**CURRICULUM VITAE**

*SOUMIK PURKAYASTHA, Ph.D.*

---

*Current research interests*

1. Information-theoretic framework for **association** and **causality**: development of statistical methods for studying association and causality without relying on traditional causal inference assumptions, with applications in areas like mediation analysis and instrumental variables.

2. Compartmental models for **infectious disease modeling**: development of spatiotemporal forecasting techniques to study transmission and fallout of infectious diseases, with specific emphasis on COVID-19.

## 1. CONTACT INFORMATION


**OFFICE** A751, Crabtree Hall,
Department of Biostatistics and Health Data Science
Graduate School of Public Health, University of Pittsburgh,
130 De Soto Street, Pittsburgh, PA 15261, USA.

**PHONE** 734-881-5075

**EMAIL** soumik@pitt.edu

**CITIZENSHIP** Indian citizen


## 2. EDUCATION

| | |
|---|---|
| 2014 - 17 | B.Sc. (Hons.)<br>St. Xavier's College (Autonomous), Kolkata, WB, INDIA. |
| 2017 - 19 | M.Stat. (Specialisation in Biostatistics)<br>Indian Statistical Institute, Kolkata, WB, INDIA. |
| 2019 - 21 | M.S. in Biostatistics<br>University of Michigan, Ann Arbor, MI, USA. |
| 2019 - 24 | Ph.D. in Biostatistics<br>University of Michigan, Ann Arbor, MI, USA. |

## 3A. ACADEMIC APPOINTMENTS AND POSITIONS

| | |
|---|---|
| 2019 - 20 | Graduate Student Research Assistant,<br>Department of Biostatistics: Abecasis Lab, University of Michigan, Ann Arbor, MI, USA. |

| | |
|---|---|
| 2020 - 23 | Graduate Student Research Assistant,<br>Department of Biostatistics: Statistical Analysis of Biomedical and Educational Research (SABER) group, University of Michigan, Ann Arbor, MI, USA. |
| 2023 - 24 | Rackham Predoctoral Fellow,<br>Department of Biostatistics, University of Michigan, Ann Arbor, MI, USA. |
| 2024 - present | Assistant Professor (tenure stream),<br>Department of Biostatistics and Health Data Science, University of Pittsburgh, Pittsburgh, PA, USA. |
| 2024 - present | Affiliate Investigator,<br>US Department of Veteran Affairs Centre for Health Equity Research and Promotion, Pittsburgh Healthcare System, Pittsburgh PA, USA. |

## 3B. NON-ACADEMIC APPOINTMENTS AND POSITIONS

| | |
|---|---|
| 2018 | Data Science intern, Walmart Labs, Bengaluru, KA, INDIA. |
| 2021 | AI/ML intern, Apple Inc., Cupertino, CA, USA. |

## 4. AWARDS

| | |
|---|---|
| 2017 - 19 | Merit-based scholarship for good academic performance,<br>Indian Statistical Institute, Kolkata, WB, INDIA. |
| 2019 | Sabyasachi Roy Memorial Gold Medal,<br>Indian Statistical Institute, Kolkata, WB, INDIA. |
| 2019 | Michigan Data Science Challenge Winner,<br>Michigan Institute of Data Science, University of Michigan, Ann Arbor, MI, USA. |
| 2020 | Richard G. Cornell Fellowship,<br>Department of Biostatistics, University of Michigan, Ann Arbor, MI, USA. |
| 2022 - 24 | Rackham Conference Travel Grant (awarded annually),<br>Rackham Graduate School, University of Michigan, Ann Arbor, MI, USA. |
| 2023 | Rising Star Award,<br>School of Public Health, University of Michigan, Ann Arbor, MI, USA. |
| 2023 - 24 | Rackham Predoctoral Fellowship,<br>Rackham Graduate School, University of Michigan, Ann Arbor, MI, USA. |
| 2023 - 24 | Best Paper Award and Best Presentation (Runner-up) Award,<br>Western North American Region of the International Biometrics Society. |
| 2024 | Excellence in Research (Honourable Mention), |

Department of Biostatistics, University of Michigan, Ann Arbor, MI, USA.

## 5. MEMBERSHIP IN PROFESSIONAL AND SCIENTIFIC SOCIETIES

2021 - present International Biometrics Society, Eastern North American Region (ENAR).

2021 - present American Statistical Association (ASA).

2021 - present Institute of Mathematical Statistics (IMS).

