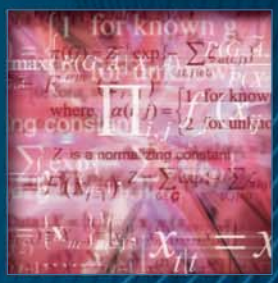
Research Organization of Information and Systems

The Institute of Statistical Mathematics

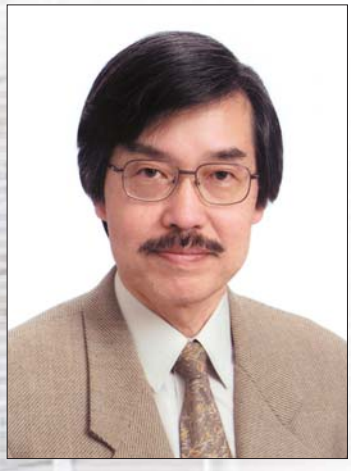
2009-2010

ISM



CONTENTS

Message from Director-General



It has already been six years since the Institute of Statistical Mathematics was incorporated into and became a member of the Research Organization of Information and Systems. During this period, we have proceeded with research projects focused on current important subjects, including the establishment of strategic research centers in 2005 in addition to the existing departments that conduct basic research. The Risk Analysis Research Center has since steadily developed, and currently, its four research groups are engaged in the fields of food and drug safety; environmental risks; financial risks and insurance; and reliability and quality assurance of services and products. In particular, since the center's NOE-forming project was approved for special funds for education and research this fiscal year, the center has started more earnest efforts. The Prediction and Knowledge Discovery Research Center has proceeded with four research projects consisting of research on statistical seismology, data assimilation, molecular evolution, and statistical genome diversity. Among other activities, the Data Assimilation Research Group's activity is being established as a next-generation simulation method that can be adopted not only in global environmental prediction but also in other extensive areas. At the Research Innovation Center, which was established two years ago, four research groups are performing activities in order to create research areas on statistical science. The Advanced Monte Carlo Algorithm Research Group has employed a foreign professor for the first time and started a top-level study.

The relocation to Tachikawa, which was approved at a Cabinet meeting in 1988, finally realized in October this year. It was also decided to construct a guest house required for joint research projects involving stays, which had been a pending issue, and its construction is about to start and is expected to be completed within this fiscal year. After moving to Tachikawa, the institute will strive to play the role as an inter-university research institute based on new concepts, such as NOE formation, in addition to conventional application-based joint research projects, by making the most of the larger joint research space and the facilities of the new campus. In line with the remarkable progress of information and communications technologies, both the society and the academic research circles are about to drastically change. Noticeably, the use of large-scale data will be the key to the progress of science and technology in the future, and it is imperative to establish data-centric science, which is regarded as a fourth scientific methodology. Statistical research, whose major goal is to perform scientific reasoning based on data, will become even more important in the future. We look forward to your 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 Centers

Molecular Evolution Research Group

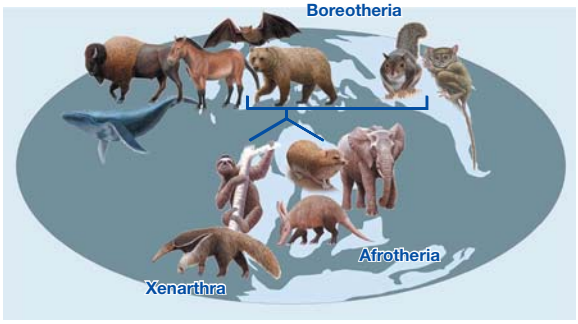
Toward Integrative Understanding of Biodiversity

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

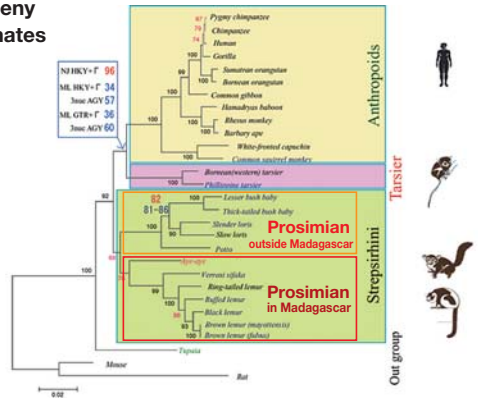
Continental configuration of 100 Myr BP and Eutherian evolution



Convergent evolution: Hedgehog and Tenrec

Lemurs present remarkable diversity in Madagascer

Phylogeny of Primates



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

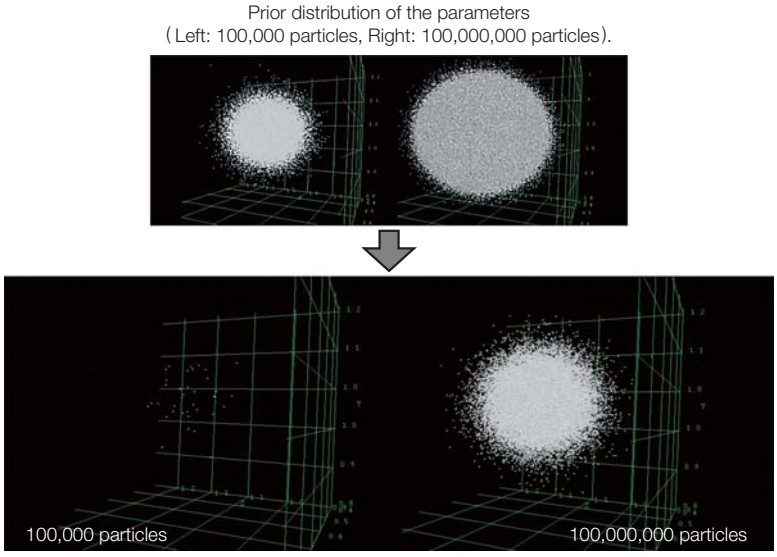


Figure 1: Numerical experiments of the data assimilation based on the particle filter technique for developing the personalized simulation tools for the living matter.

Figure 2:
Data assimilation research is carried out with a wide variety of the parallel computing systems including the handmade parallel computing system (above) as well as a general-purpose super computer.

Statistical Seismology Research Group

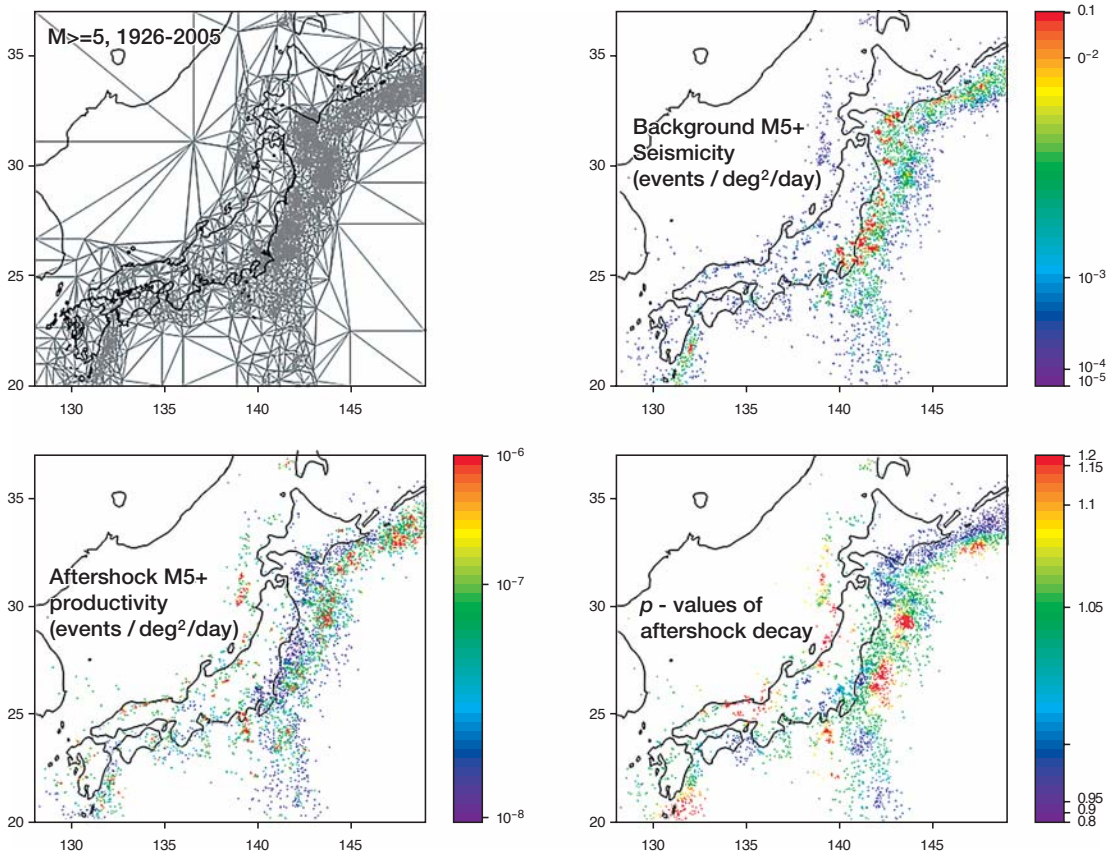
Location Dependent Space-time ETAS Model for Wide Regional and Global Seismicity

We consider the space-time ETAS (epidemic-type aftershock sequence) model which takes account of different regional physical characteristics of the earth crusts. Its parameters such as *p*-value for aftershock decay vary from place to place. The parameter variations of the estimated space-time model are visualized and used to investigate features of the regional seismic activity.

