2026 ISI ISM ISSAS Joint Conference

February 9th-11th

2nd Conference Room 3F

Humanities and Social Sciences Building
Academia Sinica
Taipei, Taiwan

India

ISI

Indian Statistical Institute
भारतीय सांख्यकीय संस्थान

Japan

ISM

Institute of Statistical Mathematics
情報・システム研究機構
統計数理研究所

Taiwan

ISSAS

Institute of Statistical Science, Academia Sinica
中央研究院
統計科學研究所

https://www3.stat.sinica.edu.tw/3ijc2026

Content

Conference Highlight

Conference Venue

The conference will be held in the Second Conference Room of the Humanities and Social Sciences Building (HSS) at Academia Sinica. Please note that smoking is not allowed on the Academia Sinica campus. All mobile phones and electronic devices should be turned off before attending any sessions or meetings.

Registration

The registration desk and program/badge pickup are located on Level 3 of the Humanities and Social Sciences Building (HSS).

On-Campus Internet Access

Wireless internet service is available in the public areas of Academia Sinica.
Users in public spaces, such as conference rooms or other areas on campus, can use mobile devices — such as laptops and mobile phones — to access the internet.
To connect, please select the Wi-Fi network (SSID) AS_Guest. No password is required.

Conference Group Photo

The conference group photo will be taken at the Humanities and Social Sciences Building at 09:00 on February 09.

Welcome Reception

IRODORI Japanese Cuisine Buffet - 彩日本料理自助饗宴

Time: February 9 2026 18:00
Location: 3Fl., No. 2 Song Shou Rd., Taipei 110061
Website: https://www.hyatt.com/grand-hyatt/zh-HK/taigh-grand-hyatt-taipei/dining/irodori

Conference Banquet

Almost Kitchen -差不多御藏私房

Time: February 10 2026 18:00
Location: No. 183, Sec. 1, Xingzhou Rd., Sanxing Township, Yilan County 266, Taiwan
Website: https://www.facebook.com/almostkitchen

List of Sessions

<table>
<tr><td>Day 01 (02/09)</td><td>Day 02 (02/10)</td><td>Day 03 (02/11)</td></tr>
<tr><td>Registration
(08:40 – 09:00)</td><td>Registration
(08:30 – 08:50)</td><td>Registration
(08:30 – 09:00)</td></tr>
<tr><td>Opening Ceremony
(09:00 – 09:15)</td><td rowspan="3">Session 05
Poster Lightning Session
(8:50 – 11:00)</td><td rowspan="3">Session 06
Spatial Statistics
(09:00 – 10:30)</td></tr>
<tr><td>Group Photo
(09:15 – 09:30)</td></tr>
<tr><td>Session 01
Time Series and Applications
(09:30 – 10:30)</td></tr>
<tr><td>Coffee Break
(10:30 – 10:50)</td><td>Coffee Break
(11:00 – 11:20)</td><td>Coffee Break
(10:30 – 10:50)</td></tr>
<tr><td>Session 02
Machine Learning and Neural Network
(10:50 – 12:20)</td><td>Poster Session
(11:20 – 12:20)</td><td>Session 07
High-Dimensional Data Analysis
(10:50 – 12:20)</td></tr>
<tr><td>Lunch and
Administrative Meeting
(12:20 – 13:30)</td><td>Lunch
(12:20 – 13:30)</td><td>Lunch
(12:20 – 13:30)</td></tr>
<tr><td>Session 03
Recent Challenges in Probability and Statistics
(13:30 – 15:00)</td><td rowspan="3">Guided Tour
(13:30 – 18:00)</td><td>Session 08
Matrix and Multivariate Structure in Statistics
(13:30 – 15:00)</td></tr>
<tr><td>Coffee Break
(15:00 – 15:20)</td><td>Coffee Break
(15:00 – 15:20)</td></tr>
<tr><td>Session 04
Functional Data Analysis
(15:20 – 16:50)</td><td>Session 09
Distribution Shift, Data Integration and Transfer Learning
(15:20 – 16:50)</td></tr>
<tr><td>Welcome Reception
(18:00 – 20:00)</td><td>Conference Banquet
(18:00 – 20:00)</td><td>Closing Ceremony
(16:50 – 17:10)</td></tr>
</table>

Campus Map

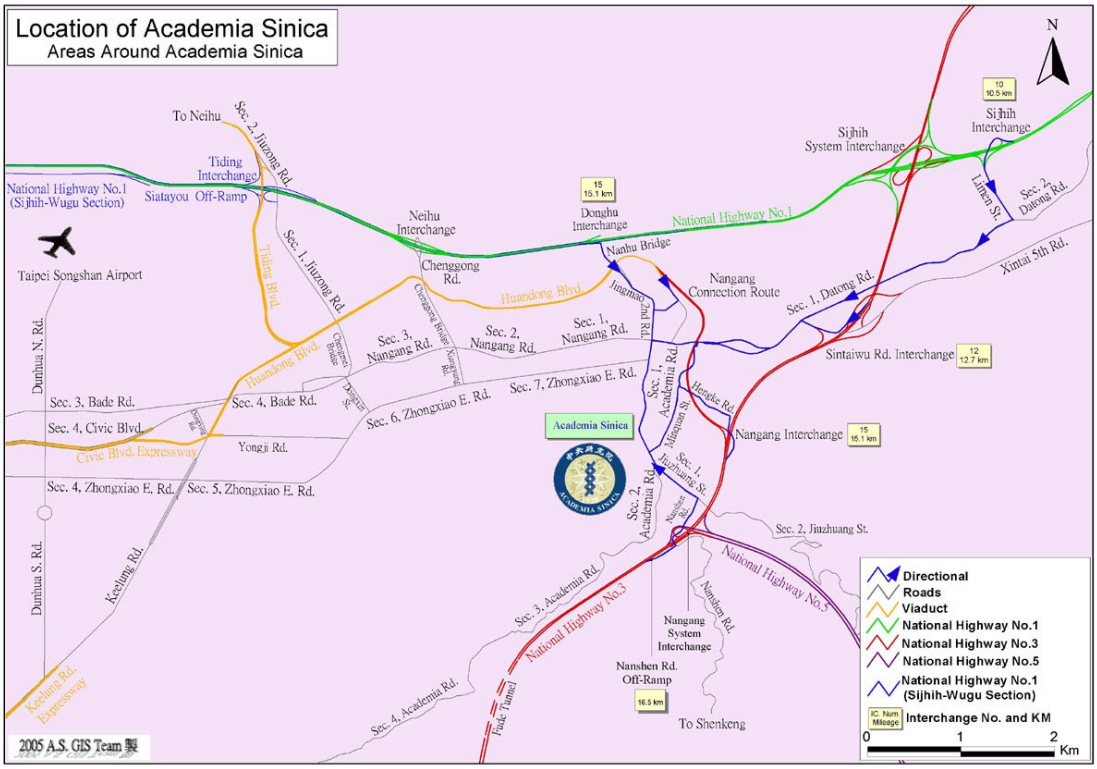
1 Main Entrance
2 Institute of Biomedical Sciences
3 Wastewater Management Facilities
4 Institute of Cellular and Organismic Biology
4 Biodiversity Research Center
5 Institute of Molecular Biology
6 Institute of Biological Chemistry
6 Life Science Library
7 National Laboratory Animal Center, NLAC
8 Interdisciplinary Research Building for Science and Technology (under construction)
9 Greenhouse
10 Central Office of Administration
11 Biodiversity Research Center
11 Biodiversity Research Museum
12 Institute of Plant and Microbial Biology
13 Research Center for Information Technology Innovation
14 Tsai Yuan-Pei Memorial Hall
14 Center for Survey Research
15 Institute of Statistical Science
16 Post office, garage, grocery store and Hi-Life convenient store
17 Ecological Pond
18 Genomics Research Center
19 Agricultural Technology Building
19 Agricultural Biotechnology Research Center (Agricultural Technology Building 1st-2nd Floor and 5-7th Floor)
19 Institute of Plant and Microbial Biology (Agricultural Technology Building 1st Floor and 3rd-4th Floor)
20 Center of Academic Activities (Bookstore, auditorium, conference rooms, guest rooms, Chinese and Western restaurants, coffee shop)
21 Institute of Chinese Literature and Philosophy
22 Institute of Earth Sciences
23 Gymnasium (Breastfeeding Room)
24 Humanities and Social Sciences Building (HSSB)
24 Joint Library of Humanities and Social Sciences (1st-2nd Floor HSSB)
24 Institute of Linguistics (5-7th Floor of South Wing, HSSB)
24 Institute of Sociology (8-10th Floor of South Wing, HSSB)
24 Research Center for Applied Sciences (11th Floor of South Wing, HSSB)
24 Research Center for Environmental Changes (11th Floor of South Wing, HSSB)
24 Institute of Political Science (Preparatory Office) (5-6th Floor of North Wing, HSSB)
24 Institute of Taiwan History (7-8th Floor of North Wing, HSSB)
24 Institutum Iurisprudentiae (Preparatory Office) (9-10th Floor of North Wing, HSSB)
25 Research Center for Environmental Changes (Laboratories)
26 Plant Molecular Breeding Greenhouse
27 Academia Sinica Greenhouse (under construction)
30 Institute of Chemistry
31 Research Center for Humanities and Social Sciences
32 Institute of Information Science
33 Institute of Physics
33 Research Center for Applied Sciences (Laboratories)
33 Wu Ta-You Memorial Hall
34 Hu Shih Memorial Hall
35 Institute of Modern History
36 Institute of European and American Studies
37 Institute of History and Philology
37 Museum of the Institute of History and Philology
38 Fu Ssu-nien Library
39 Institute of Economics
40 Institute of Ethnology
40 Museum of the Institute of Ethnology
41 Lingnan Fine Arts Museum (Jin Mei Building)
42 Archive Building, Institute of Modern History
43 Building of Taiwan Archaeological Studies
50 Taiwan International Graduate Program (TIGP) Student Dormitory
60 Teaching and Administration Building (TA Building)
60 Administrative office for TIGP, Degree Program, and Foreigners Services (2nd-3rd Floor of TA Building)
60 Kindergarten (1st Floor of TA Building)

