

## Syllabus SURV662

### Intro to Small Area Estimation 1 credit

Instructor: Partha Lahiri, PhD

Video lectures prepared by Partha Lahiri, PhD

March 5, 2024 – March 26, 2024

#### Short Course Description

The course introduces the students to the fundamental concepts of small area estimation. The course is organized in three main sections: basic concepts and challenges of small area estimation, how SAE methods can improve on direct estimation, and examples to demonstrate effective use of SAE methods in applications.

#### Course Objectives

By the end of the course, students will...

- Understand why standard design-based methods may fail to provide reliable small area estimates.
- Understand how to combine information from disparate data sources using appropriate statistical models and classical prediction and Bayesian methods.
- Learn differences between mixed models and regression models and why mixed models are more suited in small area estimation.
- Learn how to conduct small area analyses using complex survey data.

#### Prerequisites

An introductory applied statistics course that includes regression analysis and survey sampling.

#### Class Structure and Course Concept

This is an online course, using a flipped classroom design. It covers the same material and content as an on-site course but runs differently. In this course, you are responsible for watching video-recorded lectures and reading the required literature for each unit prior to participating in mandatory weekly one-hour online meetings where students have the chance to discuss the materials from a unit with the instructor. Just like in an on-site course, homework will be assigned and graded and there will be a final exam at the end of the course.

Although this is an online course where students have more freedom in when they engage with the course materials, students are expected to spend the same amount of time overall on all activities in the course – including preparatory activities (readings, studying), in-class-activities (watching videos, participating in online meetings), and follow-up activities (working on assignments and exams) – as in an on-site course. As a rule of thumb, you can expect to spend approximately 3h/week on in-class-activities and 9 hours per week on out-of-class activities (preparing for class, readings, assignments, projects, studying for quizzes and exams). Therefore, the workload in all courses will be approximately 12h/week. This is a 1-credit/2-ECTS course that runs for 4 weeks. Please note that the actual workload will depend on your personal knowledge.

## Mandatory Weekly Online Meetings

*Tuesdays at 6:00 PM CET (12:00 PM EST), starting 5, March 2024*

Meetings will be held online through Zoom. Follow the link to the meeting sessions on the course website on <https://umd.zoom.us/> . If video participation via Internet is not possible, arrangements can be made for students to dial in and join the meetings via telephone.

In preparation for the weekly online meetings, students are expected to watch the lecture videos and read the assigned literature before the start of the meeting. In addition, students are encouraged to post questions about the materials covered in the videos and readings of the week in the forum before the meetings (deadline for posting questions is two days before the online meetings, i.e. on Friday each week).

Students have the opportunity to use a different Zoom meeting room to connect with peers outside the scheduled weekly online meetings (e.g., for study groups). Detailed information is posted on the course page in Canvas. Students are encouraged to post the times that they will be using the room to the course website forum to avoid scheduling conflicts. Students are not required to use Zoom and can use other online meeting platforms, such as Microsoft Teams, Google Hangout or Skype.

## Grading

Grading will be based on:

- 4 homework assignments (worth 80%)
- Participation in the weekly online meetings (20% of grade): engagement in discussions during the meetings and submission of questions to the weekly discussion forums

Each assignment will be due three days after the corresponding online meeting. Deadline is on each Monday starting with 5 March 2024 11:30 PM CET (5:30 PM)

EST). Exact dates are indicated in the syllabus and on the Canvas page. Extensions will be granted sparingly and are at the instructor's discretion. The following scale will be used for determining the final course grade.

|             |                 |
|-------------|-----------------|
| <i>A+</i>   | <i>100 - 97</i> |
| <i>A</i>    | <i>97 - 93</i>  |
| <i>A-</i>   | <i>93 - 90</i>  |
| <i>B+</i>   | <i>90 - 87</i>  |
| <i>B</i>    | <i>87 - 83</i>  |
| <i>B-</i>   | <i>83 - 80</i>  |
| <i>Etc.</i> |                 |

The final grade will be communicated under the assignment "Final Grade" in the Canvas course. Please note that the letter grade written in parentheses in Canvas is the correct final grade. The point-grade displayed alongside the letter grade is irrelevant and can be ignored.

### Technical Equipment Needs

The learning experience in this course will mainly rely on the online interaction between the students and the instructors during the weekly online meetings. Therefore, we encourage all students in this course to use a web camera and a headset. Decent quality headsets and web cams are available for less than \$20 each. We ask students to refrain from using built-in web cams and speakers on their desktops or laptops. We know from our experience in previous online courses that this will reduce the quality of video and audio transmission and therefore will decrease the overall learning experience for all students in the course. In addition, we suggest that students use a wire connection (LAN), if available, when connecting to the online meetings. Wireless connections (WLAN) are usually less stable and might be dropped.

### Long Course Description

The demand for various socio-economic and health statistics for small geographical areas is steadily increasing at a time when survey agencies are constantly looking for ways to reduce costs to meet fixed budgetary requirements. In the current survey environment, the application of standard sample survey methods, which require a large sample, is generally not feasible for small domains, as the cost would be prohibitive. One of the key factors that leads to the success of small area estimation (SAE) methodology in many applications, is the availability of strong auxiliary variables. The accessibility of Big Data from disparate sources (e.g., administrative/register records, social media data, mobile phone data, sensor data, satellite data, etc.) brings new opportunities for statisticians to develop innovative SAE methods.