Specifically, each parameter is a 2-dimensional piecewise linear function whose value at a location is linearly interpolated by the three values at the location of the nearest three earthquakes (Delaunay triangle vertices) on the tessellated plane by epicenters. Such modeling by using Delaunay tessellation is suited for the observation on clustered points. The estimates of the parameter functions are simultaneously adjusted by the penalized log-likelihood that defines a trade-off between the goodness of fit to the data and uniformity constraint of the function (i.e., each facet of the piecewise linear function being as flat as possible). The constraint can be objectively adjusted from the data by means of an empirical Bayesian method using the Akaike's Bayesian information criterion (*ABIC*)

We are particularly concerned with the spatial estimates of the first two parameters of the space-time model: namely, μ-values of the background seismicity for the long-term prediction of the large earthquakes; and aftershock productivity *K*-values for the immediate aftershock probability forecast. The reasons and their utility of the model have been demonstrated by applying it to Japan and global seismicity. For both objectives, we further need the spatial estimates of the Gutenberg-Richter *b*-values that are also modeled using similar function on Delaunay tessellated space and optimized by the *ABIC* procedure.

Yosihiko Ogata



Statistical Genome Diversity Research Group

Gene Selection for Association Study from Gene Expressions to Phenotypes

Gene expression association study

We aim at proposing statistical methodology to provide efficiently knowledge discovery on the basis of genomics and omics data sets that are produced by tremendous developments in bioscience and biotechnology. In gene expression association studies methods for data analyses are quite contrast between research fields of bioinformatics and medicine. This comes from how does one trade off double-edged errors of false discovery and missing discovery in the presence of typical aspects in high-dimensional data and small sample size. However that two different methods are existing is problematic.

Nevertheless such controversy one strongly expects realization for personalized treatments based on genome and omics information. In fact, the prognosis prediction kit, called MammaPrint was accepted by Food and Drug Administration in U.S. In this way there is left the basic problems unsolved as mentioned as above. Our research goal is to build a consistent method to analyze genome data by fusion between the methods in clinical medicine and bioinformatics.

Gene selection

We presently analyze gene expressions data and clinical data with treatment effect by joint work with a research group in National Cancer Center. The objective is to exploit a prediction method for treatment effect given a set of gene expressions by microarray. For this we build the following three steps:

1. Around 26,000 gene expressions are filtered to around 2,500 gene expressions by referring to knowledge of gene function discussed in the literature.
2. Pick up the best 100 gene expressions according to the absolute value of 2,500 correlation coefficients with the binary outcome for prognosis.
3. AdaBoost for pattern recognition is sequentially applied to the data sets from the best 5 gene expressions by variable increment method.

Figure 1 is a hierarchical clustering plot in 100 gene expressions and the binary outcome. This shows possibly a strong association between gene expression and treatment effect. Figure 2 is the result of AdaBoost for pattern recognition from 5 gene expressions to the outcome based on 25 training samples and 16 test samples. We result good performance with 15 hit out of 16 samples. In future we plan to get a new data set to validate this result in several aspects for clinical information.

Shinto Eguchi

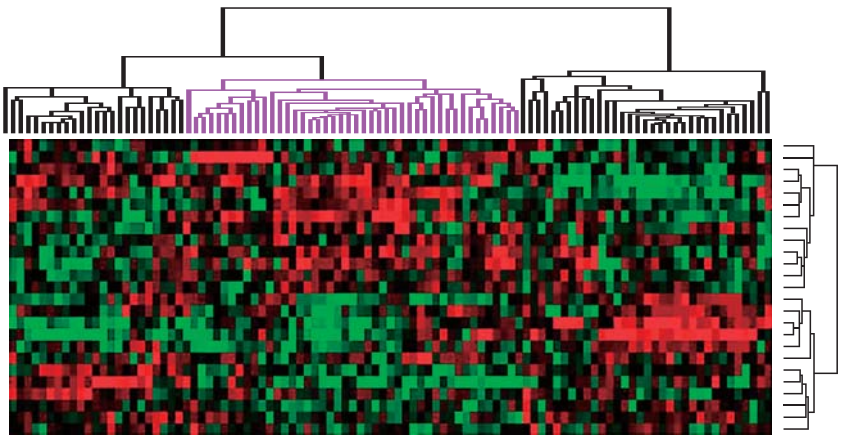


Figure 1: Hierarchical clustering for 100 genes and 25 subjects.

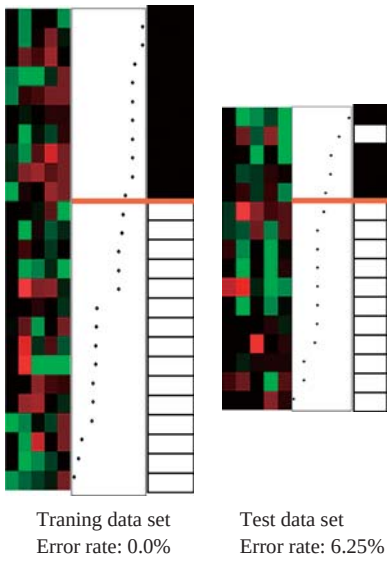


Figure 2: 5 probe classification by AdaBoost

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.

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. 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, and we conduct benefit / risk analyses of drug effect.

Genome-wide association study between phenotype and large-scale SNPs data

We study the relationship between effectiveness of treatment for hypertension and about 500,000 SNPs genotype data to find candidate genes for the good selection of drug for hypertension. After a separation of training group or test group from the sample, we examine a genome-wide association test for training group. Then it is gotten Figure 1 shows the loci and LD (D') map with chromosome position and LD blocks on chromosome 2. LD maps have been made using Haploview (Barrett et al., 2005) that is a very popular program in the statistical genetic world. For each LD block, we estimate a penetrance and test the likelihood ratio for SNPs and haplotypes to obtain candidate genes. If results of candidate genes are reproduced in the test group, it may be useful information for good selection of antihypertensive drug.

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. We participate in an ad hoc study on a specific drug safety issue as experts of statistical science and epidemiology.

Toshiharu Fujita

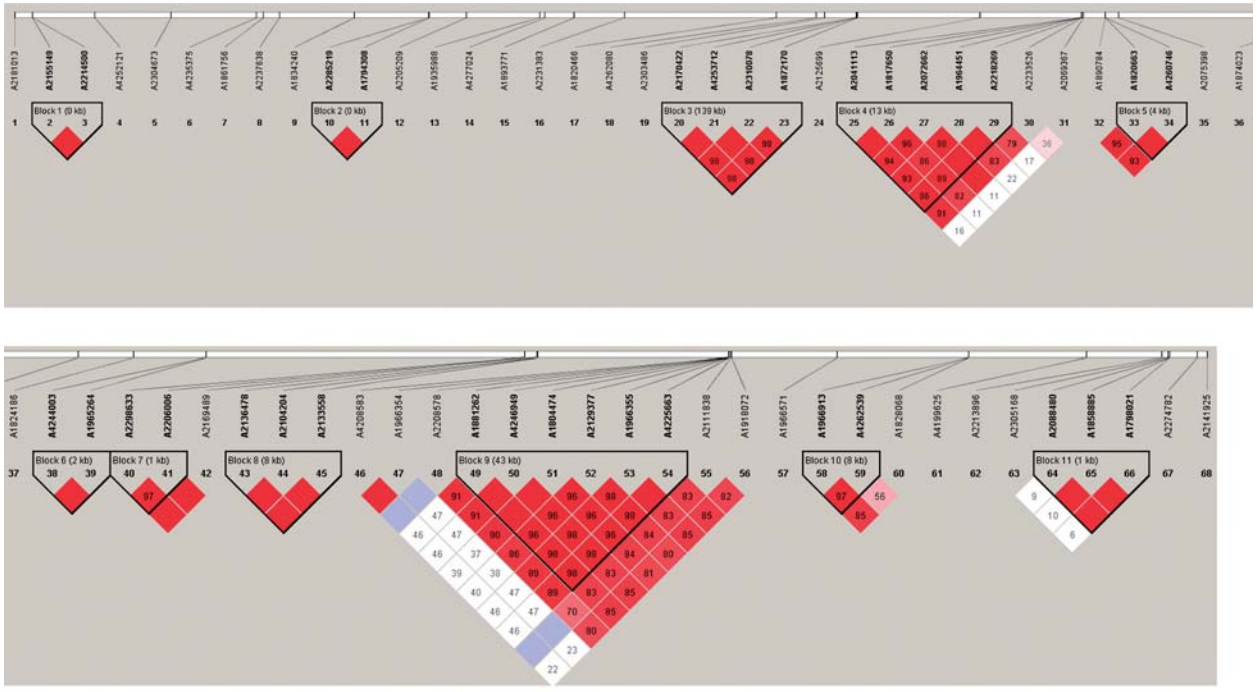


Figure 1: Ldmap(D') and LD blocks on Chr.2 by Haploview

Environmental Risk Research Group

The Effort of the Statistical Science towards Solution of Environmental Problems

■ 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.