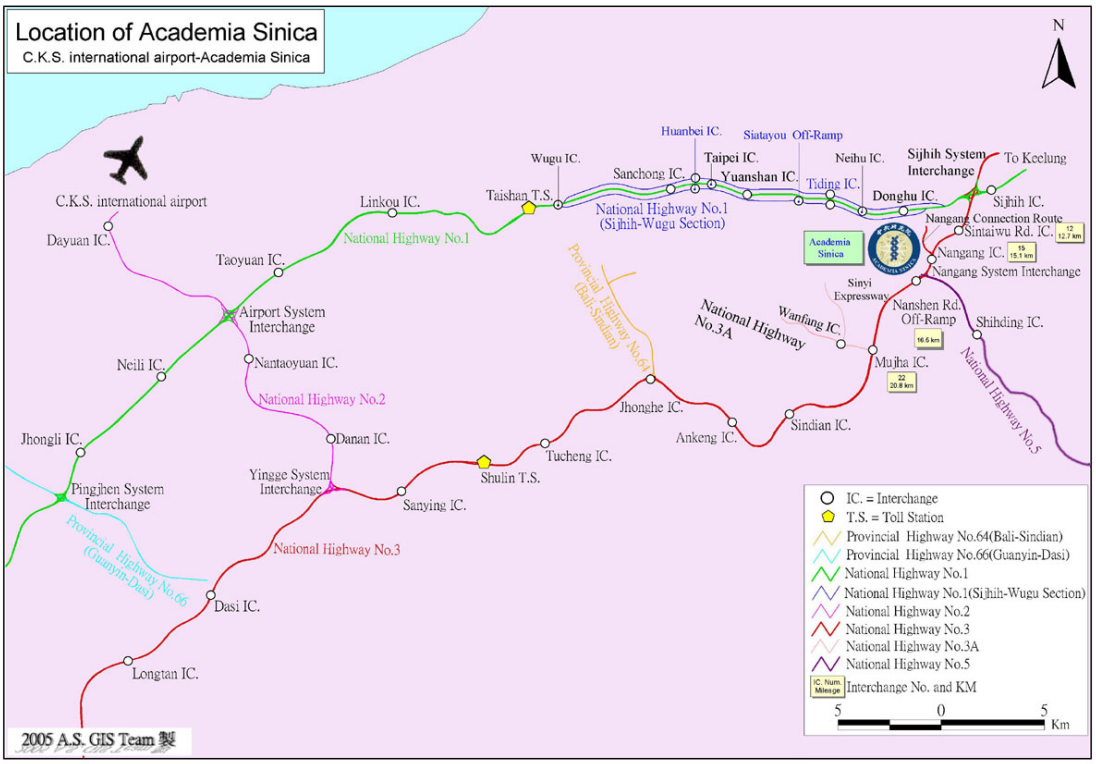
* The Institute of Mathematics, Institute of Atomic and Molecular Sciences, Institute of Astronomy and Astrophysics and some biochemistry institutes are located on the National Taiwan University campus.

Directions:

