

Workshop on Research Metrics

Date 15 - 16 November 2019
Time 9am to 5 pm
Registration https://docs.google.com/forms/d/e/1FAIpQLScPsIsNeTIFr0-XDZXeaxjzbx48m9lDvbonTzd5c8nFr-5mw/viewform?usp=sf_link
Venue Science Library – Meeting Room, Block S6, Level 4, 10 Science Drive 2, Singapore 117548
 More information on getting there can be found at this website - <https://libportal.nus.edu.sg/frontend/ms/science-library/about-science-library/getting-to-science-library>

Friday 15 November 2019	
9:00 – 9:30	Registration
9:30-9:40	Opening by Mrs LEE Cheng Ean, NUS University Librarian
9:40-10:20	Development of Research Diversity Index and Its System Keisuke Honda (The Institute of Statistical Mathematics, Japan)
10:20-11:00	Time-Delayed Effects Among the Grant and Research Performance of University Faculty Ray-Bing Chen (National Cheng Kung University, Taiwan)
11:00-11:30	Tea Break
11:30 – 12:10	Research Evaluation in the Social Sciences and Humanities: Recent Trends in Europe and Japan Keiichi Osiumi (Ryukyu University, Japan)
12:10 – 13:30	Lunch break and Discussion
13:30-14:10	Understanding Themes and Trends in Postsecondary Research Using Topic Modeling and Journal Abstracts Mio Takei, Stephan Porter and Paul Umbach, (The Institute of Statistical Mathematics, Japan)
14:10-14:50	Is Scientific Performance a Function of Funds? Alona Zharova (Humboldt-Universität zu Berlin, Germany)
14:50-15:30	Quantification of Economic Uncertainty: An Application of Deep Learning

	Alisa Kim (Humboldt-Universität zu Berlin, Germany)
15:30-16:00	Tea break
16:00 – 16:40	A Dynamic Study on the Evolution of the Structure and Clusters of Coauthorship Network from the Web of Science Frederick Phoa (Academia Sinica, Taiwan)
16:00-16:40	Statistical modeling of the article citation network in Statistics Junji Nakano (The Institute of Statistical Mathematics and Cuo University, Japan)
18:00	Dinner (By Invitation)
Saturday 16 November 2019	
	Discussion

Workshop Organizers:

Ying Chen, Department of Mathematics, Risk Management Institute, National University of Singapore

Keisuke Honda, The Institute of Statistical Mathematics, Japan

Thorsten Koch, Zuse Institute Berlin, Germany

Workshop Coordinator:

Nazgul Zakiyeva, Department of Statistics and Applied Probability, National University of Singapore

* The workshop is jointly organized by the Department of Mathematics at National University of Singapore (NUS), University Library at NUS, and The Institute of Statistical Mathematics, Japan and Zuse Institute Berlin, Germany

Development of research diversity index and its system

Keisuke Honda

The Institute of Statistical Mathematics, Tokyo, Japan

Abstract

Research organizations, along with their funding agencies, routinely need to measure the performance of their organization. However, for funding agency or grant review panels which are supporting seeds research, must make difficult predictions about the likelihood of future scientific success. In this study, we propose a new citation index to see structure of relationship among of various areas of science. This index is based on stochastic block model (SBM) on big scale citation network data. SBM generates a matrix which divides several blocks which represent relationship among research fields. We introduce implemented web system which is designed as analysis tool.

Time-Delayed Effects among the Grant and Research Performance of University Faculty

Ray-Bing Chen

*Department of Statistics & Institute of Data Science
National Cheng Kung University*

Abstract

For the faculty in a university in Taiwan, the most important duties are teaching and research. Every semester, different teaching evolution would be conducted to show the teaching quality of faculty and such measurements can provide some guidelines for each faculty to improve his/her teaching technique. In terms of the measurements of the research performances, we usually use the number of the publications and the number of citations per year of each faculty. However, these simple measurements may be limited in reflecting the true research performance of faculty. In this study, we apply more variables to explain the relationships among the factors regarding research outputs of faculty as follow: faculty age, grant size, the number of the publications, the number of PhD students and so on. Due to the different hypotheses, the PVARX or its generalized model is used to build the relationships among these factors. The data from a top-tier university in Taiwan is used for the illustration. Based on our analysis results, the time-delayed effects for the different factors are identified.

Research Evaluation in the Social Sciences and Humanities: Recent Trends in Europe and Japan

Keiichi Oshiumi

University of the Ryukyus, Okinawa, Japan

Abstract

Although bibliometric research evaluation in Science, Technology, Engineering and Mathematics (STEM) fields is accepted to a certain extent, it is more problematic in Social Sciences and Humanities (SSH) fields. In my presentation, I explain the differences between STEM and SSH related to research evaluation and the problems of evaluating SSH by using bibliometrics. Then, I introduce some activities in Europe and Japan.