■ Long-term variation of water quality in Tokyo Bay

Although the pollution load flowing into Tokyo Bay has been reduced, red tide still breaks out frequently. Red tide increases the amount of oxygen deficient water mass, and this sometimes causes blue tide. The occurrence of red and blue tides seems to be connected with elevated water temperature and change in salinity of recent date in Tokyo Bay. This study is making a collection of water quality data of Tokyo Bay and is analyzing these data to reveal long-term variation of water quality. (photo 1, photo 2)

■ On persistent organic pollutants

To identify the sources of pollution by persistent organic pollutants (POPs), this study is developing statistical methods to make inference on pollution sources besides organizing POPs data. Further, this study is making contribution to the development of technology to reduce the risk of POPs in vegetable farming. (photo 3, photo 4)

■ Risk-hedge oriented optimization system against natural and managerial disaster toward sustainable forest resource management

After the Kyoto protocol entered into force, reduction of carbon dioxide in the atmosphere is promised to slow down the degree of global warming. Among the proposed mechanisms under the Kyoto protocol, forest resources are regarded as a source of carbon sequestration, where forestation, deforestation and reforestation are the main human-induced activities. Housing and furniture manufacturing is also contributing to carbon mitigation indirectly. Despite of these facts, unfavorable current economic situation for forestry business discourages forest owners to abandon forestry practice in mountainous regions. Besides, typhoon disaster and snow damage on forest stands is accelerating abandoning phenomena. In our research project, we construct a risk-hedge oriented optimization system toward sustainable forest resource management. In the system, we try to predict events of natural disaster and managerial disaster spatially and temporally, and control the management scheme to avoid these disasters with a view of these phenomena. (photo 5, Fig 1, Fig 2)

Koji Kanefuji

Photo 1: Red tide (The Tokyo Metropolitan Research Institute for Environmental Protection)

Photo 2: Blue tide (Chiba Prefectural Environmental Research Center)

Photo 3: Growing experiment of cucumbers (Niigata Agricultural Research Institute)

Photo 4: Measurement of POPs (National Institute for Agro-Environmental Sciences)

Photo 5: Windstorm in Slovakia

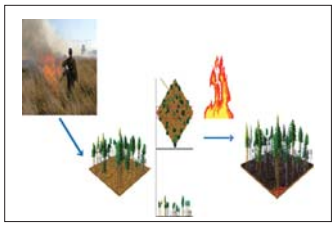


Figure 1: Fire Management

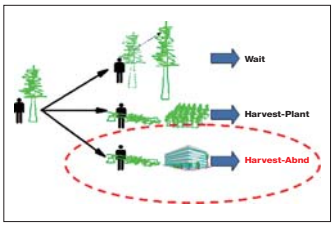


Figure 2: Decision Tree

Financial Risk and Insurance Research Group

Additional Loan and Credit Risk

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 a cooperative research by Dr Satoshi Yamashita (Associate Professor) and Mr Toshinao Yoshiba (Visiting Associate Professor, Bank of Japan).

Additional Loan and Loss of the Bank

Credit risk models investigate the default probability of a firm and the recovery rate of the loan to the firm. Popular one-period structural models capture the credit risk by the relationship between asset and liability of the firm. And they assume that the loan amount does not change until the maturity. However, a bank sometimes supplies an additional loan to the firm, in practice.

Even if expected growth rate of the firm is high, the firm's asset may become worse. In Figure 1, time t in the horizontal axis is the case. If nothing is done, further asset deterioration may lead to the bankruptcy in such a case most of the account cannot be collectible for the bank. On the other hand, the bank can stimulate the firm by supplying an additional loan so that it may decrease the expected loss of the bank. In that sense, supplying the additional loan is a rational behavior of the bank.

Expected Loss of the Bank under Downturn Macroeconomic Condition

Even though the additional loan is rational, if the firm's asset movement is highly correlated with the macroeconomic condition, the loss of the bank may increase when the macroeconomic condition worsens. Under the major premise that the bank supplies an additional loan to minimize the expected loss at some time, this research analytically derives the conditional expected loss with some stressed condition when the macroeconomy is in recession. Figure 2 depicts how the conditional expected loss varies with respect to the correlation between the firm's asset movement and the macroeconomic condition. When the correlation is large, so is the conditional expected loss, which implies banks should hold large capital in recession.

Yoshinori Kawasaki

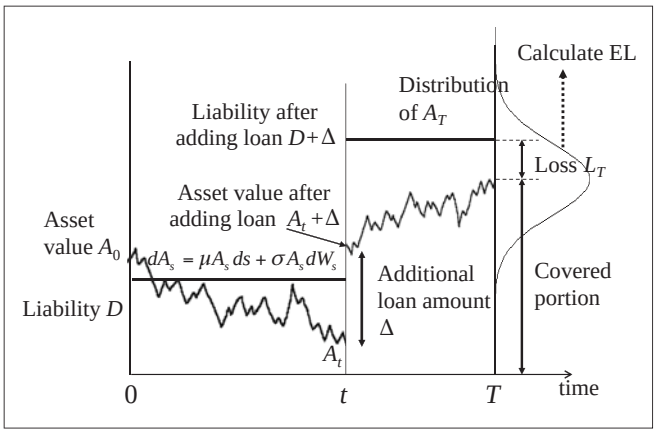


Figure 1: Schematic diagram of the model

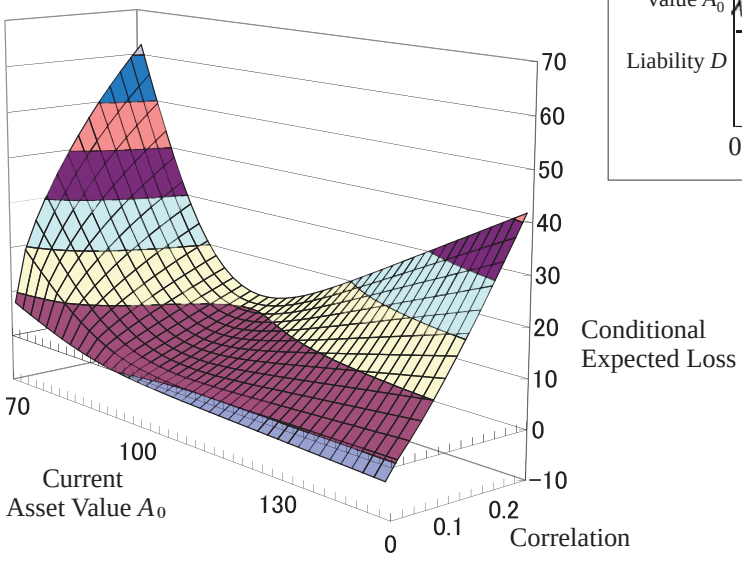


Figure 2: Correlation and Conditional Expected Loss

Research Group for Reliability and Quality Assurance of Service and Product

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 method

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 ascertain the current situation, investigate the causes and analyze the important factors. On the other hand, Taguchi method 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.

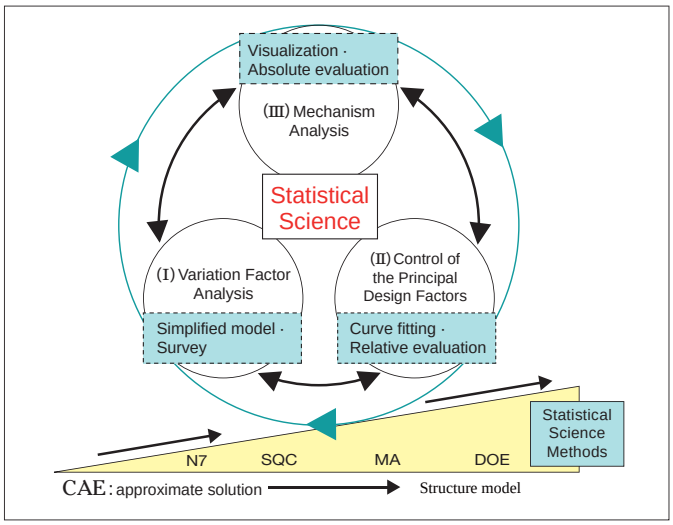


Figure 1: Highly Reliable CAE Analysis System Approach Method

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 error factors into the model as parameter factors, and considering the probability distribution of SN ratios as a statistical quantity, thereby enabling a discussion of statistical inference.