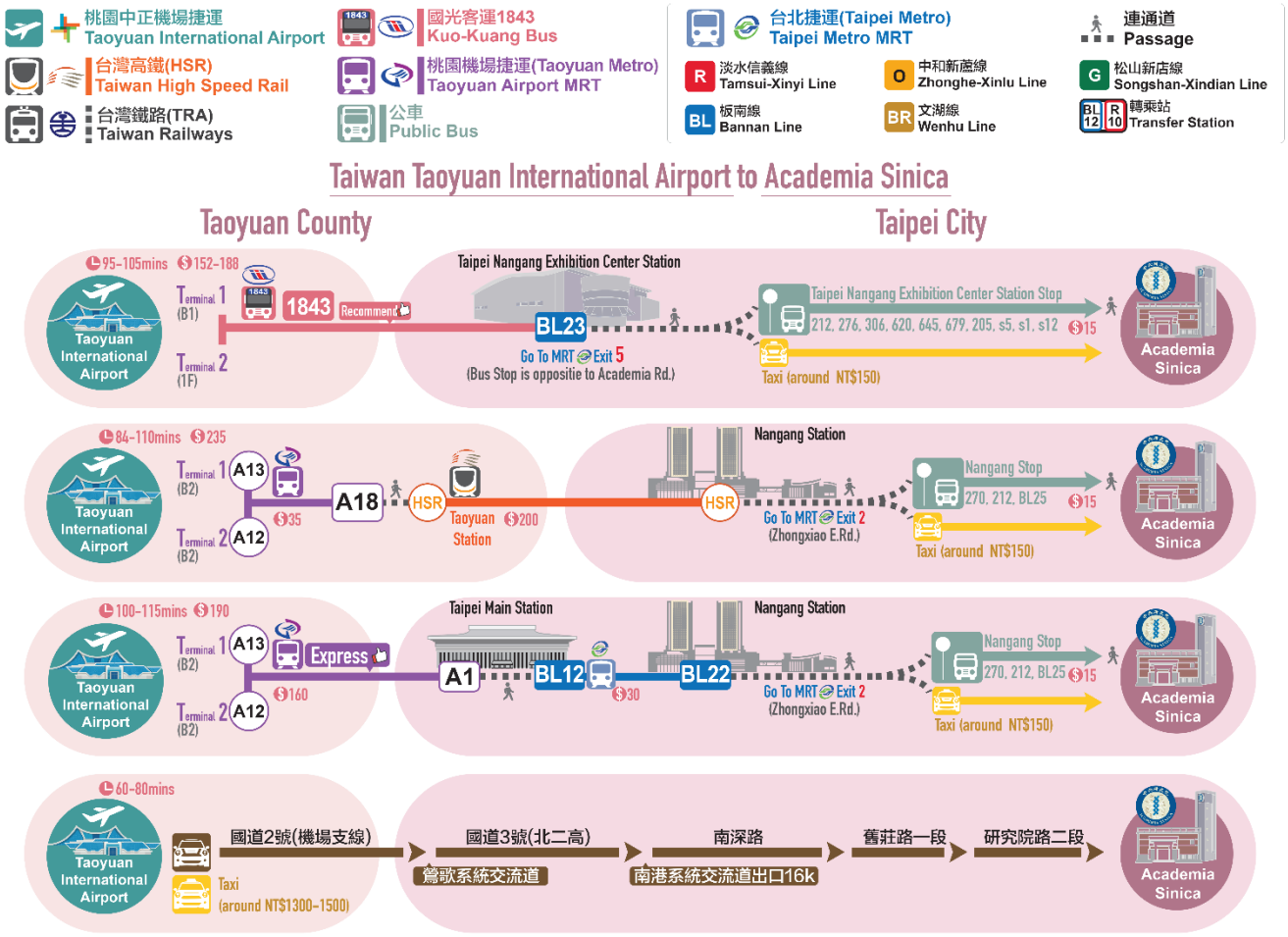
- By Bus: Bus lines 205, 212, 270, 276, 306, 620, 645 all go to Academia Sinica.
- By MRT: Take the Nankang-Banqiao-Tucheng line to Nankang Station, and then transfer to Bus 212, 270 or Blue 25 to go to Academia Sinica.
- By Train: Transfer to Bus 212, 270 or Blue 25 at Nankang Train Station to go to Academia Sinica .

Location of Academia Sinica





Transportation to Academia Sinica



A. By bus and taxi.
   1. Take the Kuo-Kuang Bus (Route 1843) to the Taipei Nangang Exhibition Center.
   2. Go to Taipei Metro MRT Bannan Line (BL23) Exit 5. The bus stop is across Academia Road from Exit 5.
   3. Take a bus or taxi to Academia Sinica. Please refer to the map for available bus routes.

B. By Taoyuan Airport MRT, Taiwan High Speed Rail and bus or taxi.
   1. Take the Taoyuan Airport MRT to Taoyuan High Speed Rail Station.
   2. Take the Taiwan High Speed Rail to Nangang Station.
   3. Go to Taipei Metro MRT Bannan Line (BL22) Exit 2. The bus stop is on Zhongxiao E. Rd.
   4. Take a bus or taxi to Academia Sinica. Please refer to the map for available bus routes.

C. By Taoyuan Airport MRT, Taipei Metro MRT, and bus or taxi.
   1. Take the Taoyuan Airport MRT to Taipei Main Station.
   2. Take the Taipei Metro MRT Bannan Line to Nangang station (BL22)
   3. Go to Taipei Metro MRT Bannan Line (BL22) Exit 2. The bus stop is on Zhongxiao E. Rd.
   4. Take a bus or taxi to Academia Sinica. Please refer to the map for available bus routes.

D. By car or taxi
   1. Take National Freeway No. 2 (Airport Spur Route) and transfer via the Yingge System Interchange to National Freeway No. 3 (Formosa Freeway).
   2. Continue via the Nangang System Interchange (Exit 16K) to Nanshen Road.
   3. Follow Section 1 of Jiuzhuang Road, then Section 2 of Yanjiuyuan Road to reach Academia Sinica.

Tour

Heritage Craft Tour & Whisky Discovery

—free half-day tour

It is our pleasure to invite you to visit the National Center for Traditional Arts and the Kavalan Whisky Distillery.

The National Center for Traditional Arts (NCFTA) in Yilan is one of Taiwan's most important cultural destinations for preserving and showcasing traditional heritage. Designed with the charm of classic Taiwanese streets and architecture, the center brings together craftsmanship, folk art, and cultural performances in an immersive environment. Visitors can explore artisan studios, browse local handicrafts, and experience hands-on workshops in various traditional crafts. With its blend of historical atmosphere and interactive cultural experiences, NCFTA offers a meaningful glimpse into Taiwan's living traditions and creative spirit.
(The information above is from https://www.px-sunmake.org.tw/)

The Kavalan Whisky Distillery, operated by the King Car Group, is internationally renowned for producing award-winning single malt whisky. Set against the pristine natural landscape of Yilan, the distillery combines pure mountain water with advanced distilling techniques to create whisky of exceptional quality. During the visit, guests can learn about the whisky-making process—from fermentation to maturation—and explore the visitor center featuring exhibitions, a whisky shop, and tasting opportunities. With its global reputation and refined flavors, Kavalan provides an unforgettable journey into Taiwan's modern craftsmanship and world-class spirits.
(The information above is from https://www.kavalanwhisky.com/zh-tw)

Date of Tour: December 10 (Tuesday), 2026

- **13:30 – 13:40 Meet at Academia Sinica · Departure**
- **14:30 – 16:20 The National Center for Traditional Arts (Map)**
- **16:50 – 17:40 The Kavalan Whisky Distillery (Map)**
- **18:00 – 20:00 Conference Banquet (Map)**
- **20:00 – 21:00 Return to Academia Sinica**

Notes:

1. Winter in northern Taiwan is often rainy, so please bring rain gear.
2. For people joining both the **Local Tour & Banquet**, take the shuttle bus departing at **13:30 from the Humanities and Social Sciences Building**.
3. For people joining **the Banquet only**, take the shuttle bus departing at **16:30 from the Institute of Statistical Science**.

