

SURV 625
Applied Sampling
Spring 2010

SURVMETH 612
Methods of Survey Sampling
Winter 2010

CLASS MEETINGS

1:00-4:00 PM, Wednesdays
Room 1208 Lefrak, University of Maryland, College Park, Maryland
Room G300 ISR-Perry, University of Michigan, Ann Arbor, Michigan
T10, US Census Bureau, Suitland, Maryland

INSTRUCTOR

James M. Lepkowski
Room 4053 ISR (U of Michigan) and Room 1218S Lefrak (U of Maryland)
Telephone: 734/678-6964 (cell) and 301/314-7911 (U of Maryland)
Email: jimlep@umich.edu or jlepkowski@survey.umd.edu

ASSISTANTS

Mahmoud ElKasabi
Room 4062 ISR, 426 Thompson Street, University of Michigan
Telephone: 734/647-0421
Email: mahmelka@umich.edu

Patricia Gregory
Room 4050 ISR, 426 Thompson Street, University of Michigan
Telephone: 734/764-6595
Email: (pagregor@umich.edu)

OFFICE HOURS

By appointment.
See also the 'Chat Room' schedule

COURSE CONTENT

Methods of Survey Sampling/Applied Sampling is an applied statistical methods course, but differs from most statistics courses, though, because it is concerned almost exclusively with the *design* of data collection. Little of the analysis of collected data will be discussed in the course. The course will concentrate on problems of applying sampling methods to human populations, since survey practices are more widely used in that area, and since sampling human populations poses a number particular problems not found in sampling of other types of units. The principles of sample selection, though, can be applied to many other types of populations.

The course is presented at a moderately advanced statistical level. While we will not develop mathematical aspects of sampling theory, statistical notation and outlines of some algebraic proofs will be given. A sound background in applied statistics is necessary, since a few algebraic derivations will be presented. Little emphasis will be placed on the derivations. Nonetheless, a thorough understanding of the notation and results will be needed.

The aims of the course are to teach basic ideas of sampling from an applied perspective

and to provide experience with life-like problems. The course will cover the main techniques used in sampling practice: simple random sampling, stratification, systematic selection, cluster sampling, multistage sampling, and probability proportional to size sampling. These methods will be examined further in the context of two particular types of sample designs, area sampling and telephone sampling. The course will also cover sampling frames, cost models, sampling error estimation techniques, non-sampling errors, and compensating for missing data.

The course will be simultaneously transmitted through compressed video technology between the Joint Program in Survey Methodology at the University of Maryland, the Institute for Social Research at the University of Michigan, and the US Census Bureau in Suitland, Maryland. The three-way interactive audio-video system allows those at the transmitting site to see and hear those at the remote site, and those at the remote site to see and hear those at the transmitting site. Approximately one-third of the class sessions will be taught in College Park; the remaining sessions will be taught from the University of Michigan and transmitted to the University of Maryland. (It is yet to be determined if I will also be teaching at the Census Bureau some class sessions.) The expected dates when the instructor is scheduled to be in Maryland or Michigan are indicated in the attached syllabus.

Class sessions are once per week. Class will start at 1:00 PM each day (except for the first day when the class will start at 1:10 PM). There will be one 10-15 minute break during the class session, and class will end at 4:00 PM.

Lecture notes and other materials will be presented on projection systems in each room. Students have access to a copy of nearly all materials presented on the projectors through a course web site, although handwritten notes will be inserted in lecture frequently.

The course web site is an integral part of the course. All registered students will have access to the site through registration (at the University of Michigan) or guest log in (for University of Maryland and Census Bureau students). The web site will contain lecture notes, homework problems and related materials, homework solutions, readings, the course project and related materials, discussion items, an email log, and the chat room. Materials posted on the web site will not be distributed on paper in class (except the first class session when paper copies of the syllabus, lecture notes, and homework problem will be distributed in class).

HOMEWORK

Homework and examination are important learning devices. The homework assignments are of two types, problem sets and a project. Both types are to be turned in by the beginning of the class session when due. The regular problems will be graded on a five level system: check-plus (100), check (90), check-minus (80), late (60), not submitted (0).

Homework must be submitted electronically via the course web site Assignment tool. Each homework assignment is an attachment in the Assignment tool. Students must submit solutions, handwritten or typed, in a single .pdf format file, with name and homework set number at the top of the first page, and page numbers at the bottom of each page. The .pdf format files are to be attached to the Assignment tool homework problem on the course web site.

Files must be submitted in a standard name convention as follows: 'Surname First Initial HW #.pdf'. For example, if the instructor were to submit a solution to the first homework assignment, the file would be named 'Lepkowski J HW 1.pdf'. The submitted homework will be marked electronically and returned via the Assignment tool as an attachment, along with a copy of the solution to the homework.

Study groups are useful, and encouraged, for preparing answers to homework exercises and the project. Group answers are not acceptable. Each student must submit individual homework exercise solutions.