■ Highly Reliable CAE Analysis System Approach Method

In this study, consideration is given to the validity of "statistical science", which is necessary for innovation of the development designing process, for use in a new issue involving numerical simulation technology (CAE) in development designing. This new application of statistical science is designed to aid the transition from the "prototype / experimental based real machine evaluation method" to the "prediction evaluation method based on highly reliable CAE analysis." More specifically, concrete examples of "statistical science study cases that contributed to CAE analysis" from automotive manufacturers are examined from the standpoint of "statistical science contributing to design science". This was done so as to propose the "highly reliable CAE analysis system approach method" as shown in Fig. 1 that contributes to the process innovation of development designing.

■ Data Mining for Hospital Management

Organizations in our modern society grow larger and more complex to provide advanced services due to the varieties of social demands. Such organizations are highly efficient for routine work processes but known to be not robust to unexpected situations. According to this observation, the importance of the organizational risk management has been noticed in recent years. On the other hand, a large amount of data on the work processes has been automatically stored since information technology was introduced to the organizations. Thus, it has been expected that reuse of collected data should contribute to risk management for large-scale organizations. In this research, we focus on hospital information system, to which we applied temporal data mining and exploratory data analysis techniques and successfully quantified several risk factors for hospital management.

Toshihiko Kawamura

Social Survey Information Research Group

Building a Database of Social Survey Results

■ Objective of the Social Survey Information Research Group

The Social Survey Information Research Group of the Research Innovation Center has been collecting social survey results with the aim of building a database of social surveys. A single survey provides limited information, particularly in the recent deteriorated survey environment. Multidimensional analysis using various types of social surveys is essential for comprehending the complicated aspects of modern society. This research group studies the methodologies for building a database and develops statistical methods for analyzing data. The group also intends to make the database available for public use.

■ A comprehensive report of a survey on the Japanese national character

The core data used to develop the database is obtained from a survey on the Japanese national character, which is conducted on a regular basis by the Institute of Statistical Mathematics. This survey has been conducted every five years since the first 1953 survey. The most recent was its twelfth survey and was conducted in the autumn of 2008. The questionnaire items in the survey concerned the attitudes of the participants toward life, environment, religion, politics, race, the Japanese people, and so forth. The results of this longitudinal survey, which are crucial for evincing the transition of Japanese character, are already available on the Internet. More comprehensive results, including cross-tabulation tables and graphs, will be also provided on the Internet.

■ Database of related surveys

The survey on the Japanese national character is by no means the only survey conducted by the Institute. Various interrelated comparative surveys and experimental surveys have also been conducted. The results of these surveys will also be included in the database and made accessible to the public one at a time. The potential users of this database, including not only researchers but also the general public, will be able to analyze the data from various perspectives.

Takahiro Tsuchiya

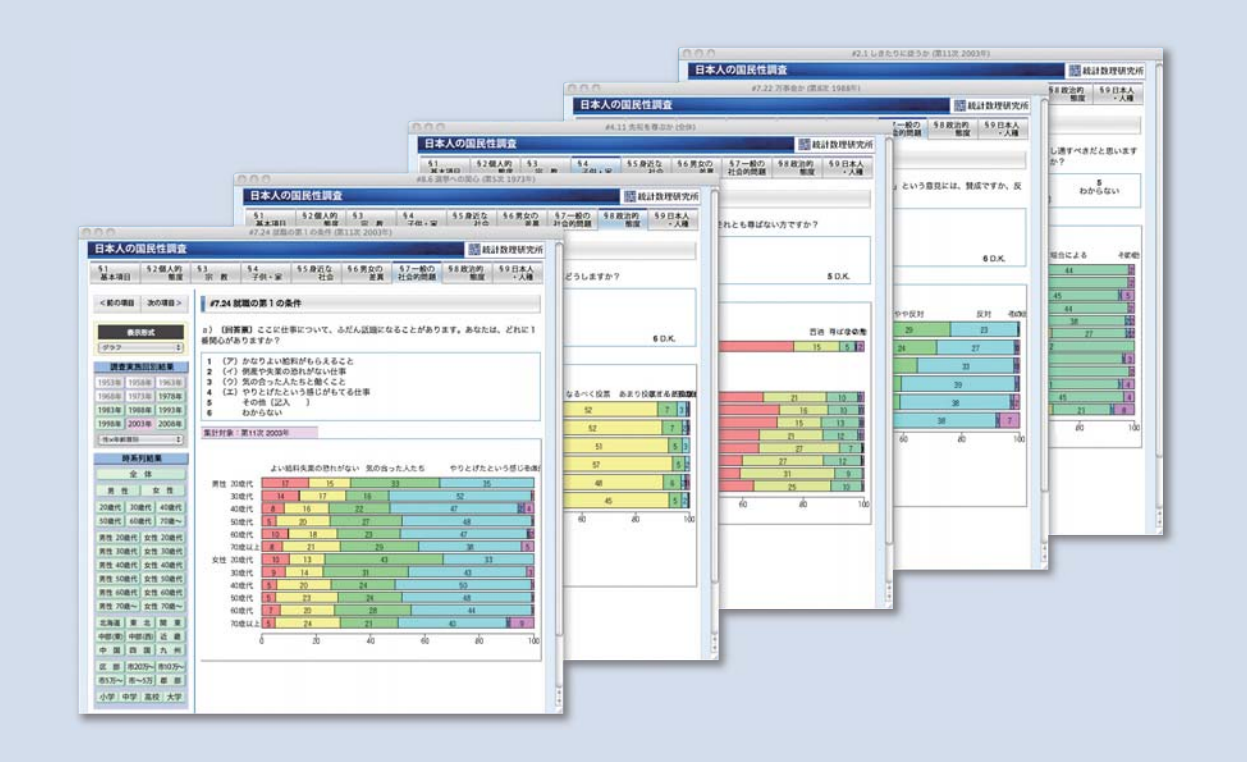


Figure : Results of a survey on the Japanese national character

Functional Analytic Inference Research Group

Kernel Method: Nonlinear Data Analysis with Efficient Computation

What is kernel method?

Kernel method is a recent methodology for analyzing data of complex structure. Instead of pursuing elaborate modeling of data, the method uses a feature map of data into a high-dimensional functional space to extract the high order moments of the data. While methods of mapping data by power series expansion or basis function expansion have been long used for the high order moments, they have a problem of computational explosion for high-dimension data. The kernel methods avoid this problem by using the clever device of "positive definite kernel", which computes the inner product of the high dimensional functional space by simply evaluating the kernel values. This gives an efficient method for incorporating high order moments of data. In addition, another strong point of the kernel method is that it provides a seamless methodology for analysis of non-vectorial data such as network data and symbol sequence such as texts.

Dependence analysis with kernels

The kernel method was first developed as "kernelization" of existing linear methods: support vector machine and kernel principal component analysis are among such examples. More recent studies, however, have revealed that it provides a method for analyzing homogeneity, independence and conditional independence of variables by representing high order moments with feature maps. Our group has been making essential contributions in this line of research, in particular in the characterization of conditional independence and suitable kernels for such analysis. In addition, we have proposed a new kernel method of estimating causal networks from data, which is constructed based on the kernel device of relevant conditional independence tests. We also carry out a project on approximation of large matrix computation necessary for kernel methods.

International research network

Our group has active collaboration networks with researchers in the world, in particular, with top researchers in University of California, Berkeley, and Max Planck Institute in Germany. Our research results have been presented in tutorials of various international conferences such as Machine Learning Summer School 2007 (Germany) and International Conference on Machine Learning 2008 (Finland).

Kenji Fukumizu

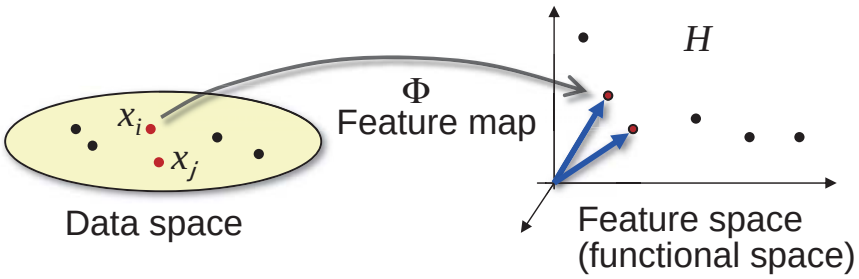


Figure 1: Feature map of kernel method

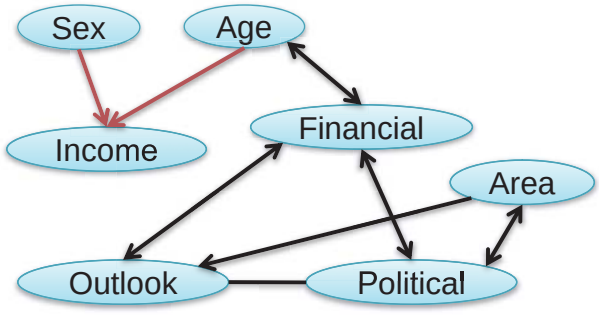


Figure 2: An example of causal network given by the kernel method

Advanced Monte Carlo Algorithm Research Group

Rare Event Sampling by Markov Chain Monte Carlo

Our research focuses on the development of novel and nonstandard applications of Markov Chain Monte Carlo algorithm (MCMC). MCMC is introduced in physics in 1950s by Metropolis and coworkers and proved to be a useful tool in the field of Bayesian modeling in 1990s. MCMC is, however, a general strategy for sampling multivariate distributions with unknown normalization constants and in principle can be applied to many other problems.