Daily Schedule

Day 01 (02/09, Monday)	
08:40-09:00	Registration
09:00-09:15	Opening Ceremony
09:15-09:30	Group Photo
09:30-10:30	Session 01: Time Series and Applications
	Chair: Dr. Daisuke Murakami, Associate Professor, ISM
09:30	Analysis of Seismic Time Series Data via Statistical Tests and Imaging Techniques.
	Speaker: Dr. Frederick Kin Hing Phoa, Research Fellow, ISSAS
10:00	Asset Pricing Models Utilizing the Equi-Correlation Structure of Multivariate Volatility Models.
	Speaker: Prof. Yoshinori Kawasaki, Professor, ISM
10:30-10:50	Coffee Break
10:50-12:20	Session 02: Machine Learning and Neural Network
	Chair: Prof. Dipti Prasad Mukherjee, Professor, ISI Kolkata
10:50	Algebraic Approach to Ridge-Regularized Mean Squared Error Minimization in Minimal ReLU Neural Network.
	Speaker: Dr. Akifumi Okuno, Assistant Professor, ISM
11:20	Density Ratio Estimation with Doubly Strong Robustness.
	Speaker: Prof. Hironori Fujisawa, Professor, ISM
11:50	Constructive Universal Approximation and Sure Convergence for Multi-Layer Neural Networks.
	Speaker: Dr. Chien-Ming Chi, Assistant Research Fellow, ISSAS
12:20-13:30	Lunch and Administrative Meeting
13:30-15:00	Session 03: Recent Challenges in Probability and Statistics
	Chair: Dr. Rituparna Sen, Assistant Professor, ISI Bangalore
13:30	Robust clustered multi-task learning via non-convex group penalties.
	Speaker: Dr. Akira Okazaki, Assistant Professor, ISM
14:00	Residual Analysis in Point Process and State Space Modelling.
	Speaker: Prof. Jiancang Zhuang, Professor, ISM
14:30	Challenges in the Airspace Safety Monitoring.
	Speaker: Prof. Antar Bandyopadhyay, Professor, ISI Delhi
15:00-15:20	Coffee Break

15:20-16:50 Session 04: Functional Data Analysis

Chair: Dr. Akifumi Okuno, Assistant Professor, ISM

15:20 Separable Expansions for Covariance Estimation via the Partial Inner Product.

Speaker: Dr. Soham Sarkar, Assistant Professor, ISI Kolkata

15:50 Function-on-Function Prediction via Deep Generative Models.

Speaker: Dr. Tso-Jung Yen, Associate Research Fellow, ISSAS

16:20 Change Point Detection in Functional Time Series.

Speaker: Dr. Rituparna Sen, Associate Professor, ISI Bangalore

18:00-20:00 Welcome Reception

Day 02 (02/10, Tuesday)

08:30-08:50 Registration

08:50-11:00 Session 05: Poster Lightning Session

Chair: Dr. Junho Yang, Assistant Research Fellow, ISSAS

Group A: Statistical Theory, Design & Causal Inference

01. Chao-Hui Huang	05. Li-Sheng Zhuang
02. Sourav Chakrabarty	06. Jing-Wen Huang
03. Javed Hazarika	07. Tzu-Chi Wang
04. Kazuki Nakajima	

Group B: Computational Statistics, Machine Learning & Optimization

08. Hsin-Ping Liu	12. Po-Wei Chen
09. Wei-Chu Chiang	13. Chia-Tse Wang
10. Hirofumi Shiba	14. Jyun-Yu Chen
11. Anirban Chakraborty	

Group C: Applied Data Science Across Domains

15. Dmytro Luzhbin	19. Bing-Ru Jhou
16. Cai-Sian Liao	20. Yi-Ting Huang
17. Urmisha Chatterjee	21. Kai-Yuan Wu
18. Yu-Hsiang Lien	22. Jia-Ying Su

11:00-11:20 Coffee break

11:20-12:20 Poster Session

Group A: Statistical Theory, Design & Causal Inference

01. Chao-Hui Huang	05. Li-Sheng Zhuang
02. Sourav Chakrabarty	06. Jing-Wen Huang
03. Javed Hazarika	07. Tzu-Chi Wang
04. Kazuki Nakajima	

Group B: Computational Statistics, Machine Learning & Optimization

08. Hsin-Ping Liu	12. Po-Wei Chen
09. Wei-Chu Chiang	13. Chia-Tse Wang
10. Hirofumi Shiba	14. Jyun-Yu Chen
11. Anirban Chakraborty	

Group C: Applied Data Science Across Domains

15. Dmytro Luzhbin	19. Bing-Ru Jhou
16. Cai-Sian Liao	20. Yi-Ting Huang
17. Urmisha Chatterjee	21. Kai-Yuan Wu
18. Yu-Hsiang Lien	22. Jia-Ying Su

Not Participating in Poster Competition

23. Kengo Kamatani	24. Keishi Sando
25. Shen-Han Chiu	

12:20-13:30 Lunch

13:30-18:00 Guided Tour

18:00-20:00 Conference Banquet

Day 03 (02/11, Wednesday)

08:30-09:00 Registration

09:00-10:30 Session 06: Spatial Statistics

Chair: Prof. Antar Bandyopadhyay, Professor, ISI Delhi

09:00 Bayesian Variable Selection for Censored Spatial Responses with Application to PFAS Concentrations in California.

Speaker: Dr. Suman Majumder, Assistant Professor, ISI Kolkata

09:30 Coarse-to-Fine Spatial GLMMs for Large-Scale Modeling.

Speaker: Dr. Daisuke Murakami, Associate Professor, ISM

10:00 Spatially-Dependent Indian Buffet Processes.

Speaker: Prof. Daichi Mochihashi, Professor, ISM

10:30-10:50 Coffee Break

10:50-12:20 Session 07: High-Dimensional Data Analysis

Chair: Dr. Yen-Tsung Huang, Research Fellow, ISSAS

10:50 Variable Selection for High-Dimensional Regression Models with Higher-Order Interactions.

Speaker: Dr. Hsueh-Han Huang, Assistant Research Fellow, ISSAS

11:20 Uniform-over-Dimension Asymptotic Theory with Application to High-Dimensional Testing of Locations.

Speaker: Dr. Subhajit Dutta, Associate Professor, ISI Kolkata

11:50 High-Dimensional Two-Sample Tests Based on Optimal Transport and Related Ideas.

Speaker: Prof. Anil Kumar Ghosh, Professor, ISI Kolkata

12:20-13:30 Lunch

13:30-15:00 Session 08: Matrix and Multivariate Structure in Statistics

Chair: Dr. Akira Okazaki, Assistant Professor, ISM

13:30 Orthogonalized Moment Aberration for Multi-Stratum Factorial Designs.

Speaker: Dr. Ming-Chung Chang, Associate Research Fellow, ISSAS

14:00 Inference Based on the Generalized Spectrum for Integer-Valued Heavy-Tailed Time Series.

Speaker: Dr. Gaspard Bernard, Assistant Research Fellow, ISSAS

14:30 Extending Widely–Applicable Information Criterion: Posterior Covariance and Its Applications.

Speaker: Dr. Keisuke Yano, Associate Professor, ISM

15:00-15:20 Coffee Break

15:20-16:50 Session 09: Distribution Shift, Data Integration and Transfer Learning

Chair: Dr. Hsin-wen Chang, Associate Research Fellow, ISSAS

15:20 Computer Vision Based Compositional Zero-Shot Learning.

Speaker: Prof. Dipti Prasad Mukherjee, Professor, ISI Kolkata

15:50 Prediction of Response to Immunotherapy in Cancers via Transfer Learning.

Speaker: Dr. Shwu-Rong Grace Shieh, Research Fellow, ISSAS

16:20 Integration of Individual Participant and Aggregate Data.

