Module 5: Sampling

**Lesson 1: Introduction and Foundations**

**Estimated Time**: ~7.5 hours

**Concepts:** probability; populations; observational studies

**Lesson Description**: This lesson will introduce students to core concepts in the module.

**Instructor Preparation**: Overview of course texts, read through lesson slides, create Slido polls for 00-Introduction and 01-Probability

|  |  |
| --- | --- |
| **Materials and Resources** | **Learning Goals** |
| 1. The slides 2. Copies of/links to core course texts    1. Pitman, 1993, *Probability*, Springer.    2. Wu and Thompson, 2020, *Sampling Theory and Practice*, Springer.    3. Lohr, 2019, *Sampling Design and Analysis*, 2nd Edition, CRC Press.    4. Salganik, 2018, Bit by Bit: Social research in the Digital Age. 3. Python and R script files    1. Probability-CLT Demo    2. Probability-LLN Demo 4. Optional: deck of cards 5. Optional: dice 6. Optional: coin | * Be familiar with overall module content * Have access to course texts * Connect probability concepts to sampling at a high level * Understand the basics of sampling populations and strategies |

|  |  |  |
| --- | --- | --- |
| **Time** | **Lesson Content** | **Instructor Notes** |
| 20 minutes  35 minutes  10 minutes | **00-Introduction**  **Introduction**   1. Welcome students to the class, introductions   **Lesson**   1. [00-Introduction-slides](https://github.com/UofT-DSI/sampling/blob/main/lectures/00-Introduction-slides.pptx) 2. Walk through each course text as it comes up in the slides   **Activity / Discussion**   1. Use [Slido](https://www.sli.do/) to poll students on the following questions (either list sections from intro deck or use open ended questions):    1. What topic are you most excited for?    2. What topic are you most nervous for? | Would be good at this stage to ensure students all have access to relevant texts and assist those who don’t |
| 10 minutes  180 minutes | **01-Probability**  **Introduction**   1. Use [Slido](https://www.sli.do/) to poll students on knowledge of math and probability concepts    1. Have you ever taken a course involving: Calculus; Linear algebra; Statistics; Probability theory; Sampling theory? Select all that apply 2. Share files with students    1. 5.1-Probability-CLT Demo (Python and R)    2. 5.1-Probability-LLN Demo (Python and R)    3. DSI 5.1 Probability Cheat Sheet 1/2/3.pdf   **Lesson**   1. Distribute “DSI 5.1 Probability Cheat Sheet 1/2/3.pdf” to students prior to lesson so they have them accessible 2. [01-Probability-slides](https://github.com/UofT-DSI/sampling/blob/main/lectures/01-Probability-slides.pptx)    1. **Slide 40** (“Law of Large Numbers”): Open “5.1-Probability-LLN Demo” and run Python or R demo    2. **Slide 57** (“Central Limit Theorem”): Open “5.1-Probability-CLT Demo” and run Python or R demo | Use coins, dice, and cards as desired for examples/demonstrations |
| 50 minutes  10 minutes | **02-Populations, censuses, surveys, and observational data**  **Lesson**   1. [02-Populations, censuses, surveys, and observational data-slides](https://github.com/UofT-DSI/sampling/blob/main/lectures/02-Populations%2C%20censuses%2C%20surveys%2C%20and%20observational%20data-slides.pptx)    1. See **Activity/Discussion** for in-lesson activities   **Activity/Discussion**   1. **Slide 13**: Host discussion surrounding Example 1.1 from Wu and Thompson (answers in slide speaker notes) 2. **Slide 22**: Go through examples of observational studies, ask students to note specific methodologies that distinguish these from surveys    1. Alexander, M., Polimis, K. and Zagheni, E. (2019), The Impact of Hurricane Maria on Out-migration from Puerto Rico: Evidence from Facebook Data. Population and Development Review, 45: 617-630. <https://doi.org/10.1111/padr.12289>    2. Mishra S, Ma H, Moloney G, Yiu KCY, Darvin D, Landsman D, Kwong JC, Calzavara A, Straus S, Chan AK, Gournis E, Rilkoff H, Xia Y, Katz A, Williamson T, Malikov K, Kustra R, Maheu-Giroux M, Sander B, Baral SD; COVID-19 Heterogeneity Research Group. Increasing concentration of COVID-19 by socioeconomic determinants and geography in Toronto, Canada: an observational study. Ann Epidemiol. 2021 Jul 25:S1047-2797(21)00216-7. doi: [10.1016/j.annepidem.2021.07.007](https://pubmed.ncbi.nlm.nih.gov/34320380/). Epub ahead of print. PMID: 34320380. | Connect population concepts to probability concepts from previous lesson content. |
| 40 minutes  30 minutes  30 minutes | **03-Essentials of sampling, asking, and observing-slides**  **Introduction/Assessment**   1. [ASSIGNMENT-Questionnaire Design (Part A)](https://github.com/UofT-DSI/sampling/blob/main/assessment/ASSIGNMENT%20-%20Questionnaire%20Design%20(Part%20A).md)    1. Can be done in groups or individually   **Lesson**   1. [03-Essentials of sampling, asking, and observing-slides](https://github.com/UofT-DSI/sampling/blob/main/lectures/03-Essentials%20of%20sampling%2C%20asking%2C%20and%20observing-slides.pptx)    1. Slides 13-15 will require reference to content in Lohr’s *Sampling Design and Analysis,* Section 1.5   **Assessment**   1. [ASSIGNMENT-Questionnaire Design (Part B)](https://github.com/UofT-DSI/sampling/blob/main/assessment/ASSIGNMENT%20-%20Questionnaire%20Design%20(Part%20B).md) |  |