Here we discuss rare event sampling (i.e., sampling from the tails of the given distribution) problem as an example of the promising application fields of MCMC. By using MCMC, especially with its variants "Replica Exchange Monte Carlo" and "Multicanonical Sampling", we can generate samples of very small probability, say, 10^{-15}, or even 10^{-30}, under the given model, and calculate its probability correctly. Today's most powerful computer can never achieve such an ultimate precision with naïve random sampling.

With this approach we can solve various interesting problems. For example, we can sample very rare errors in an information processing systems and access its probability. Also, we can find atypical but important regular orbits in chaotic systems. For network design problems, MCMC can generate not only a network that optimizes the given criterion, but also a collection of networks that belong to the tail of the distribution defined by the given criterion.

Because these problems are highly cross-disciplinary, we are collaborating specialists of each field and exploring frontiers of MCMC with scientists at Univ. Tokyo, Hokkaido Univ., Osaka Univ., and Kyoto Univ..

Yukito Iba

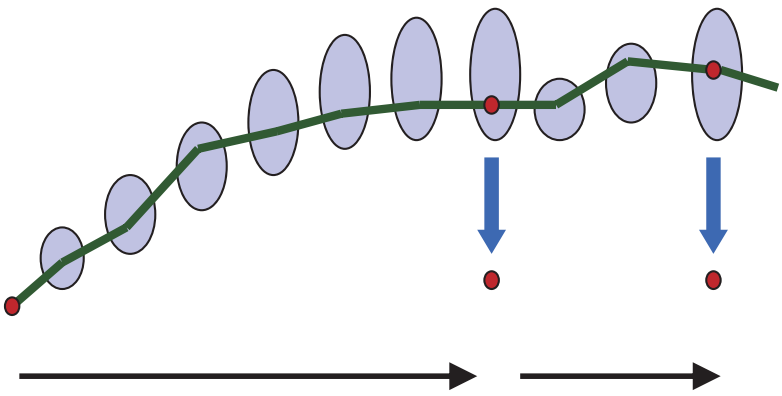


Figure 1: Markov Chain Monte Carlo

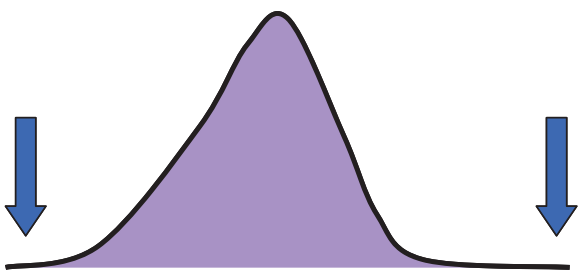


Figure 2: Rare events at the tail of the distribution

Random Number Research Group

Parallel Pseudo Random Number Generation and Hardware Random Number Generator

■ Random number

Random number is defined a sequence of random variables. We think binary random variables with equal probability. It is easy to understand that we can generate this random numbers form a result of finite number of honest coin tossing.

■ Random number generator

We call methods to generate random numbers random number generators. The random number sequence generated by a computer is called pseudo random numbers. On the other hand the random number sequence generated by a physical phenomenon is called physical random numbers. In general pseudo random numbers has uniform distribution on [0, 1). There are linear congruential method, M-sequence, Mersenne Twister and so on. The most popular phenomenon which is used for a hardware random number generator is thermal noise of electronic circuit. Recently chaos phenomenon of semiconductor laser is used for a hardware random number generator.

■ Parallel pseudo random number generation

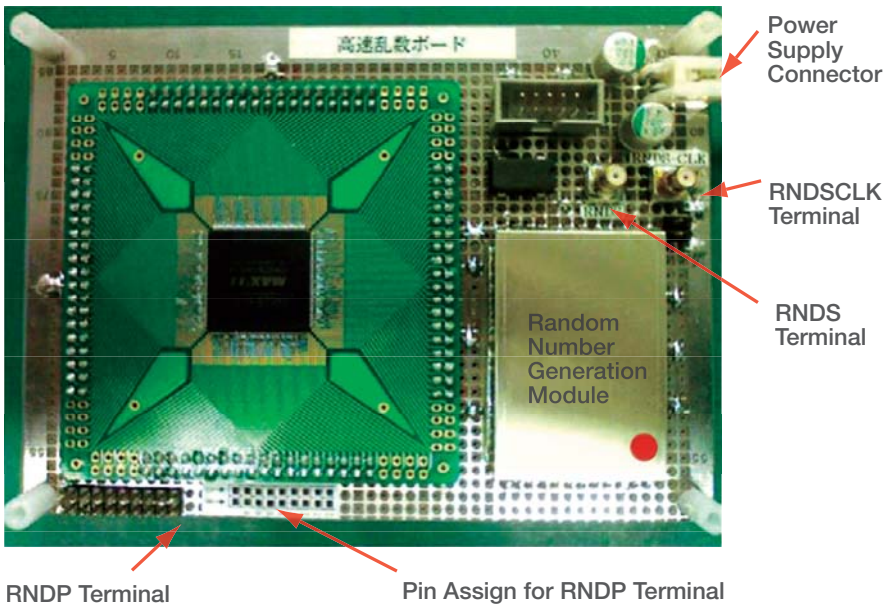
Random numbers are used for Monte Carlo method, MCMC, bootstrap and so on. Large-scale computations are necessary to large-scale problems. Large numbers of random numbers are necessary to large-scale computations. It seems that we cannot do this computation by only one CPU. Parallel computations can bring us some results. Well-known methods for parallel pseudo random number generation are leap frog method and cycle splitting method. We cannot use these methods free from care, since it is too difficult to confirm whether two sequences have correlation or not.

■ Hardware random number generator

At the Institute of Statistical Mathematics, hardware random number generators are researched for a long period. Our group also research new hardware random number generators which are faster and have better quality. We show a test board of new generator. It is important to use two noise sources. We have plan to product and deliver new hardware random number generator.

Yoshiyasu Tamura

Figure : Prototype of hardware random number generator



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

2003	2004	2005	2006	2007	2008
99	108	124	122	120	138

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

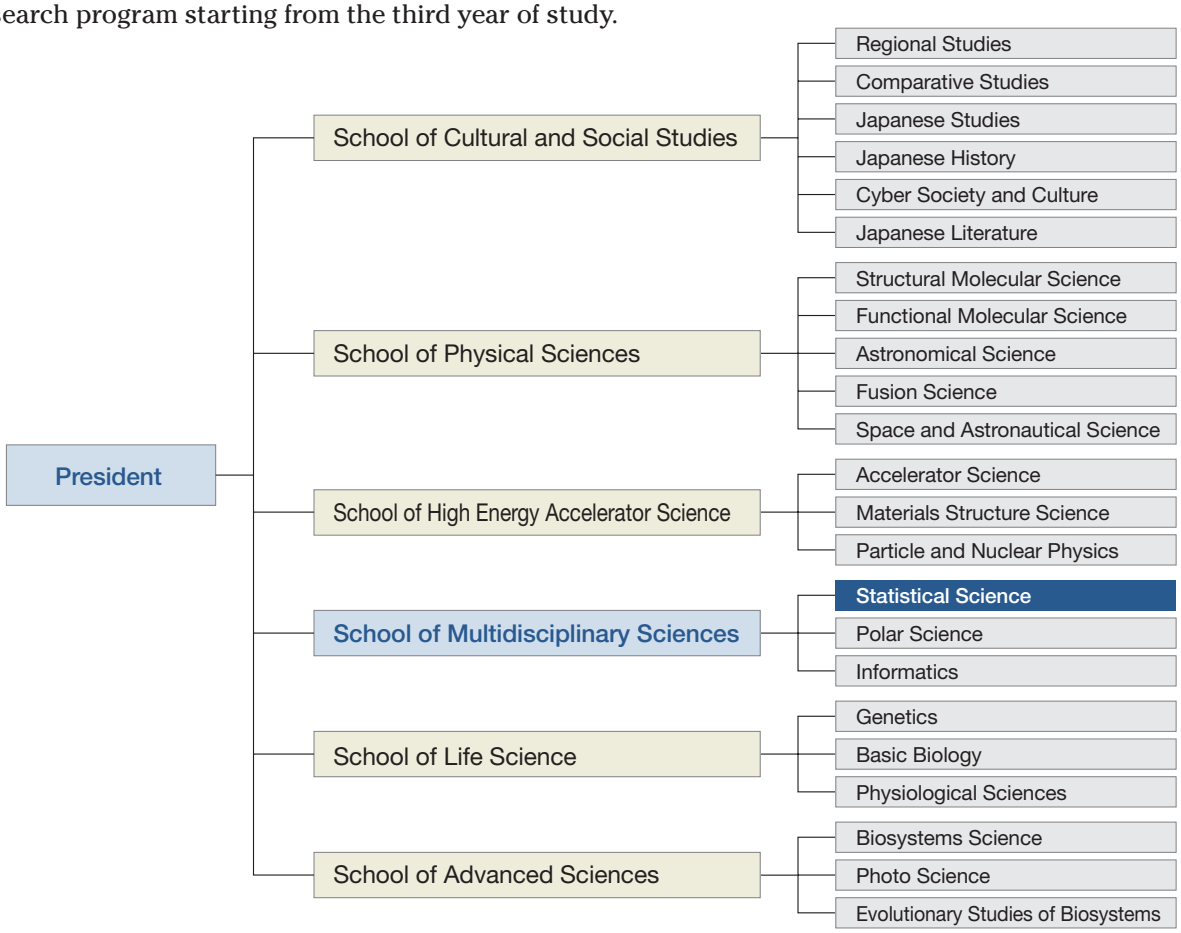
Major Research Fields

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.



