Engineering and Technical Support

Director Junji NAKANO
Vice Director Satoshi YAMASHITA
Adjunct Prof. Yasumasa BABA
Deputy Manager Yuriko WATANABE

Head of Computing Facilitise Unit	Kazuhiro NAKAMURA	Head of Media Development Unit (Concurrent)	Yuriko WATANABE
Head of Information Resources Unit	Saeko TANAKA		

Library

Head Junji NAKANO

Administration Planning and Coordination Section

Head of Intellectual Property Unit	Masaharu TANEMURA	Head of Evaluation Unit	Yoshiyasu TAMURA
Head of Information and Public Relations Unit	Tomoyuki HIGUCHI		

Administration Office

General Manager Toshifumi HAGIWARA

General Administration Section

Manager of General Administration Section	Yoshifumi KUROKAWA	Chief of General Affairs Unit (Concurrent)	Fumio SUTO
Deputy Manager of General Administration Section	Toshiyuki INADA	Chief of Personnel Unit	Masami SAKAO
Specialist (General Affairs)	Fumio SUTO	Chief of Research Cooperation Unit (Concurrent)	Tohru NAKAMURA
Specialist (Research Cooperation)	Tohru NAKAMURA		

Accounting Section

Manager of Accounting Section	Toyokichi KITAHARA	Chief of Auditing Unit (Concurrent)	Minoru HAGIWARA
Deputy Manager of Accounting Section	Tetsuya KUMAZAWA	Chief of Supplies Unit	Tatsuya TAKAGI
Chief of General Affairs Unit	Minoru HAGIWARA	Chief of Property Maintenance Unit	Hirotomo SHIMIZU

Council of The Institute of Statistical Mathematics (As of April 1, 2008)

Masanori IYE	Professor, Optical and Infrared Astronomy Division, National Astronomical Observatory of Japan, National Institutes of Natural Sciences
Yutaka KANO	Professor, Graduate School of Engineering Science, Osaka University
Toshinari KAMAKURA	Professor, Faculty of Science and Engineering, Chuo University
Naoto KUNITOMO	Professor, Graduate School of Economics, the University of Tokyo
Koji KURIHARA	Professor, Graduate School of Environmental Science, Okayama University
Yoshiharu SATO	Professor, Graduate School of Information Science and Technology, Hokkaido University
Kunio SHIMIZU	Professor, Faculty of Science and Technology, Keio University
Kazuo SEIYAMA	Professor, Graduate School of Humanities and Sociology, the University of Tokyo
Makoto TAIJI	Group Director, Computational Systems Biology Research Group, Advanced Computational Sciences Department, Advanced Science Institute, RIKEN
Takashi WASHIO	Professor, the Institute of Scientific and Industrial Research, Osaka University
Masaharu TANEMURA	Professor (Vice Director-general, ISM (General Affairs))
Yoshiyasu TAMURA	Professor (Vice Director-general, ISM (Assessment))
Tomoyuki HIGUCHI	Professor (Vice Director-general, ISM (Research Planning))
Makio ISHIGURO	Professor (Director of Department of Statistical Modeling, ISM)
Takashi NAKAMURA	Professor (Director of Department of Data Science, ISM)
Shinto EGUCHI	Professor (Director of Department of Mathematical Analysis and Statistical Inference, ISM)
Junji NAKANO	Professor (Director of Center for Engineering and Technical Support, ISM)
Hiroe TSUBAKI	Professor (Director of Risk Analysis Research Center, ISM)
Yosihiko OGATA	Professor (Department of Statistical Modeling, ISM)
Satoshi KURIKI	Professor (Department of Mathematical Analysis and Statistical Inference, ISM)
Takashi TSUCHIYA	Professor (Department of Mathematical Analysis and Statistical Inference, ISM)

Cooperative Research Committee (As of April 1, 2008)

Manabu IWASAKI	Professor, Faculty of Science and Technology, Seikei University
Hidetoshi KONNO	Professor, Graduate School of System and Information Engineering, University of Tsukuba
Masahiro MIZUTA	Professor, Information Initiative Center, Hokkaido University
Michiko WATANABE	Professor, Faculty of Economics, Toyo University
Makio ISHIGURO	Professor (Department of Statistical Modeling, ISM)
Takashi NAKAMURA	Professor (Department of Data Science, ISM)
Shinto EGUCHI	Professor (Department of Mathematical Analysis and Statistical Inference, ISM)
Nobuhisa KASIWAGI	Professor (Department of Data Science, ISM)

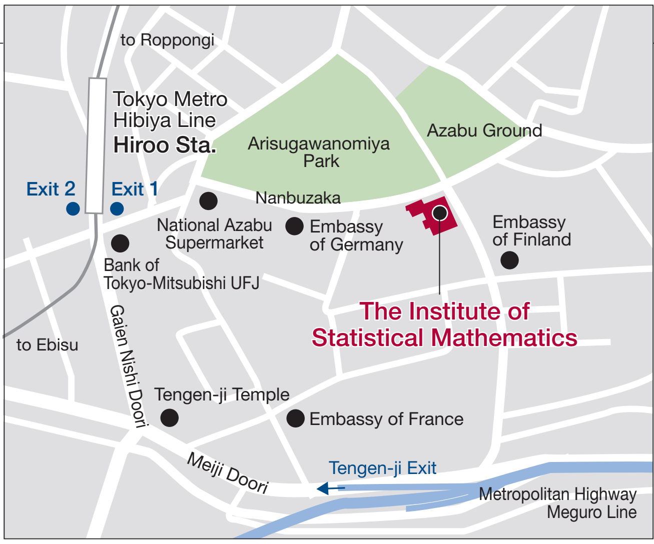
Professor Emeritus (As of April 1, 2008)