2022 - present International Biometrics Society, Western North American Region (WNAR).

2023 - present International Indian Statistical Association (IISA).

## 6. PROFESSIONAL SERVICE

**Reviewer for the following journals:**
*Annals of Applied Statistics* (2022+), *New England Journal of Statistics in Data Science* (2022+), and *PLOS One* (2021+).

**Service for School and University:**

*University of Michigan*

2020 - 22 Department of Biostatistics Seminars and Brown Bag Committee: Member.

2021 - 24 Statistics in the Community (STATCOM) at the University of Michigan: President (2022-23) and leadership team member (2021-24).

2022 - 23 Department of Biostatistics Faculty Meetings: Student representative.

2023 - 24 Michigan Institute of Data Science (MIDAS) Student Council: Leadership team member.

*University of Pittsburgh*

2024 - present Graduate Research and Development (GRAD) seminar: organiser.

## 7. PUBLICATIONS

h-index: 10 (Google scholar) as of October, 2024.
+: co-first author.

*Statistical papers (from independent methodological research)*

1. Ray, D., Salvatore, M., Bhattacharyya, R., Wang, L., Du, J., Mohammed, S., **Purkayastha, S.**, Halder, A., Rix, A., Barker, D., Kleinsasser, M., Zhou, Y., Bose, D., Song, P., Banerjee, M.,

Baladandayuthapani, V., Ghosh, P., & Mukherjee, B. (2020). Predictions, Role of Interventions, and Effects of a Historic National Lockdown in India's Response to the COVID-19 Pandemic: Data Science Call to Arms. *Harvard Data Science Review, (Special Issue 1).*

2. Zhou, Y., Wang, L., Zhang, L., Shi, L., Yang, K., He, J., Bangyao, Z., Overton, W., **Purkayastha, S.**, & Song, P. (2020). A Spatiotemporal Epidemiological Prediction Model to Inform County-Level COVID-19 Risk in the United States. *Harvard Data Science Review, (Special Issue 1).*

3. **Purkayastha, S.**, Salvatore, M. and Mukherjee, B., 2020. Are women leaders significantly better at controlling the contagion during the COVID-19 pandemic? *Journal of health and social sciences, 5(2), p.231.*

4. Tang, L., Zhou, Y., Wang, L., **Purkayastha, S.**, Zhang, L., He, J., Wang, F. and Song, P.X.K., 2020. A review of multi-compartment infectious disease models. *International Statistical Review, 88(2), pp.462-513.*

5. **Purkayastha, S.** and Song, P., 2021. Discussion on "The timing and effectiveness of implementing mild interventions of COVID-19 in large industrial regions via a synthetic control method" by Tian et al. *Statistics and Its Interface, 14(1), pp.21-22.*

6. Salvatore, M., Basu, D., Ray, D., Kleinsasser, M., **Purkayastha, S.**, Bhattacharyya, R. and Mukherjee, B., 2020. Comprehensive public health evaluation of lockdown as a non-pharmaceutical intervention on COVID-19 spread in India: national trends masking state-level variations. *BMJ open, 10(12), p.e041778.*

7. **Purkayastha, S.**, Bhattacharyya, R., Bhaduri, R., Kundu, R., Gu, X., Salvatore, M., Ray, D., Mishra, S. and Mukherjee, B., 2021. A comparison of five epidemiological models for transmission of SARS-CoV-2 in India. *BMC infectious diseases, 21, pp.1-23.*

8. **Purkayastha, S.,** Kundu, R., Bhaduri, R., Barker, D., Kleinsasser, M., Ray, D. and Mukherjee, B., 2021. Estimating the wave 1 and wave 2 infection fatality rates from SARS-CoV-2 in India. *BMC research notes, 14, pp.1-7.*

9. Zimmermann, L., Bhattacharya, S., **Purkayastha, S.**, Kundu, R., Bhaduri, R., Ghosh, P. and Mukherjee, B., 2021. SARS-CoV-2 infection fatality rates in India: systematic review, meta-analysis and model-based estimation. *Studies in Microeconomics, 9(2), pp.137-179.*

10. Bhaduri, R., Kundu, R., **Purkayastha, S.**, Kleinsasser, M., Beesley, L.J., Mukherjee, B. and Datta, J., 2022. Extending the susceptible-exposed-infected-removed (SEIR) model to handle the false negative rate and symptom-based administration of COVID-19 diagnostic tests: SEIR-fansy. *Statistics in medicine, 41(13), pp.2317-2337.*