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, 2009)

■ Doctor's course three years

Year of enrollment	2002	2003	2004	2005	2006	2007	2008	2009
Number of students	1 ①	1 ①	3 ③	2 ②	1 (1)	5 ③	6 ④	4 ③

* 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	2009
Number of students	1	2	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 (5) • Chiba University • The University of Tokyo (14) • 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 (6) • Tokyo University of Science (5) • Toyo University • Nihon University (2)
• Japan Women's University • Hosei University (7) • Waseda University (5) • Nanzan University (1) • 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
2002	4
2003	8 [1]
2004	4
2005	4
2006	8 [1]
2007	7 [1]
2008	4 [1]

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

Current Position of Alumni (As of April 1, 2009)

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 (7)
• 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

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 2008, the number of courses was 13.

Courses

The total number of courses held from 1969 to March, 2009 was 246, with a total of 18,266 participants. These courses covered a wide range of fields from basic to applied statistics. The following table lists the courses held in the past 3 years:

Year	Category	Title	Month	Number of participants
2006	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
	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

Year	Category	Title	Month	Number of participants
	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
	Basic course	Data Analysis and Simulation with R	May	69
	Basic course	The Grammar of Science and Descriptive Statistical Methods	June	47
	Basic course	Introduction to Statistics	July	79
	Standard course	Spatial Point Pattern Analysis – Introduction from Biological Examples	August	46
	Standard course	Statistics of Extremes	August	48
	Standard course	An Introduction to Statistical Analysis by the Theory of Martingales	September	48
	Basic course	Analysis of Sample Surveys with R	September	41
	Basic course	Introduction to Multivariate Analysis	October	71
	Standard course	Data Assimilation: State Space Model and Simulation	October	72
	Basic course	Quality Control, Quality Engineering (Taguchi Method)	November	17
	Basic course	Introduction to Pharmacoepidemiology	November	71
	Standard course	Understanding of Evolutionary Computation: Statistical Viewpoint	December	56
2009	Standard course	Markov Chain Monte Carlo : Basics and Examples	February	78

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

Statistical Mathematics 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. 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 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/

Consultation about Statistical Science

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 advice receivers 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 Meeting for 2008

The Annual Research Meeting of the Institute was held on March 18 and 19, 2009. This annual meeting is an opportunity for the Institute's researchers and visiting researchers 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 13 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 48 research education staff and 6 visiting researchers then gave their presentations. In addition, a poster session was held by 35 research fellows and graduate students of Sokendai. 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.

http://www.ism.ac.jp/

Finance and Buildings

Administration Subsidy and Others (2008)

Type	Personnel expenses	Non-personnel expenses	Total
Expenditure	762,681	1,080,970	1,843,651

Unit: ¥1,000

Accepted External Funds (2008)

Type	Subcontracted research	Joint research	Contribution for scholarship	Total
Items	6	3	2	11
Income	84,624	13,650	32,000	130,274

Unit: ¥1,000

Grant-in-Aid for Scientific Research (2008)

Research Category	Items	Amount Granted
Grant-in-Aid for Scientific Research on Priority Areas	1	2,700
Grant-in-Aid for Scientific Research (A)	6	65,650
Grant-in-Aid for Scientific Research (B)	5	25,610
Grant-in-Aid for Scientific Research (C)	13	17,290
Grant-in-Aid for Exploratory Research	2	1,600
Grant-in-Aid for Young Scientists (A)	1	13,390
Grant-in-Aid for Young Scientists (B)	8	8,190
Grant-in-Aid for Young Scientists (Start-up)	1	1,716
Grant-in-Aid for JSPS Fellows	1	1,100
Total	38	137,246

Unit: ¥1,000

Site and Buildings (As of October 1, 2009)

Site Area	62,450m^2
Area for Buildings (total)	11,855m^2

Introduction to the Tachikawa Campus

With three newly added strategic research centers, our campus in Hiroo has become increasingly crowded in recent years. The Tachikawa new campus offers ample space for conducting joint researches in a more extensive and flexible manner. Taking advantage of the large space available for joint researches and facilities for visitors, we, under a new concept, would like to reinforce the functions of inter-university research institute by helping develop Networks of Excellence (NOEs) and allowing for long-stay joint researches as well as the traditional form of open-type joint researches, and contribute to academic and social development.


Front of Building


East Side of Building

Main Gate


Entrance Hall


Atrium Space


Library


Auditorium

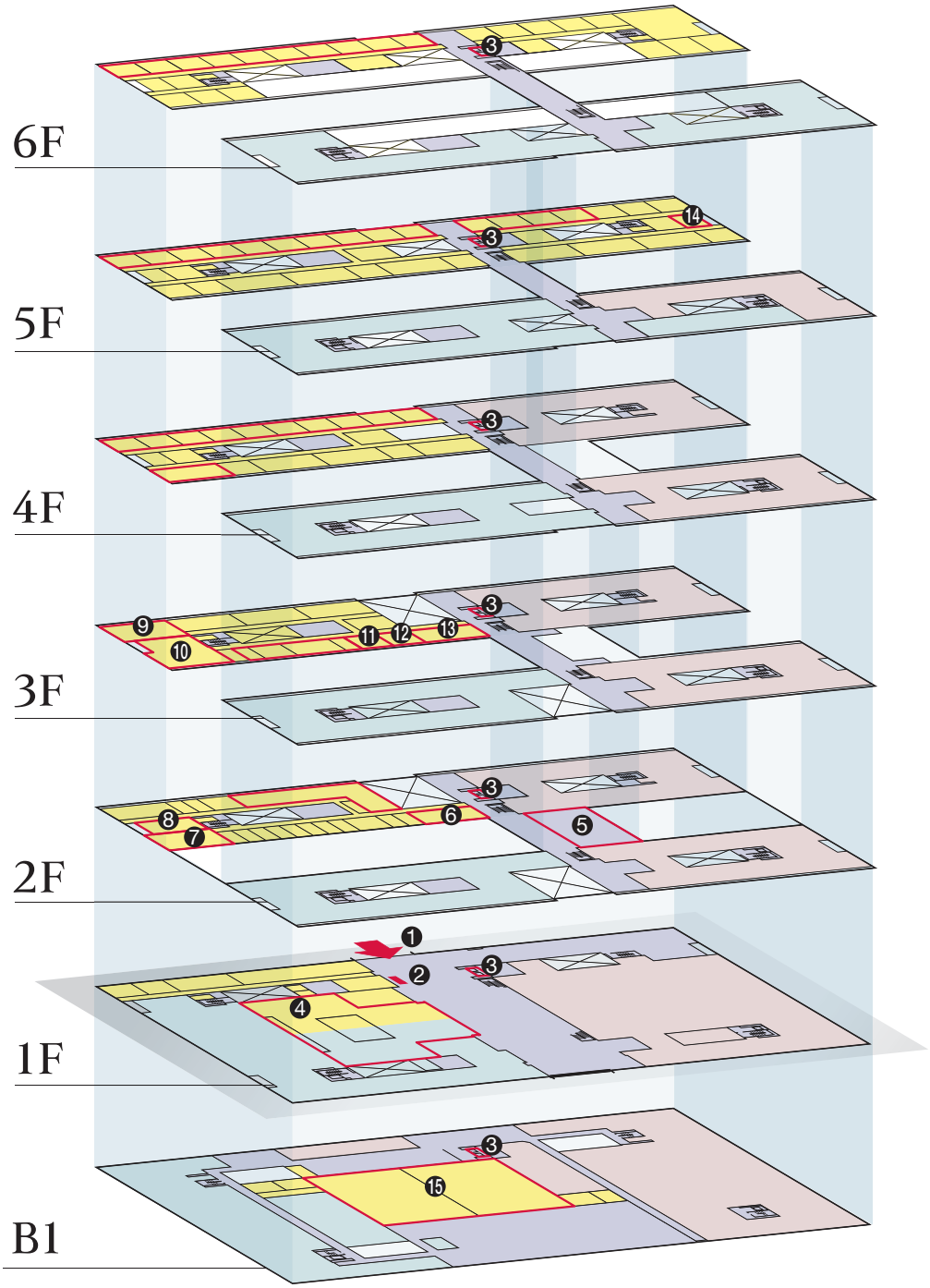

Conference Room


Seminar Room


Lounge


Refresh Corner


Courtyard



The Institute of Statistical Mathematics
National Institute of Polar Research
National Institute of Japanese Literature
Common Area

❶ Main Entrance
❷ Reception
❸ Lift
❹ Library
❺ Auditorium
❻ Conference Room 1
❼ Conference Room 2
❽ Seminar Room 1
❾ Seminar Room 2
❿ Lecture Hall
⓫ Seminar Room 3
⓬ Seminar Room 4
⓭ Seminar Room 5
⓮ Seminar Room 6
⓯ Computer Room

Facilities and Equipment

Computation Resources (As of October 1, 2009)

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.