Speaker: Dr. Ming-Yueh Huang, Associate Research Fellow, ISSAS

16:50-17:10 Closing Ceremony

Session 01

Time Series and Applications

[02/09 09:30 – 10:30]

Chair: **Dr. Daisuke Murakami, Associate Professor, ISM**

09:30 **Analysis of Seismic Time Series Data via Statistical Tests and Imaging Techniques.**
Speaker: Dr. Frederick Kin Hing Phoa, Research Fellow, ISSAS

10:00 **Asset Pricing Models Utilizing the Equi-Correlation Structure of Multivariate Volatility Models.**
Speaker: Prof. Yoshinori Kawasaki, Professor, ISM

S01-1 (02/09 09:30-10:00)

Analysis of Seismic Time Series Data via Statistical Tests and Imaging Techniques

Frederick Kin Hing Phoa
Institute of Statistical Science, Academia Sinica, Taipei, Taiwan

Abstract

Distributed Acoustic Sensing (DAS) has emerged as a powerful tool for monitoring geophysical hazards such as earthquakes. In this study, we propose a systematic framework for analyzing time-series DAS data recorded along an optical fiber installed beneath a coastal highway on the east coast of Taiwan, capturing seismic signals at millisecond resolution. Through algebraic transformations and visualization techniques, the raw DAS signals are first converted into image representations. Unlike conventional U-Net-based approaches, we develop a uniform segmentation strategy and further extend it to a non-uniform scheme to identify candidate partition times corresponding to seismic phase arrivals. To determine whether a detected partition corresponds to a true P-wave arrival or a spurious detection, we theoretically derive a statistical hypothesis test for distinguishing between two time series or functional data objects. Extensive simulation studies and comparisons with existing methods demonstrate that the proposed approach consistently outperforms competing techniques from seismic analysis, machine learning, and deep learning. Potential applications to seismic early warning systems are also discussed.

S01-2 (02/09 10:00-10:30)

Asset Pricing Models Utilizing the Equi-Correlation Structure of Multivariate Volatility Models.

Yoshinori Kawasaki
The Institute of Statistical Mathematics, Tokyo, Japan

Abstract

This study evaluates the performance of traditional asset pricing models in the Japanese stock market and then examines the effectiveness of new factors designed to capture correlation structures. We augment the Capital Asset Pricing Model as well as the Fama–French three-factor, Carhart four factor, and Fama–French five-factor models with two types of correlation-based factors. The first is the Industry Equi-Correlation (IEC) index, which captures the average inter-industry correlation, following Wang et al. (2020). The second is a factor derived from principal component analysis (PCA) and applied to intra-IEC (IIEC) indices, which are constructed for 33 Tokyo Stock Exchange industries using the Block DECO framework of Engle and Kelly (2012). Using 100 size-and-characteristic sorted portfolios as test assets, we evaluate 16 distinct models with Fama–MacBeth regressions, the Hansen–Jagannathan distance, and the Gibbons–Ross–Shanken test. Our results confirm that the explanatory power of traditional Fama–French factors is limited in the context of the Japanese market, whereas the IEC and PCA factors exhibit low correlations with existing factors, suggesting that they capture distinct risk sources. Most significantly, incorporating these correlation factors, particularly the principal component factor derived from IIEC, is found to consistently and substantially improve model performance across all evaluation metrics. Models that include the PCA factor demonstrate the best overall performance in our comprehensive model rankings. These results strongly suggest that inter- and intra-industry correlation structures are priced risk factors that are essential for explaining the cross-section of Japanese equity returns.

Session 02

Machine Learning and Neural Network

[02/09 10:50 – 12:20]

Chair: **Prof. Dipti Prasad Mukherjee, Professor, ISI Kolkata**

10:50 **Algebraic Approach to Ridge-Regularized Mean Squared Error Minimization in Minimal ReLU Neural Network.**
Speaker: Dr. Akifumi Okuno, Assistant Professor, ISM

11:20 **Density Ratio Estimation with Doubly Strong Robustness.**
Speaker: Prof. Hironori Fujisawa, Professor, ISM

11:50 **Constructive Universal Approximation and Sure Convergence for Multi-Layer Neural Networks.**
Speaker: Dr. Chien-Ming Chi, Assistant Research Fellow, ISSAS

Algebraic Approach to Ridge-Regularized Mean Squared Error Minimization in Minimal ReLU Neural Network

Akifumi Okuno
The Institute of Statistical Mathematics, Tokyo, Japan

Abstract

This paper investigates a perceptron, a simple neural network model, with ReLU activation and a ridge-regularized mean squared error (RR-MSE). Our approach leverages the fact that the RR-MSE for ReLU perceptron is piecewise polynomial, enabling a systematic analysis using tools from computational algebra. In particular, we develop a Divide-Enumerate-Merge strategy that exhaustively enumerates all local minima of the RR-MSE. By virtue of the algebraic formulation, our approach can identify not only the typical zero-dimensional minima (i.e., isolated points) obtained by numerical optimization, but also higher-dimensional minima (i.e., connected sets such as curves, surfaces, or hypersurfaces). Although computational algebraic methods are computationally very intensive for perceptrons of practical size, as a proof of concept, we apply the proposed approach in practice to minimal perceptrons with a few hidden units.

S02-2 (02/09 11:20-11:50)

Density Ratio Estimation with Doubly Strong Robustness

Hironori Fujisawa
The Institute of Statistical Mathematics, Tokyo, Japan

Abstract

We develop two density ratio estimation (DRE) methods with robustness to outliers. These are based on the divergence with a weight function to weaken the adverse effects of outliers. One is based on the Unnormalized Kullback-Leibler divergence, called Weighted DRE, and its optimization is a convex problem. The other is based on the γ-divergence, called γ-DRE, which improves a normalizing term problem of Weighted DRE. Its optimization is a DC (Difference of Convex functions) problem and needs more computation than a convex problem. These methods have doubly strong robustness, which means robustness to the heavy contamination of both the reference and target distributions. Numerical experiments show that our proposals are more robust than the previous methods.

Constructive Universal Approximation and Sure Convergence for Multi-Layer Neural Networks

Chien-Ming Chi
Institute of Statistical Science, Academia Sinica, Taipei, Taiwan

Abstract

We propose o1Neuro, a new neural network model built on sparse indicator activation neurons, with two key statistical properties. (1) Constructive universal approximation: At the population level, a deep o1Neuro can approximate any measurable function of $\boldsymbol{X}$, while a shallow o1Neuro suffices for additive models with two-way interaction components, including XOR and univariate terms, assuming $\boldsymbol{X} \in [0,1]^p$ has bounded density. Combined with prior work showing that a single-hidden-layer non-sparse network is a universal approximator, this highlights a trade-off between activation sparsity and network depth in approximation capability. (2) Sure convergence: At the sample level, o1Neuro's optimization reaches an optimal model with probability approaching one after sufficiently many update rounds, and we provide an example showing that the required number of updates is well bounded under linear data-generating models. Empirically, o1Neuro is compared with XGBoost, Random Forests, and TabNet for learning complex regression functions with interactions, demonstrating superior predictive performance on several benchmark datasets from OpenML and the UCI Machine Learning Repository with $n = 10{,}000$, as well as on synthetic datasets with $100 \le n \le 20{,}000$.