Kameo MATSUSHITA	Giichiro SUZUKI	Kunio TANABE	Takemi YANAGIMOTO
Shigeki NISHIHIRA	Ryoichi SHIMIZU	Tadashi MATSUNAWA	Yoshiaki ITOH
Tatsuzo SUZUKI	Noboru OHSUMI	Masami HASEGAWA	Yasumasa BABA
Hirotsugu AKAIKE	Masakatsu MURAKAMI	Yoshiyuki SAKAMOTO	Katsuomi HIRANO

History

June, 1944	Based on a proposal submitted at an academic study conference in December 1943, the organization was founded as an institute under the direct control of the Ministry of Education. This proposal aimed to provide supervision for studies looking into the mathematical principles of probability and their application, and was also intended to facilitate, unify and promote the publication of research results.
April, 1947	The affiliated statistical specialists' school was opened.
May, 1947	The Institute was divided into the 1st Research Dept. (fundamental theories), the 2nd Research Dept. (statistical theories for the natural sciences), and the 3rd Research Dept. (statistical theories for the social sciences).
June, 1949	The Institute was placed under the control of the Ministry of Education because of the enforcement of the Ministry of Education Establishment Law.
September, 1955	Reorganized into the 1st Research Dept. (fundamental theories), the 2nd Research Dept. (natural and social science theories), and the 3rd Research Dept. (operations, research, statistical analysis theories). The laboratory system, comprising 9 laboratories and the research guidance promotion room, was adopted.
October, 1969	A new office building was constructed.
April, 1971	The 4th Research Dept. (informatics theories) was instituted.
April, 1973	The 5th Research Dept. (prediction and control theories) was instituted.
October, 1975	The 6th Research Dept. (statistical theories of human behavior) was instituted.
November, 1979	The Information Research Building was constructed.
April, 1985	Reorganized as an Inter-University Research Institute owing to the revision of the Order for the Enforcement of the National school Establishment Law. The revised law required that the Institute would, as an National Inter-University Research Institute, 1) conduct studies on statistical mathematics and its application, 2) provide opportunities for university teachers or other researchers majoring in this field to utilize the facility, and 3) contribute to the development of academic studies in universities. At the same time, the 6 research departments were reorganized into 4 research departments (Fundamental Statistical Theory, Statistical Methodology, Prediction and Control, and Interdisciplinary Statistics). The Statistical Data Analysis Center and the Statistical Education and Information Center were instituted, and the affiliated Statistical Technician Training Center was abolished.
October, 1988	The Dept. of Statistical Science was instituted in the School of Mathematical and Physical Science, part of the Graduate University for Advanced Studies (SOKENDAI).
June, 1999	The Institute was reorganized as an Inter-University Research Institute based on the National School Establishment Law.
April, 1993	The Planning Coordination Chief System was instituted.
April, 1997	The affiliated Statistical Data Analysis Center was reorganized into the Center for Development of Statistical Computing, and the Statistical Education and Information Center was reorganized into the Center for Information on Statistical Sciences.
September, 2003	The Prediction and Knowledge Discovery Research Center was instituted in the affiliated facility.
April, 2004	The Institute was reorganized into the Institute of Statistical Mathematics, part of the Research Organization of Information and Systems of the Inter-University Research Institute based on the National University Corporation Law. The Planning Coordination Chief System was abolished and the position of Vice Director-General was instituted instead. The Dept. of Statistical Science in the School of Mathematical and Physical Science, SOKENDAI, was reorganized. In addition, the Dept. of Statistical Science and the School of Multidisciplinary Sciences were instituted.
April, 2005	The research organization was reorganized into three research departments (the Department of Statistical Modeling, the Department of Data Science, and the Department of Mathematical Analysis and Statistical Inference). The affiliated Center for Development of Statistical Computing, the Center for Information on Statistical Sciences, and the Engineering and Technical Services Section were integrated into the Center for Engineering and Technical Support. The affiliated facilities were reorganized as research departments, and the Risk Analysis Research Center was instituted.
April, 2008	The Research Innovation Center was instituted in the affiliated facility. The Intellectual Property Unit was instituted.

The Institute of Statistical Mathematics



Access to the ISM

◎ Tokyo Metro Hibiya Line, Hiroo Sta.
About 7 min on foot

◎ Tokyo Metro Namboku Line, Azabu-jūban Sta.
Toei Ōedo Line, Azabu-jūban Sta.
About 20 min on foot

Inter-University Research Institute Corporation
Research Organization of Information and Systems
The Institute of Statistical Mathematics
4-6-7 Minami-Azabu, Minato-ku, Tokyo 106-8569, Japan
Tel: +81-(0)3-3446-1501(Reception) Fax: +81-(0)3-3443-3552(General Affairs)
http://www.ism.ac.jp/