We will begin the course by presenting a brief history of small area estimation, basic concepts, and challenges. We will then discuss shortcomings of direct estimation for small area estimation problems and how different SAE methods can improve on direct estimation. We will then discuss a few illustrative examples to demonstrate how SAE can be effectively used in applications ranging from the production of sustainable goal indicators at disaggregated levels, to the production of granular coronavirus pandemic data, to election projections. We will illustrate available R packages for data analysis.

## Readings

Primary readings will be from the following lecture notes:

Lahiri, P. (2022), An Introduction to Small Area Estimation, Unpublished Lecture Notes

Lists of required and recommended readings during the course are provided below for each session.

## Academic Conduct

Clear definitions of the forms of academic misconduct, including cheating and plagiarism, as well as information about disciplinary sanctions for academic misconduct may be found at

<https://www.president.umd.edu/sites/president.umd.edu/files/documents/policies/III-100A.pdf> (University of Maryland)

Knowledge of these rules is the responsibility of the student and ignorance of them does not excuse misconduct. The student is expected to be familiar with these guidelines before submitting any written work or taking any exams in this course. Lack of familiarity with these rules in no way constitutes an excuse for acts of misconduct. Charges of plagiarism and other forms of academic misconduct will be dealt with very seriously and may result in oral or written reprimands, a lower or failing grade on the assignment, a lower or failing grade for the course, suspension, and/or, in some cases, expulsion from the university.

## Accommodations for Students with Disabilities

In order to receive services, students at the University of Maryland must contact the Accessibility & Disability Service (ADS) office to register in person for services. Please call the office to set up an appointment to register with an ADS counselor. Contact the ADS office at 301.314.7682; <https://www.counseling.umd.edu/ads/>.

## Course Evaluation

In an effort to improve the learning experience for students in our online courses,

students will be invited to participate in an online course evaluation at the end of the course. Participation is entirely voluntary and highly appreciated.

## Sessions

Canvas pages for each week, containing video lectures, slides, materials and assignments will be released one week before each video lecture.

Assignments are due three days after each online meeting. Questions should be posted on the Forum two days before each online meeting.

*Please note that assignments and dates are subject to change. Information (e.g., articles and assignments) posted to the course website supersedes the information noted here.*

### Week 1: Introduction

By the end of this unit, you will know

- Basic concepts and challenges in small area estimation
- Why direct survey-weighted estimation does not work for small area estimation problems

Video lecture: available on Tuesday, February 27, 2024

Online meeting: Tuesday, March 5, 2024, 6:00 PM CET (12:00 PM EST)

Assignment 1:

- Due: Friday, March 8, 2024, 11:30 PM CET (5:30 PM EST)

#### **Required readings:**

Lahiri, P. (2022), An Introduction to Small Area Estimation, Unpublished Lecture Notes

#### **Recommended readings (not mandatory):**

Rao, J.N.K. and Molina, I. (2015) Small Area Estimation. Second Edition, Wiley Series in Survey Methodology

## Week 2: Synthetic Methods

By the end of this unit, you will know...

- How to combine information from disparate data sources using a variety of simple methods that may not use explicit statistical models
- Uncertainty measures for synthetic estimates
- Advantages of synthetic estimation methods
- Different issues and challenges in synthetic methods

Video lecture: available on Tuesday, March 5, 2024

Online meeting: Tuesday, March 12, 2024, 6:00 PM CET (12:00 PM EST)

Assignment 2:

- Due: Friday, March 15, 2024, 11:30 PM CET (5:30 PM EST)

### **Required readings:**

Lahiri, P. (2022), An Introduction to Small Area Estimation, Unpublished Lecture Notes

### **Recommended readings (not mandatory):**

Rao, J.N.K. and Molina, I. (2015) Small Area Estimation. Second Edition, Wiley Series in Survey Methodology

## Week 3: Area Level Models

By the end of this unit, you will know...

- How to model survey-weighted estimates at the area levels
- Data preparation for implementing area level models
- How to obtain smoothed sampling variance estimates of direct estimates
- How to produce Empirical Bayes/Best (EB) and Hierarchical Bayes (HB) estimates
- Uncertainty measures of EB and HB estimates.
- How to produce estimates and the associated uncertainty measures using software
- Challenges in implementing area level models.

Video lecture: available on Tuesday, March 12, 2024

Online meeting: Tuesday, March 19, 2024, 6:00 PM CET (12:00 PM EST)

### Assignment 3:

- Due: Friday, March 22, 2024, 11:30 PM CET (5:30 PM EST)

#### **Required readings:**

Lahiri, P. (2022), An Introduction to Small Area Estimation, Unpublished Lecture Notes

#### **Recommended readings (not mandatory):**

Rao, J.N.K. and Molina, I. (2015) Small Area Estimation. Second Edition, Wiley Series in Survey Methodology

## Week 4: Unit Level Models

By the end of this unit, you will know

- How to model respondent level observations
- Data preparation for implementing unit level models
- EB and HB estimates for unit level models
- Uncertainty measure of EB and HB estimates
- How to produce estimates and associated uncertainty measures using software
- Challenges in implementing unit level models.

Video lecture: available on Tuesday, March 19, 2024

Online meeting: Tuesday, March 26, 2024, 6:00 PM CET (12:00 PM EST)

### Assignment 4:

- Due: Friday, March 29, 2024, 11:30 PM CET (6:30 PM EDT)

#### **Required readings:**

Lahiri, P. (2022), An Introduction to Small Area Estimation, Unpublished Lecture Notes

#### **Recommended readings (not mandatory):**

Rao, J.N.K. and Molina, I. (2015) Small Area Estimation. Second Edition, Wiley Series in Survey Methodology