Session 03

Recent Challenges in Probability and Statistics

[02/09 13:30 – 15:00]

Chair: **Dr. Rituparna Sen, Assistant Professor, ISI Bangalore**

13:30 **Robust Clustered Multi-Task Learning Via Non-Convex Group Penalties.**
Speaker: Dr. Akira Okazaki, Assistant Professor, ISM

14:00 **Residual Analysis in Point Process and State Space Modelling.**
Speaker: Prof. Jiancang Zhuang, Professor, ISM

14:30 **Challenges in the Airspace Safety Monitoring.**
Speaker: Prof. Antar Bandyopadhyay, Professor, ISI Delhi

S03-1 (02/09 13:30-14:00)

Robust Clustered Multi-Task Learning Via Non-Convex Group Penalties

Akira Okazaki
The Institute of Statistical Mathematics, Tokyo, Japan

Abstract

Multi-task learning (MTL) is a methodology that aims to improve the general performance of estimation and prediction by sharing common information among related tasks. In the MTL, one of the natural assumptions is that tasks are classified into some clusters with their characteristics. MTL methods under this assumption are performed by simultaneously estimating the task parameters and clustering them. However, the assumption that all tasks are classified into some clusters is too strict, because outlier tasks that have no common information may exist in some practical situations. If the task set is contaminated by some outlier tasks, the estimation accuracy of MTL deteriorates. To overcome this problem, we propose an MTL method based on clustering with outlier parameters. The outlier parameters represent non-explained parts by center of the cluster parameters and constrained task-specific parameters. Outlier parameter vectors are selected via group sparse regularization, providing robustness of clustering against outlier tasks. We show the effectiveness of the proposed method through Monte Carlo simulations and applications to real data.

S03-2 (02/09 14:00-14:30)

Residual Analysis in Point Process and State Space Modelling

Jiancang Zhuang
The Institute of Statistical Mathematics, Tokyo, Japan

Abstract

Residual analysis provides an important and powerful tool for model diagnostics and improvement. It allows the evaluation of model fit without the need to formulate new, computationally expensive alternatives, while highlighting both the strengths and limitations of existing models. Building on recent developments in innovation-based residuals for spatial and spatiotemporal point processes, as well as extensions to state-space models, this talk outlines the general principles of innovation-based residual analysis and its role in advancing statistical modeling. Through selected applications, I illustrate how appropriately constructed residual statistics can guide model refinement and lead to more robust statistical inference across different classes of models.

S03-3 (02/09 14:30-15:00)

Challenges in the Airspace Safety Monitoring

Antar Bandyopadhyay
Indian Statistical Institute, New Delhi, India

Abstract

In this talk we will describe the Safety Analysis of the Indian Oceanic Airspace which we are conducting jointly with the Airports Authority of India (AAI), starting from the year 2011. The talk will concentrate on the understanding of the goal of this work, typical data structure and the current "state of the art". We will also discuss some of the unconventional statistical challenges which we face for analysis of such data.

Session 04

Functional Data Analysis

[02/09 15:20 – 16:50]

Chair: **Dr. Akifumi Okuno, Assistant Professor, ISM**

15:20 **Separable Expansions for Covariance Estimation via the Partial Inner Product.**
Speaker: Dr. Soham Sarkar, Assistant Professor, ISI Kolkata

15:50 **Function-on-Function Prediction via Deep Generative Models.**
Speaker: Dr. Tso-Jung Yen, Associate Research Fellow, ISSAS

16:20 **Change Point Detection in Functional Time Series.**
Speaker: Dr. Rituparna Sen, Associate Professor, ISI Bangalore

S04-1 (02/09 15:20-15:50)

Separable Expansions for Covariance Estimation via the Partial Inner Product

Soham Sarkar
Indian Statistical Institute, Kolkata, India

Abstract

The non-parametric estimation of covariance lies at the heart of functional data analysis, whether for curve or surface-valued data. The case of a two-dimensional domain poses both statistical and computational challenges, which are typically alleviated by assuming separability. However, separability is often questionable, sometimes even demonstrably inadequate. We propose a framework for the analysis of covariance operators of random surfaces that generalises separability, while retaining its major advantages. Our approach is based on the expansion of the covariance into a series of separable terms. The expansion is valid for any covariance over a two-dimensional domain. Leveraging the key notion of the partial inner product, we generalise the power iteration method to general Hilbert spaces and show how the aforementioned expansion can be efficiently constructed in practice at the level of the surface observations. Truncation of the expansion and retention of the leading terms automatically induces a non-parametric estimator of the covariance, whose parsimony is dictated by the truncation level. The resulting estimator can be calculated, stored and manipulated with little computational overhead relative to separability. Consistency and rates of convergence are derived under mild regularity assumptions, illustrating the trade-off between bias and variance regulated by the truncation level. The merits and practical performance of the proposed methodology are demonstrated in a comprehensive simulation study. [Joint work with Tomas Masak and Victor M. Panaretos]

S04-2 (02/09 15:50-16:20)

Function-on-Function Prediction via Deep Generative Models

Tso-Jung Yen
Institute of Statistical Science, Academia Sinica, Taipei, Taiwan

Abstract

Function-on-function prediction involves using one sequence to predict another sequence. It is a problem commonly seen in many scientific fields. Its model is trained on paired sequence data. However, sequences may have different lengths, observed at irregular spaces and at different locations. These irregularities make model training difficult to proceed. In this talk, we present a graph-based method for training deep generative models to tackle function-on-function prediction problems. This method addresses a common challenge in training data, where functional sequences are only partially observed at irregular locations. Under this method, a functional sequence is represented as a graph in which each node corresponds to a location–value pair, and each link is defined in terms of the distance between locations of two nodes. This formulation allows training data to have functional sequences with different lengths and observed at different locations. Simulation results show that deep generative models trained under our method outperform the ground-truth model when only incomplete observations are available. This work is joint with Chia-Tse Wang, Ming-Chung Chang, Su-Yun Huang, and Tailen Hsing.

S04-3 (02/09 16:20-16:50)

Change Point Detection in Functional Time Series

Rituparna Sen
Indian Statistical Institute, Bengaluru, India

Abstract

We consider functional autoregressive processes observed on a sparse grid with noise. Using robust Kalman filtering and blocked Gibbs sampling, we fit a state-space model. The interest is in detection of change point in any of the three parameters, the mean, the variance and the autocovariance operator. Identifying the posterior mode as the change point, we obtain very promising results in simulations under several situations. We apply the method to obtain change points in daily patterns of high frequency financial data.

Session 06

Spatial Statistics

[02/11 09:00 – 10:30]

Chair: **Prof. Antar Bandyopadhyay, Professor, ISI Delhi**

09:00 **Selection for Censored Spatial Responses with Application to PFAS Concentrations in California.**
Speaker: Dr. Suman Majumder, Assistant Professor, ISI Kolkata

09:30 **Coarse-to-Fine Spatial GLMMs for Large-Scale Modeling.**
Speaker: Dr. Daisuke Murakami, Associate Professor, ISM

10:00 **Spatially-Dependent Indian Buffet Processes.**
Speaker: Prof. Daichi Mochihashi, Professor, ISM

S06-1 (02/11 09:00-09:30)

Bayesian Variable Selection for Censored Spatial Responses with Application to PFAS Concentrations in California

Suman Majumder
Indian Statistical Institute, Kolkata, India

Abstract

Per- and polyfluoroalkyl substances (PFAS) are persistent environmental contaminants of major public health concern due to their resistance to degradation, widespread occurrence, and potential adverse health effects. Statistical analysis of PFAS concentrations in groundwater is complicated by left-censoring from detection limits, strong spatial dependence, and high-dimensional covariates. While PFAS levels are believed to be influenced by diverse sociodemographic, industrial, and environmental factors, their relative contributions remain unclear, motivating rigorous statistical approaches that can isolate key predictors from a large candidate set. We develop a Bayesian hierarchical framework that embeds censoring within a spatial process model via approximate Gaussian processes and employs a global–local shrinkage prior for effective high-dimensional variable selection. To further refine inference, we compare three post-selection strategies, credible interval rules, shrinkage weight thresholding, and clustering-based inclusion, based on their predictive accuracy, robustness to censoring, and stability of variable inclusion. Applying this framework to PFOS concentrations in California groundwater, we identify a parsimonious yet scientifically meaningful set of drivers. Significant covariates include demographic factors (percentage of American Indian population, local gender composition), industrial sources (metal coating, textile and leather, oil and gas, electronics, chemical manufacturing and cement manufacturing facilities), proximity to airports, traffic density, and environmental features such as herbaceous cover, elevation and ozone concentration. Our findings show that the proposed Bayesian variable selection framework provides interpretable and stable inference in high-dimensional censored spatial settings, while simultaneously offering actionable insights into the factors driving PFAS contamination in groundwater.