Understanding themes and trends in postsecondary research using topic modeling and journal abstracts

Mio Takei¹, Stephen R. Porter², Paul D. Umbach² and Junji Nakano³

¹ *Institute of the Statistical Mathematics, Japan*

² *North Carolina State University*

³ *Chuo University, Japan*

Abstract

Investigating trends in research activities is important for understanding our current position and future trends. Therefore, we analyze abstracts for three main postsecondary journals from the past three decades to understand themes and trends of postsecondary research. We use topic models, especially Hierarchical Dirichlet Process, to extract topics and display words using network graph layout to interpret the topics. Moreover, we extend the result to reveal time variation of topics.

Is Scientific Performance a Function of Funds?

Alona Zharova

Humboldt-Universität zu Berlin, Germany

Abstract

The management of universities requires data on teaching and research performance. While teaching quality can be measured via student performance and teacher evaluation programs, the connection of research output and its antecedents is much harder to test, check and understand. To inform research governance and policy making at universities, we clarify the relationship between grant money and research performance. We examine the interdependence structure between third-party expenses, publications, citations and academic age. To describe the relationship between these factors, we analyze individual-level data from a sample of professorships and a Scopus database for the period 2001 to 2015. Using estimates from a panel vector autoregressive model with exogenous variables, impulse response functions and forecast error variance decomposition, we show that decision making based on the university-level data is inappropriate and does not reflect the behavior of individual faculties. Our results quantify the difference between the quality and quantity of research output, a better understanding of which is important to design incentive schemes and promotion programs. The paper also proposes a visualization of the cooperation between faculties and research interdisciplinarity via the co-authorship structure among publications. We discuss the implications for policy and decision making and make recommendations for university research management.

QUANTIFICATION OF ECONOMIC UNCERTAINTY: AN APPLICATION OF DL

Alisa Kim¹ and Niels Gillmann²

¹ *Humboldt-Universität zu Berlin, Germany*

² *Dresden University of Technology*

Abstract

The creation of the Economic Policy Uncertainty Index in 2016 started a new wave of research on the impact of uncertainty on macroeconomic variables like production, consumption, and investment in the United States. The original index was constructed based on major USA media publications with the use of manual labeling and word count. Several attempts of automating this procedure have been undertaken since, using Support Vector Machines (less accurate but automated labeling) and LDA analysis (replacing the word count with topic modeling). The current paper takes these efforts one step further and offers an algorithm for the quantification of economic policy uncertainty based on natural language processing technics and deep learning. The new approach allows the accurate

distillation of the latent “uncertainty” underlying newspaper articles, enables an automated construction of a new EPU index and demonstrates pronounced explanatory powers when compared to established benchmarks like the previous EPU Index, the VIX and other measures of uncertainty like a disagreement between survey participants or dispersion of forecast densities. The potential applications extend to the areas of state policy evaluation, business cycle analysis, financial forecasting and, potentially, derivative pricing.

A Dynamic Study on the Evolution of the Structure and Clusters of Coauthorship Network from the Web of Science

Frederick Phoa

*Institute of Information Science,
Academia Sinica,
Taipei, Taiwan*

Abstract

Coauthorship is common in scientific researches as at least two researchers work together on a collaborative project. When these authors are connected when they write a scientific article together, a large-scale network called coauthorship network is formed. The structure and property of this network reveals many interesting phenomena in scientific researches. This work aims at studying the structure of the large-scale coauthorship network over the past 36 years. In specific, we look into the properties of the network structure, and applied community detection techniques to cluster the whole network into multiple clusters of interest. We also observe how the network is grown over time and how the clusters evolve in this period. The coauthorship network data is extracted from the Web of Science. This is a joint work with Ms. Livia Lin-Hsuan Chang (Institute of Statistical Mathematics, Japan) and the Phoa research group (Academia Sinica, Taiwan).

Statistical modeling of the article citation network in Statistics

Livia Lin-Hsuan Chang¹, Junji Nakano², Frederick Kin Hing Phoa³

¹ *SOKENDAI (The Graduate University for Advanced Studies), Japan*

² *Chuo University, Japan*

³ *Academia Sinica, Taiwan*

Abstract

Citation networks of scientific papers are used for Institutional Research purposes. They have some specific features. First, there is a special time structure, i.e., each article has a published year, and it is only possible that the newer published articles can cite the earlier published articles and it is impossible to be reversed. Second, all the articles have a similar life cycle, i.e. the citation number received by other articles will reach to the peak after a few years of the publication, and it will gradually decrease afterward. Third, each article may contain a latent variable "importance" on which the citation probability will be based. We are interested in the modeling of the generating process of citation network. In this work, we focus on the article citation network in statistics research community which is extracted from Web of Science database. Based on the specified features above, we developed a generative model of the article citation network. Simulated networks from the generative model are compared to the real article citation network by checking their degree distributions.