11. Salvatore, M., **Purkayastha, S.**, Ganapathi, L., Bhattacharyya, R., Kundu, R., Zimmermann, L., Ray, D., Hazra, A., Kleinsasser, M., Solomon, S. and Subbaraman, R., 2022. Lessons from SARS-CoV-2 in India: A data-driven framework for pandemic resilience. *Science advances, 8(24), p.eabp8621.*

12. **Purkayastha, S.** and Song, P.X.K., 2024. fastMI: A fast and consistent copula-based nonparametric estimator of mutual information. *Journal of Multivariate Analysis, 201, p.105270.*

*Collaborative papers (from interdisciplinary collaborative research)*

1. Giri, S., **Purkayastha, S.**, Hazra, S., Chanda, A., Das, I. and Das, S., 2020. Prediction of monthly Hilsa (Tenualosa ilisha) catch in the Northern Bay of Bengal using Bayesian structural time series model. *Regional Studies in Marine Science, 39, p.101456.*

*Statistical papers under preparation/review*

***1.*** **Purkayastha, S.** and Song, P. X. K. *Asymmetric predictability in causal discovery: an information theoretic approach.*

***2.*** **Purkayastha, S.** and Song, P. X. K. *Generative causality: using Shannon's information theory to infer underlying asymmetry in causal relations.*

***3.*** **Purkayastha, S.** and Basu, A. *On minimum Bregman divergence inference.*

*Collaborative papers under preparation/review*

1. Zhang, L., **Purkayastha, S.**, Kirsner, R., Spino, C. & Song, P. X. K. *Determinants of Enrolment in 284 Clinical Trials for Healing Diabetic Foot Ulcers: A Systematic Review.*

2. Kadura, S., **Purkayastha, S.**, Spino, C., Benditt, J., Anand, A., De Quadros, M., Hobson, M., Biswas, M. J., Collins, B., Ho, L., Raghu, G. *Yoga effect on quality-of-life among patients with idiopathic pulmonary fibrosis (YES-IPF): A prospective, randomized control pilot trial by virtual means during the COVID-19 pandemic in Seattle, USA.*

## 8. SOFTWARE

1. Bhaduri R., Kundu R., **Purkayastha S.**, Beesley L., Mukherjee B., Kleinsasser, M. 2021. *SEIRfansy*: Extended Susceptible-Exposed-Infected-Recovery Model [https://CRAN.R-project.org/package=SEIRfansy]

2. **Purkayastha, S.** and Song, P.X.K., 2024. *fastMI*: A fast and consistent copula-based nonparametric estimator of mutual information. [https://github.com/soumikp/fastMI]

3. **Purkayastha, S.** and Song, P.X.K., 2024. *comet*: Collider-mediator testing using information theory [https://github.com/soumikp/comet]

## 9. CONFERENCES AND WORKSHOPS

1. Eastern North American Region (ENAR) of the International Biometric Society Spring Meeting (March 2023). *"An information-theoretic framework for causal discovery in epigenetic data"*

2. Western North American Region (WNAR) of the International Biometric Society Spring Meeting (June 2023). *"Asymmetric predictability in causal discovery: an information theoretic approach."*

3. Joint Statistical Meetings (August 2023). *"Asymmetric predictability in causal discovery: an information theoretic approach."*

4. Workshop on *Foundations of Causal Graphical Models and Structure Discovery*. Travel award provided by the National Science Foundation (NSF DMS-2227849) and Texas A&M Institute for Applied Mathematics and Computational Science (IAMCS).

## 10. SELECTED PRESS

1. Basu, D., Salvatore, M., Kleinsasser, M., Purkayastha, S., Bhattacharya, R., and Mukherjee, B., We're Focusing on National Data on COVID-19 When We Should Be Looking at State- Level Trends. The Wire, 2020.
2. Laguipo, A., A study of the COVID fatality rates in India during waves 1 and 2. News-Medical.Net, 2021.
3. Mukherjee, B., Purkayastha, S., Salvatore, M., & Mishra, S. Underreporting does hurt the COVID fight. The Hindu, 2021.
4. Ellis-Petersen, H. India's 1.3bn population locked down to beat coronavirus. The Guardian, 2021.
5. Bastian, H., Women Versus Men Leaders in the Pandemic: An Update and Dig Into the Latest Data. PLOS Blogs, 2022.