S06-2 (02/11 09:30-10:00)

Coarse-to-Fine Spatial GLMMs for Large-Scale Modeling

Daisuke Murakami
The Institute of Statistical Mathematics, Tokyo, Japan

Abstract

Although a recent study showed that a coarse-to-fine modeling approach provides a fast and flexible alternative for large-scale spatial process modeling, the method was originally developed for Gaussian responses, limiting its applicability. To address this limitation, we extend the coarse-to-fine approach to spatial generalized linear mixed models (GLMMs), enabling the analysis of count, binary, and other non-Gaussian responses. The performance of the proposed spatial GLMMs is evaluated through Monte Carlo experiments, demonstrating its predictive accuracy and computational efficiency. Finally, we apply the proposed method to an empirical study.

S06-3 (02/11 10:00-10:30)

Spatially-Dependent Indian Buffet Processes

Daichi Mochihashi
The Institute of Statistical Mathematics, Tokyo, Japan

Abstract

We develop a new stochastic process called spatially-dependent Indian buffet processes (sIBP) for binary feature matrices of unbounded columns with spatial correlations between subjects, and propose general spatial factor models for various multivariate response variables. We introduce spatial dependency through the stick-breaking representation of the original Indian buffet process (IBP) (Griffiths and Ghahramani, 2005, 2011) and latent Gaussian process for the logit-transformed breaking proportion to capture underlying spatial correlation. We show that the marginal limiting properties of the number of non-zero entries under SIBP are the same as those in the original IBP, while the joint probability is affected by the spatial correlation. Using binomial expansion and Pólya-gamma data augmentation, we provide a novel Gibbs sampler for posterior computation. The usefulness of our SIBP is demonstrated through simulation studies and two applications for large-dimensional multinomial data of areal dialects and geographical distribution of multiple tree species.

Session 07

High-Dimensional Data Analysis

[02/11 10:50 – 12:20]

Chair: Dr. Yen-Tsung Huang, Research Fellow, ISSAS

10:50 Variable Selection for High-Dimensional Regression Models with Higher-Order Interactions.

Speaker: Dr. Hsueh-Han Huang, Assistant Research Fellow, ISSAS

11:20 Uniform-over-Dimension Asymptotic Theory with Application to High-Dimensional Testing of Locations.

Speaker: Dr. Subhajit Dutta, Associate Professor, ISI Kolkata

11:50 High-Dimensional Two-Sample Tests Based on Optimal Transport and Related Ideas.

Speaker: Prof. Anil Kumar Ghosh, Professor, ISI Kolkata

S07-1 (02/11 10:50-11:20)

Variable Selection for High-Dimensional Regression Models with Higher-Order Interactions

Hsueh-Han Huang
Institute of Statistical Science, Academia Sinica, Taipei, Taiwan

Abstract

This work proposes the network orthogonal greedy algorithm (Network OGA), an efficient method designed to capture higher-order (beyond second-order) interactions. By integrating the concepts of ranking and stepwise forward regression, Network OGA leverages the advantages of both approaches. The algorithm is applicable to high-dimensional interaction models of arbitrary unknown orders. We establish the sure screening property for Network OGA and demonstrate that, when coupled with a high-dimensional information criterion (HDIC), the method achieves variable selection consistency. Simulation studies further validate its superior performance.

.

S07-2 (02/11 11:20-11:50)

Uniform-over-Dimension Asymptotic Theory with Application to High-Dimensional Testing of Locations

Subhajit Dutta
Indian Statistical Institute, Kolkata, India

Abstract

Asymptotic methods for hypothesis testing in high-dimensional data usually require the dimension of the observations to increase to infinity, often with an additional relationship between the dimension (say, p) and the sample size (say, n). On the other hand, multivariate asymptotic testing methods are valid for fixed dimension only and their implementations typically require the sample size to be large compared to the dimension to yield desirable results. In practical scenarios, it is usually not possible to determine whether the dimension of the data conforms to the conditions required for the validity of the high-dimensional asymptotic methods for hypothesis testing, or whether the sample size is large enough compared to the dimension of the data. In this work, we first describe the notion of uniform-over-p convergences and subsequently, develop a uniform-over-dimension central limit theorem. An asymptotic test for the two-sample equality of locations is developed, which now holds uniformly over the dimension of the observations. Using simulated and real data, it is demonstrated that the proposed test exhibits better performance compared to several popular tests in the literature for high-dimensional data as well as the usual scaled two-sample tests for multivariate data, including the Hotelling's T^2 test for multivariate Gaussian data.

S07-3 (02/11 11:50-12:20)

High-Dimensional Two-Sample Tests Based on Optimal Transport and Related Ideas

Anil Kumar Ghosh
Indian Statistical Institute, Kolkata, India

Abstract

In recent years, the idea of optimal transport has become popular in statistics, mainly for constructing distribution-free tests for high-dimensional data. The idea is to transport the observations to a known reference distribution and construct tests based on the transported data. Under suitable regularity conditions, these resulting tests usually have large sample consistency in finite dimensions. But they often fail to yield satisfactory performance for high-dimensional data, especially when the dimension is comparable to or larger than the sample size. In this article, we first investigate the high-dimensional behaviour of some two-sample tests based on optimal transport and show that a judicious choice of reference distribution and transportation cost may lead to a better performance in high dimensions. Our theoretical investigation also leads to the construction of a class of distribution-free two-sample tests based on the idea of minimum cost derangement. Interestingly, some popular distribution-free two-sample tests belong to this class. Several simulated and benchmark data sets are analysed to study the empirical performance of our proposed test in high-dimension, low-sample-size situations.

(Joint work with Vaibhab Sherkar, Abhradipta Ghosh and Bilol Banerjee)

Session 08

Matrix and Multivariate Structure in Statistics

[02/11 13:30 – 15:00]

Chair: **Dr. Akira Okazaki, Assistant Professor, ISM**

13:30 **Orthogonalized Moment Aberration for Multi-Stratum Factorial Designs.**
Speaker: Dr. Ming-Chung Chang, Associate Research Fellow, ISSAS

14:00 **Inference Based on the Generalized Spectrum for Integer-Valued Heavy-Tailed Time Series.**
Speaker: Dr. Gaspard Bernard, Assistant Research Fellow, ISSAS

14:30 **Extending Widely–Applicable Information Criterion: Posterior Covariance and Its Applications.**
Speaker: Dr. Keisuke Yano, Associate Professor, ISM

S08-1 (02/11 13:30-14:00)

Orthogonalized Moment Aberration for Multi-Stratum Factorial Designs

Ming-Chung Chang
Institute of Statistical Science, Academia Sinica, Taipei, Taiwan

Abstract

Multi-stratum factorial designs, such as block designs and row–column designs, are widely used for screening treatment factors in experiments with complex experimental-unit structures arising from multiple sources of variability. In this presentation, I will introduce a unified, model-free approach—termed orthogonalized moment aberration—for comparing similarities among level combinations of treatment factors assigned to heterogeneous experimental units. The proposed approach evaluates the rows of design matrices through kernel functions, rather than the columns, enabling the assessment of a broad class of mixed-level regular and nonregular factorial designs under heterogeneous experimental-unit structures known as partially relaxed orthogonal block structures. This framework is highly flexible: different choices of kernel functions allow adaptation to various experimental scenarios, with certain choices recovering well-known minimum aberration criteria from the literature. Although model-free in nature, the proposed method admits rigorous justification via linear mixed-effects models and Gaussian process models. Theoretical results and numerical examples demonstrate that this approach can generate multi-stratum factorial designs with high Bayesian D-efficiencies.