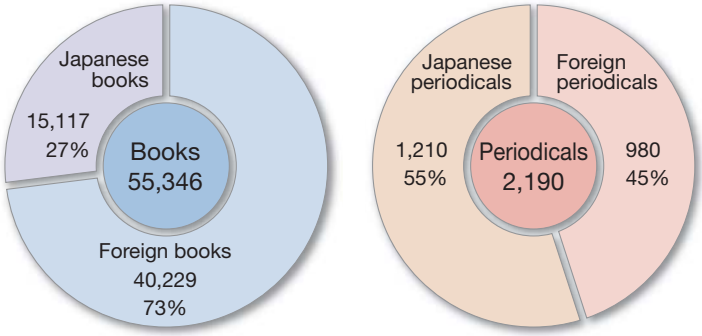
In October 2009, the Institute moved to a new building in Tachikawa, Tokyo. Present supercomputer systems will work until December 2009 in the old building in Hiroo. New supercomputer systems will be introduced in January 2010 in Tachikawa.

SGI Altix3700 Super Cluster, a parallel computer

Library and Materials (As of April 1, 2009)

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.

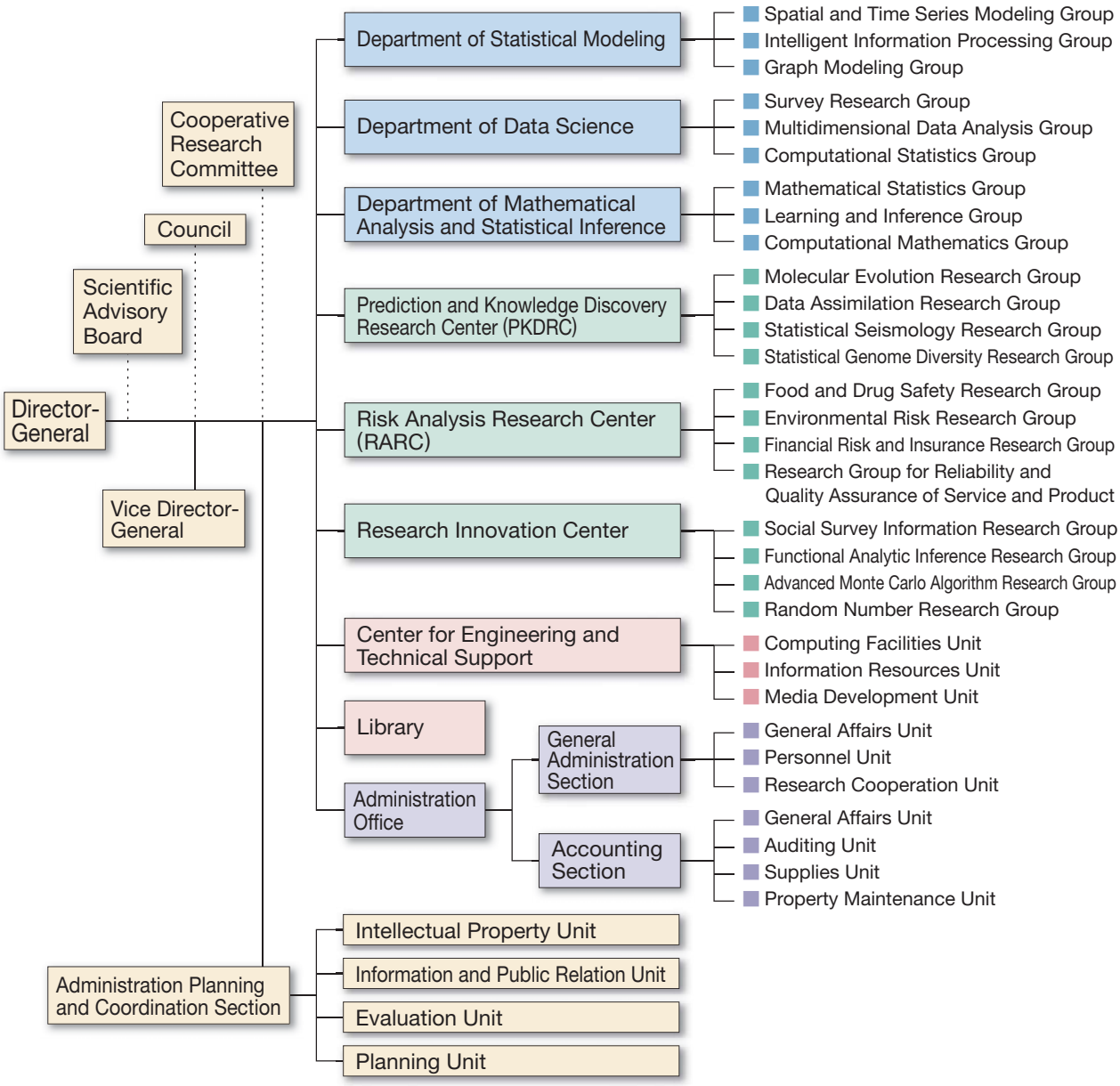
Besides contributed to Japanese and foreign publications, we also have a collection of journals that we publish ourselves: Annals of the Institute of Statistical Mathematics (English; Springer), Proceedings of the Institute of Statistical Mathematics (Japanese), Research Report (Statistical Researches 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, 2009)



Number of Staff (As of October 1, 2009)

Type	Director-General	Professor	Associate Professor	Assistant Professor	Administrative Staff	Technical Staff	Total
Director-General	1						1
Department of Statistical Modeling		6	6	5			17
Department of Data Science		8	6	5			19
Department of Mathematical Analysis and Statistical Inference		5	4	5			14
Center for Engineering and Technical Support						8(1)	8(1)
Administration Office					13	1	14
Total	1	19	16	15	13	9(1)	73(1)

( ) Staff of reemployment

Staff (As of October 1, 2009)

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.	Jiancang ZHUANG
Prof.	Yosihiko OGATA	Assoc. Prof.	Kenichiro SHIMATANI	Assist. Prof.	Shinya NAKANO
Prof.	Tomoyuki HIGUCHI	Assoc. Prof.	Genta UENO	Project Researcher	Hai Yen SIEW
		Assist. Prof.	Ryo YOSHIDA	Project Researcher	Eiki TANAKA

Intelligent Information Processing Group

Prof.	Makio ISHIGURO	Assoc. Prof.	Yukito IBA	Assist. Prof.	Hiroshi SOMEYA
Prof.	Tomoko MATSUI	Assoc. Prof.	Yumi TAKIZAWA	Project Researcher	Ken ISHIKAWA
Prof.	Kenji FUKUMIZU	Visiting Assoc. Prof.	Masataka GOTO		

Graph Modeling Group

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

Department of Data Science

Director Takashi NAKAMURA

Survey Research Group

Prof.	Takashi NAKAMURA	Assoc. Prof.	Tadahiko MAEDA	Assist. Prof.	Wataru MATSUMOTO
Prof.	Ryozo YOSHINO	Assoc. Prof.	Takahiro TSUCHIYA		
Visiting Prof.	Ikuo NASU	Visiting Assoc. Prof.	Takahiro HOSHINO		

Multidimensional Data Analysis Group

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

Computational Statistics Group

Prof.	Yoshiyasu TAMURA	Visiting Prof.	Kazunori YAMAGUCHI	Visiting Assoc. Prof.	Norikazu IKOMA
Prof.	Junji NAKANO	Assoc. Prof.	Naomasa MARUYAMA	Assist. Prof.	Nobuo SHIMIZU
Visiting Prof.	Michiko WATANABE	Assoc. Prof.	Koji KANEFUJI		
Visiting Prof.	Yuichi MORI	Assoc. Prof.	Seisho SATO		

Department of Mathematical Analysis and Statistical Inference

Director Shinto EGUCHI

Mathematical Statistics Group

Prof.	Satoshi KURIKI	Assist. Prof.	Takaaki SHIMURA	Assist. Prof.	Shogo KATO
Assoc. Prof.	Yoichi NISHIYAMA	Assist. Prof.	Kei KOBAYASHI	Project Researcher	Xiaoling DOU

Learning and Inference Group

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

Computational Mathematics Group

Prof.	Takashi TSUCHIYA	Prof.	Atsushi YOSHIMOTO	Project Researcher	Hayato TAKAHASHI
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	Assist. Prof.	Ying CAO	Adjunct Prof.	Masami HASEGAWA

