

CONTACT INFORMATION Department of Statistics (MC0439)
Hutcheson Hall, Room 417-C,
Virginia Tech
250 Drillfield Drive
Blacksburg, VA 24061 sengupta@vt.edu
publish.illinois.edu/srijan-sengupta

EXPERIENCE **Virginia Tech**, Blacksburg, VA
Assistant Professor of Statistics (tenure-track) Fall 2016 - Now
Max New York Life, Gurgaon, India
Product Pricing & Risk Management Actuary 2010 - 2011
ICICI Prudential, Mumbai, India
Risk Management Actuary 2009 - 2010

EDUCATION **University of Illinois at Urbana-Champaign, IL**
Ph.D. in Statistics 2011 - 2016
(Norton prize for outstanding PhD Thesis)
Co-advisors: Prof. Yuguo Chen and Prof. Xiaofeng Shao
Dissertation title: Statistical analysis of networks with community structure
and bootstrap methods for big data
Indian Statistical Institute, Kolkata, India
M.Stat (1st Division with Distinction) 2007 - 2009
Specialization: Mathematical Statistics and Probability
Indian Statistical Institute, Kolkata, India
B.Stat (1st Division with Distinction) 2004 - 2007

PUBLICATIONS

1. **Sengupta, S.** and Chen, Y. (2016+). A blockmodel for node popularity in networks with community structure. Invited revision, *Journal of the Royal Statistical Society: Series B*.
2. **Sengupta, S.**, Volgushev, S., and Shao, X. (2016). A subsampled double bootstrap for massive data. *Journal of the American Statistical Association (Theory and Methods)*, **111**, 1222-1232.
(Finalist: ASA Nonparametric Statistics Student Paper Award, 2015)
3. **Sengupta, S.** and Chen, Y. (2015). Spectral clustering in heterogeneous networks. *Statistica Sinica*, **25**, 1081-1106.
4. **Sengupta, S.**, Shao, X., and Wang, Y. (2015). The dependent random weighting. *Journal of Time Series Analysis*, **36**, 315-326.
(**Invited paper** for the special issue on “Recent developments in bootstrap methods for dependent data”)
5. **Sengupta, S.** (2010). Modeling the zero coupon yield curve: A regression based approach. In *Proceedings of 12th Global Conference of Actuaries*.

RESEARCH INTERESTS Network data, Community structure in networks, Random graph models, Bootstrap and related resampling methods, Big data, Time series, Spatial data, Statistical computing, Machine learning.

AWARDS	1. Norton Prize for Outstanding PhD Thesis, University of Illinois	2015
	2. IISA student paper award finalist	2015
	3. ASA Nonparametric Statistics Section student paper finalist	2015
	4. NSF travel award, Statistics for Complex Systems	2015
	5. Graduate College travel award, University of Illinois	2013, 2014, 2015
	6. Birla Sun Life Excellence award, Institute of Actuaries of India	2009
	7. Dean's List for exceptional performance, Indian Statistical Institute	2004

TEACHING
EXPERIENCE
AT UIUC

1. **Primary instructor with full course responsibility**

STAT 200 (Statistical Analysis) Spring 2013, Spring 2015
– Prepared lecture materials, R modules, assignments, and exams.
– Gave three lectures per week to 50 students, for two semesters.

2. Discussion section instructor

STAT 400 (Statistics and Probability I) Fall 2011, Spring 2012
– Prepared review material covering concepts, techniques, and examples.
– Led two weekly review sections for 100 students, for two semesters.

3. Teaching assistant for PhD level courses:

STAT 510 (Mathematical Statistics I) Fall 2012
STAT 553 (Probability and Measure I) Fall 2014, Fall 2015
STAT 554 (Probability and Measure II) Fall 2013

RESEARCH
EXPERIENCE
AT UIUC

Networks with community structure

Advisor: Prof. Yuguo Chen

- Developed methodology for community detection in heterogeneous networks, including the heterogeneous stochastic blockmodel and heterogeneous spectral clustering algorithm.
- Developed methodology for accurate modeling of node popularity in networks, including the popularity-adjusted blockmodel and likelihood modularity.

Modern bootstrap methods for big data

Advisor: Prof. Xiaofeng Shao

- Developed the dependent random weighting method for resampling from irregularly spaced time series and spatial data.
- Developed the subsampled double bootstrap for fast and accurate resampling from massive data, both dependent and independent.

STATISTICAL CONSULTING PROJECTS

Client: University Library, University of Illinois

Spring 2014

1. Identifying Serialized Fiction From Text Corpus

Data: Illinois Digital Newspaper Collections (7 million articles).

Objective: Separate serialized fiction pieces from news articles.

Challenge: Noisy, massive, and complex data (articles from 1831-2011).

Solution: Developed Natural Language Processing based tools for classification.

2. Newspaper usage and ranking: a cost-benefit analysis

Data: Daily newspaper hits for 40,000 titles for 2 years (2012-13).

Objective: Compare newspaper subscriptions from various vendors.

Challenges: Quantification of an implicit problem, usage outliers, seasonality.

Solution: Cost-vs-benefit analysis with usage and ranking scores.

PRESENTATIONS

1. "Modeling node popularity in networks and bootstrap methods for massive data" at *University of Florida*, Jan 2015
2. "Modeling node popularity in networks and bootstrap methods for massive data" at *Ohio State University*, Jan 2015
3. "Modeling node popularity in networks and bootstrap methods for massive data" at *Virginia Tech*, Dec 2015
4. "Modeling node popularity in networks" *Norton Lecture* at *Bohrer Workshop*, University of Illinois at Urbana-Champaign, Nov 2015
5. "A subsampled double bootstrap for massive data" at *Joint Statistical Meetings*, Seattle, Aug 2015
6. "A subsampled double bootstrap for massive data" (Poster) at *Conference on Statistics for Complex Systems*, University of Wisconsin-Madison, Jun 2015
7. "A subsampled double bootstrap for massive data" at *Bohrer Workshop*, University of Illinois at Urbana-Champaign, Nov 2014
8. "Spectral clustering in heterogeneous networks" at *Joint Statistical Meetings*, Boston, Aug 2014.
9. "Spectral clustering in heterogeneous networks" at *Bohrer Workshop*, University of Illinois at Urbana-Champaign, Nov 2013
10. "Modeling the zero coupon yield curve: A regression based approach" at the *12th Global Conference of Actuaries*, Mumbai, India, Feb 2010

ACTUARIAL
QUALIFICATION

- Worked for two years as risk management and product pricing actuary, specializing in asset-liability matching, pricing the cost of embedded financial guarantees, value-at-risk, and economic capital
- Actuarial qualification: passed 12 exams from Institute and Faculty of Actuaries (UK)/ Institute of Actuaries of India.
 - Specialist technical subjects ST6 (Finance & Investment B) and ST9 (Enterprise risk management).
 - Core application subject CA3 (*ranked 1st in India*)
 - Core technical subjects CT1 – CT9 (complete CT series)

REFEREED
JOURNALS

- Computational Statistics & Data Analysis
- Electronic Journal of Statistics
- Technometrics
- Journal of Time Series Analysis

SERVICE AND
ACTIVITIES

- Serving as Member of Student Appeals Committee, Department of Statistics, University of Illinois at Urbana-Champaign.
- Membership in Professional Societies:
 - American Statistical Association
 - International Chinese Statistical Association
 - International Indian Statistical Association
 - Institute of Mathematical Statistics