S08-2 (02/11 14:00-14:30)

Inference based on the generalized spectrum for integer-valued heavy-tailed time series

Gaspard Bernard
Institute of Statistical Science, Academia Sinica, Taipei, Taiwan

Abstract

Integer-valued time series models are useful for representing count data that may exhibit bubble-like phenomena with sudden bursts. In this talk, we consider the integer-valued Moving Average and Autoregressive models with stable innovations and address the issue of parameter estimation. Due to the presence of heavy tails, classical spectrum-based inference is not applicable in this context, and it is necessary to develop new methods. We propose an approach based on the generalized spectrum obtained using characteristic functions, allowing root-n consistent parameter estimation without any moment assumption. We explore the asymptotic properties of the proposed estimators, as well as their finite sample performance (the latter through simulation studies).

S08-3 (02/11 14:30-15:00)

Extending Widely–Applicable Information Criterion: Posterior Covariance and Its Applications

Keisuke Yano
The Institute of Statistical Mathematics, Tokyo, Japan

Abstract

In this talk, we present a novel computationally low-cost method for estimating a general predictive measure of generalized Bayesian inference. The proposed method utilizes posterior covariance and provides estimators of the Gibbs and the plugin generalization errors. We present theoretical guarantees of the proposed method, linking it to Bayesian sensitivity analysis and the infinitesimal jackknife approximation of Bayesian leave-one-out cross validation. We demonstrate that the proposed method achieves accurate, stable generalization error estimates, across a range of settings, including differentially private learning, hierarchical models, regression with influential observations, and models with strong priors, and remains effective in high-dimensional settings.

-Yukito Iba and Keisuke Yano, Posterior Covariance Information Criterion for general loss functions, Bayesian Analysis, 2025 (DOI:https://doi.org/10.1214/25-BA1536)
-Yukito Iba and Keisuke Yano, Posterior Covariance Information Criterion for Weighted Inference, Neural Computation, vol. 35, 1340–1361, 2023 (DOI: https://doi.org/10.1162/neco_a_01592)

Session 09

Distribution Shift, Data Integration and Transfer Learning

[02/11 15:20 – 16:50]

Chair: **Dr. Hsin-wen Chang, Associate Research Fellow, ISSAS**

15:20 **Computer Vision Based Compositional Zero-Shot Learning.**
Speaker: Prof. Dipti Prasad Mukherjee, Professor, ISI Kolkata

15:50 **Prediction of Response to Immunotherapy in Cancers via Transfer Learning.**
Speaker: Dr. Shwu-Rong Grace Shieh, Research Fellow, ISSAS

16:20 **Integration of Individual Participant and Aggregate Data.**
Speaker: Dr. Ming-Yueh Huang, Associate Research Fellow, ISSAS

S09-1 (02/11 15:20-15:50)

Computer Vision Based Compositional Zero-Shot Learning

Dipti Prasad Mukherjee
Indian Statistical Institute, Kolkata, India

Abstract

Computer vision based machine learning (CVML) for zero-shot compositions recognizes composite concepts. For example, CVML tries to recognize a rotten banana when the CVML system has seen ripe banana and rotten apple during training. It is zero-shot learning as rotten banana was not seen by the CVML during training. It is learning of compositions as state rotten and object banana describe a new concept unknown during training while individual parts (like, state rotten and object banana) are already seen by the learning model. Naturally, the challenge during training of compositions is to disentangle features of state from features of objects. The better disentanglement of features of state and object during training results in better description of previously unseen compositions during inference. The state features also influence the definition of compositions. For example, in the compositions peeled apple and peeled orange, the state peeled has different visual connotations. This context dependency of state features poses additional challenge to feature disentanglement. Therefore, one of the key research questions in CVML is to measure the quality and extent of disentanglement of state features from features of object. We propose to discuss the CVML architecture and possible disentanglement measures.

S09-2 (02/11 15:50-16:20)

Prediction of Response to Immunotherapy in Cancers via Transfer Learning

Shwu-Rong Grace Shieh
Institute of Statistical Science, Academia Sinica, Taipei, Taiwan

Abstract

To date, immunotherapies such as immune checkpoint inhibitors (ICIs) have emerged as a leading treatment for metastatic cancer, significantly improving patient survival while causing relatively few side effects. However, the objective response rate for ICIs remains low, approximately 25% for metastatic urothelial carcinoma (mUC) and renal cell carcinoma (mRCC), and ~40% for melanoma, underscoring the urgent need for predictive response biomarkers. Several state-of-the-art signatures have been revealed in top-tier journals, highlighting the importance of this field. As the number of genes (~20,000) far exceeds the sample sizes of typical training sets (generally ≤ 300), we first developed feature selection procedures to reduce the number of features to a few hundred. We then trained multiple machine learning classifiers using the selected genes and the IMvigor210, IMmotion150, and Gide cohorts, which includes RNA-seq and clinical data from 298, 77, and 91 patients with mUC, mRCC, and melanoma, respectively. Notably, our predictor LogitDA using the identified gene signatures achieved a prediction AUC of 0.75, 0.83, and 0.71~0.75 in independent cohorts, PCD4989g (mUC, mRCC), and three melanoma cohorts, respectively. Moreover, our signatures outperformed (most of) six state-of-the-art signatures, PD-L1 IHC, and five tumor microenvironment signatures, including IFN-γ, T-effector, and T-cell exhaustion signatures in mUC (mRCC and melanoma). From our signatures, we identified key prognostic biomarkers in mUC, mRCC, and melanoma, respectively.

S09-3 (02/11 16:20-16:50)

Integration of Individual Participant and Aggregate Data

Ming-Yueh Huang
Institute of Statistical Science, Academia Sinica, Taipei, Taiwan

Abstract

Integrated IPD–AD analysis, which combines individual participant data with aggregate data, is widely used for synthesizing evidence across heterogeneous studies. This talk examines how the form of aggregate data affects integration efficiency, an issue that has received less attention than algorithmic development. Using a constrained maximum likelihood framework, I show that subgroup-specific aggregate summaries substantially improve estimation efficiency, with outcome-stratified summaries consistently outperforming covariate-stratified ones, especially for continuous outcomes. Although outcome-stratified summaries are standard for discrete outcomes, they are rarely reported for continuous endpoints; our results suggest that routinely providing such summaries could meaningfully enhance evidence synthesis. I further extend the framework to accommodate dataset shift and propose a fast, non-iterative estimation procedure, illustrated with applications to income data under covariate shift and housing data under prior probability shift.