Stat 472/572: Sampling Theory and Practice, Spring 2026

Instructor: Dr. Yan Lu, yanlu@unm.edu

Teaching Assistant:

Prerequisites: Stat 345 and ADA1: Stat 427/527 Advanced Analysis 1. Stat 461 (probability) is a helpful course but is not required.

Class website (Spring 2024 notes): https://www.math.unm.edu/~luyan/stat472572/stat472572.html

Class Time and Location: Online MAX Arranged

- This course follows Tuesday/Thursday schedule.
- On **Tuesdays and Thursdays**:
  - Read the assigned chapter and lecture slides.
  - Watch the corresponding video lectures to reinforce your understanding.
  - Practice R coding with data.

Office Hours:

- Tuesdays and Thursday 5:00pm-6:00pm via Zoom.
  https://unm.zoom.us/j/6727234389
  passcode:123
  or by appointment
  Please feel free to email me if you have any questions.

Textbook:

Lohr, Sharon. Sampling: Design and Analysis 3rd edition. Boca Raton, FL: CRC Press. (**required**).

Lu, Yan and Lohr, Sharon: R Companion for *Sampling: Design and Analysis, 3rd Edition*, 1st Edition. Boca Raton, FL: CRC Press (**Electronic version can be downloaded here:**) https://www.math.unm.edu/~luyan/rbook/Rbook.pdf

Datasets are available from R package SDAResources or you can download from https://www.math.unm.edu/~luyan/rbook/rbook.html

Topics: In classical finite population sampling, probability sampling is used to select a sample from a directory or map of units, often called a sampling frame. Sample surveys play a crucial role in gathering information about finite populations of interest in various fields such as education, public health, sociology, ecology, agriculture, genetics, quality improvement, marketing, and accounting. This course focuses on designing effective surveys and analyzing data from sample surveys. Topics covered include simple random sampling, stratified and cluster sampling, ratio estimation, addressing nonresponse, variance estimation, as well as regression analyses and chi-square tests tailored specifically for survey data.

Computing:

- **R:** windows: `https://cran.r-project.org/bin/windows/base/`
  mac: `https://cran.r-project.org/bin/macosx/`

- **RStudio:** Integrated development environment for R, available at
  `https://posit.co/download/rstudio-desktop/`.
  Installation guides:
  – Windows: `https://www.youtube.com/watch?v=eD07NznguA4`
  – Mac: `https://www.youtube.com/watch?v=AEebOXiMyyI`

- **Quarto:** An open-source scientific and technical publishing system. With Quarto, you can combine text, R code, and mathematical notation (via LaTeX) into a single source file, then render it into multiple formats including Word, HTML, and PDF.
  Documentation and installation instructions:
  `https://quarto.org/docs/get-started/`

- **TinyTeX:** A lightweight and easy-to-install LaTeX distribution. TinyTeX is required if you wish to render PDF files directly from Quarto. It provides all necessary LaTeX tools without the bulk of full distributions. Installation in R:

  ```
  install.packages("tinytex")
  tinytex::install_tinytex()
  ```

  To confirm installation:

```
tinytex::tinytex_root()
```

More information: `https://yihui.name/tinytex/`

- Getting started with R, Rstudio, Quarto, and Tinytex
  `https://www.youtube.com/watch?v=VbWu-oZ45Q0&t=23s`

Grading: Homework assignments will be distributed approximately every two weeks, totaling 7-8 assignments throughout the course. The grading breakdown is as follows: Homework (due Friday midnight) 30%, Midterm Exam (75 minutes, on line by zoom, with cameras on) 30%, and Final Project 40%. Please note that the grading criteria for graduate students and undergraduates will be assessed separately.

	Stat 572	Stat 472
A+	96%-100%	93%-100%
A	90%-95%	85%-92%
B	75%-89%	70%-84%
C	65%-74%	55%-69%
D	under 65%	under 55%

Final Project: Your project in this class will encompass either a theoretical or applied aspect of sampling. You have the option to work independently or collaborate with your classmates, with group sizes limited to a maximum of three individuals.Examples include: analyzing data from a complex sample survey from public website or from your research area; creating a sampling design for a long-term research project; or investigating some theoretical sampling problems. More detailed information will be given in class later.