Data Assimilation Research Group

Prof.	Tomoyuki HIGUCHI	Assist. Prof.	Ryo YOSHIDA	Project Researcher	Tomoko IMOTO
Assoc. Prof.	Genta UENO	Assist. Prof.	Shinya NAKANO	Project Researcher	Hiromichi NAGAO

Statistical Seismology Research Group

Prof.	Yosihiko OGATA	Assist. Prof.	Jiancang ZHUANG	Project Researcher	Ushio TANAKA
Visiting Prof.	Shinji TODA	Project Professor	Mitsuhiro MATSUURA		

Statistical Genome Diversity Research Group

Prof.	Shinto EGUCHI	Assoc. Prof.	Hironori FUJISAWA	Assist. Prof.	Tadayoshi FUSHIKI
Prof.	Satoshi KURIKI	Visiting Assoc. Prof.	Kanta NAITO	Assist. Prof.	Shogo KATO
Assoc. Prof.	Shiro IKEDA				

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	Visiting Prof.	Manabu IWASAKI	Assist. Prof.	Takaaki SHIMURA
Prof.	Hiroe TSUBAKI	Visiting Prof.	Tosiya SATO	Assist. Prof.	Masayuki HENMI
Prof.	Shigeyuki MATSUI	Visiting Prof.	Yoshimitsu HIEJIMA	Project Researcher	Yosuke FUJII
		Visiting Assoc. Prof.	Satoshi AOKI	Project Researcher	Takehiko YAMAGUCHI

Environmental Risk Research Group

Prof.	Nobuhisa KASHIWAGI	Visiting Prof.	Hidetoshi KONNO	Assoc. Prof.	Koji KANEFUJI
Prof.	Atsushi YOSHIMOTO	Visiting Prof.	Kazuo YAMAMOTO	Visiting Assoc. Prof.	Hirokazu TAKANASHI
Visiting Prof.	Yukio MATSUMOTO	Visiting Prof.	Yoshiro ONO	Visiting Assoc. Prof.	Tomohiro TASAKI
Visiting Prof.	Kunio SHIMIZU	Visiting Prof.	Hideshige TAKADA	Project Researcher	Masayuki KAGEYAMA

Financial Risk and Insurance Research Group

Visiting Prof.	Naoto KUNITOMO	Assoc. Prof.	Yoshinori KAWASAKI	Project Researcher	Masakazu ANDO
Visiting Prof.	Hiroshi TSUDA	Assoc. Prof.	Yoichi NISHIYAMA	Project Researcher	Kentaro AKASHI
Assoc. Prof.	Satoshi YAMASHITA	Visiting Assoc. Prof.	Toshinao YOSHIBA	Project Researcher	Yoko TANOKURA
Assoc. Prof.	Seisho SATO				

Research Group for Reliability and Quality Assurance of Service and Product

Prof.	Hiroe TSUBAKI	Visiting Prof.	Kazuo TATEBAYASHI	Visiting Prof.	Yoichi KATO
Visiting Prof.	Kakuro AMASAKA	Visiting Prof.	Sadaaki MIYAMOTO	Assist. Prof.	Toshihiko KAWAMURA
Visiting Prof.	Kosei IWASE	Visiting Prof.	Shusaku TSUMOTO	Project Researcher	Natsuki SANO

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

Prof.	Kenji FUKUMIZU	Assist. Prof.	Kei KOBAYASHI

Advanced Monte Carlo Algorithm Research Group

Assoc. Prof.	Yukito IBA	Visiting Prof.	Makoto KIKUCHI
Project Professor	Arnaud DOUCET	Visiting Assoc. Prof.	Koji FUKUSHIMA

Random Number Research Group

Prof.	Yoshiyasu TAMURA	Visiting Assoc. Prof.	Toru ONODERA	Assist. Prof.	Sumie UEDA

Project Researchers

Project Professor	Atsushi FUKASAWA	Project Researcher	Masaki OKUDA	Project Researcher	Masayuki KUMON
		Project Researcher	Takayuki FUJII	Project Researcher	Christopher Andrew ZAPART

Center for Engineering and Technical Support

Director	Junji NAKANO	Adjunct Prof.	Yasumasa BABA
Vice Director	Satoshi YAMASHITA	Deputy Manager	Yuriko WATANABE
		Senior Specialist	Saeko TANAKA

Head of Computing Facilitise Unit	Kazuhiro NAKAMURA	Head of Media Development Unit (Concurrent)	Yuriko WATANABE
Head of Information Resources Unit (Concurrent)	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	Head of Planning Unit	Tomoyuki HIGUCHI

Administration Office

General Manager Masaru SHIMODA

General Administration Section

Manager of General Administration Section	Yoshiki HAMA	Chief of General Affairs Unit (Concurrent)	Fumio SUTO
Deputy Manager of General Administration Section	Toshiyuki INADA	Chief of Personnel Unit	Motoyoshi URANO
Specialist (General Affairs)	Fumio SUTO	Chief of Research Cooperation Unit	Akihiko NAKAMURA

Accounting Section

Manager of Accounting Section	Toyokichi KITAHARA	Chief of Auditing Unit (Concurrent)	Hiroaki ARAI
Deputy Manager of Accounting Section	Hiroshi YAMAMOTO	Chief of Supplies Unit	Tatsuya TAKAGI
Chief of General Affairs Unit	Hiroaki ARAI	Chief of Property Maintenance Unit	Hirotomo SHIMIZU

Council of The Institute of Statistical Mathematics (As of April 1, 2009)

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 Emeritus, 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 June 1, 2009)

Akifumi OIKAWA	Professor and Presidential Aide, The Graduate University for Advanced Studies
Hidetoshi KONNO	Professor, Graduate School of System and Information Engineering, University of Tsukuba
Toshimitsu HAMASAKI	Associate Professor, Graduate School of Medicine, University of Osaka
Michiko WATANABE	Professor, Faculty of Economics, Toyo University
Yuichi MORI	Professor, Faculty of Informatics, Okayama University of Science
Makio ISHIGURO	Professor, (Department of Statistical Modeling, ISM)
Ryozo YOSHINO	Professor, (Department of Data Science, ISM)
Shinto EGUCHI	Professor, (Department of Mathematical Analysis and Statistical Inference, ISM)
Nobuhisa KASHIWAGI	Professor, (Department of Data Science, ISM)

Research Ethics Review Committee (As of April 1, 2009)

Specialist on Epidemiology and Social Research	Michio UMINO	Professor Extraordinary, Tohoku University
Specialist on Epidemiology and Social Research	Yoshihiko MIURA	Professor, School of Health and Social Service, Saitama Prefectural University
Person in Citizens' Position	Noboru MIYAMOTO	Principal, Koryo Junior High School, Minato-Ku
Specialist in the field of Ethics and Law	Hitomi NAKAYAMA	Lawyer, Kasumigaseki-Sogo Law Offices
The Institute of Statistical Mathematics	Toshiharu FUJITA	Professor, Department of Data Science, ISM
The Institute of Statistical Mathematics	Takashi NAKAMURA	Professor, Department of Data Science, ISM
The Institute of Statistical Mathematics	Tomoko MATSUI	Professor, Department of Statistical Modeling, ISM
The Institute of Statistical Mathematics	Shinto EGUCHI	Professor, Department of Mathematical Analysis and Statistical Inference, ISM

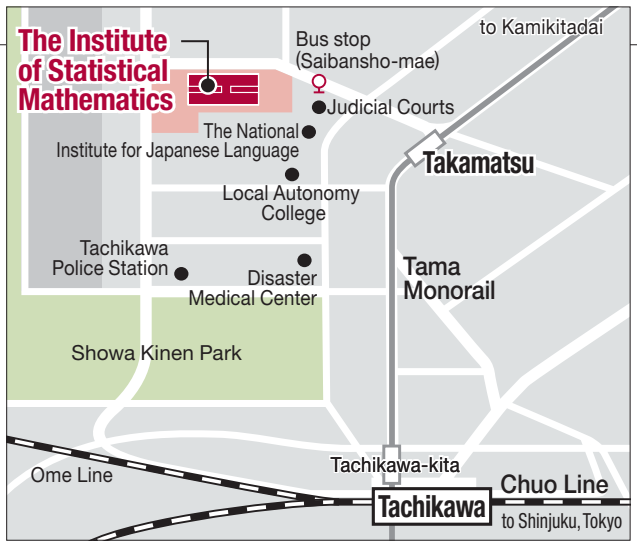
Professor Emeritus (As of October 1, 2009)

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

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.
October, 2009	The ISM was moved to 10-3 Midori-cho, Tachikawa, Tokyo.

The Institute of Statistical Mathematics



Access to the ISM

◎ Tama Monorail
ca. 7 min walk from Takamatsu Sta.

◎ Tachikawa Bus
ca. 3 min walk from Saibansho-mae bus stop

Research Organization of Information and Systems
The Institute of Statistical Mathematics
10-3 Midori-cho, Tachikawa, Tokyo 190-8562, Japan
Tel: +81-(0)50-5533-8500 Fax: +81-(0)42-527-9302 (General Affairs)
http://www.ism.ac.jp/