The project is a two-stage sampling exercise distributed early in the term, and discussed throughout the course during class sessions. The instructor will assign students to project teams. Each team will submit one copy of the project in .pdf format with name 'Team # project.pdf' (for example, 'Team A project.pdf'). All students in a team receive the same base score, from zero to 80. Each student also completes an evaluation for the other students in the team. The remainder of each student's project grade, zero to 20 points, will be based on the evaluations by fellow students in the team. Students who do not turn in evaluations of other team members will receive zero for the individual component, regardless of fellow student ratings.

EXAMINATIONS, FINAL GRADE, AND ACADEMIC INTEGRITY

There will be a short one hour in-class cumulative open book, open notes midterm examination the last hour of class on Wednesday, February 24.

The final examination will be held on Wednesday, April 28, 1-4 PM. The final exam will also be cumulative, open book, and open notes. (If this examination time presents a conflict with other regularly scheduled examination times for University of Michigan students, the students must let the instructor know as soon as possible. A course meeting once per week on Wednesday does not fit into the University of Michigan final examination schedule.)

Final grades will be a weighted composite of homework (30%), class project (35%), and examination scores (35%). The instructor may alter the relative weights, depending on overall class performance on each component of the final grade.

The course will abide by ethical standards set at the respective campuses (Census Bureau students are University of Maryland students). The University of Maryland, College Park has a nationally recognized Code of Academic Integrity, administered by the Student Honor Council. This Code sets standards for academic integrity for all undergraduate and graduate students. As a student you are responsible for upholding these standards for this course, and being aware of the consequences of cheating, fabrication, facilitation, and plagiarism. For more information on the University of Maryland Code of Academic Integrity or the Student Honor Council, please visit <http://www.studenthonorcouncil.umd.edu/whatis.html>.

TEXTBOOKS AND ASSIGNED READING

The principal text for the course will be *Survey Sampling* by Leslie Kish (John Wiley and Sons, Inc., New York, 1965). It is available at university bookstores or through online sales.

Students may find that the following texts serve as useful supplemental reading to several lecture topics: *Introduction to Survey Sampling* by Graham Kalton (Sage Publications, Beverly Hills, 1983), *Sample Survey Methods and Theory*, Volume 1, by Morris Hansen, *et al.* (New York: John Wiley and Sons, Inc., 1953), and *Sampling Techniques*, 3rd edition, by William G. Cochran (New York: John Wiley and Sons, Inc., 1977). There are also assigned readings of several papers (see list below). These will be available on the course web site.

[1] Casady, Robert J., and James M. Lepkowski. "Telephone Sampling," Chapter 15 in *Sampling of Populations: Methods and Applications*, Paul Levy and Stanley Lemeshow. New York: Wiley and Sons, Inc., 1999.

- [2] Rust, K. "Variance estimation for complex estimators in sample surveys," *Journal of Official Statistics*, 1(4) (1985): 381-397.
- [3] Kish, L. and Frankel, M. "Inference from complex samples," *Journal of the Royal Statistical Society, Series B*, **36** (1974): 1 - 37.
- [4] Kalton, G. and Kasprzyk, D. "The treatment of missing survey data," *Survey Methodology*, **12** (1986): 1 - 16.

SYLLABUS, REVISED FEBRUARY 23, 2010

The syllabus presents *approximate* dates of lecture topics, reading assignments, and homework assignments. Students must complete the assigned reading prior to the lecture. Each assignment is to be submitted by the start of the class session when due via the Assignment tool on the course web site.

Due to a winter storm in the Washington, DC, area, there will be a make up lecture Wednesday, March 3, during the University of Michigan winter break. Attendance is optional, and the class session will be recorded for viewing by students unable to attend during the break.

Week	Date	Topic	Reading ^a	Home-work	Project discussion	Lecture Site
1	Jan 6	<i>Introduction. SRS. Frames.</i>	Kish 1.0-1.7	--	--	Perry
2	12	Chat room (7:00-8:00 PM): HW 1				
	13	<i>Cluster sampling. Subsampling.</i>	Kish 2.1-2.7, 5.1	HW 1	--	Lefrak
3	20	<i>Optimal subsampling. Stratification.</i>	Kish 5.2-5.5	--	--	Perry
4	26	Chat room (7:00-8:00 PM): HW 2.				
	27	<i>Proportionate allocation. Other allocations.</i>	Kish 3.1-3.3	HW 2	--	Lefrak
5	Feb 3	<i>Paired selection and other stratification topics.</i>	Kish 3.4-3.5	--	Overview	Perry
6	9	Chat room (6:00-7:00 PM): HW 3.				
	10	<i>Class session cancelled!</i>	--	--	--	--
7	16	Chat room (6:00-7:00 PM): HW 3				
	17	<i>Systematic sampling.</i>	Kish 4.1-4.4, 6.1	HW 3	Sample size	Lefrak
8	23	Chat room (6:00-7:00				

		PM): Exam preparation.				
	24	<i>Unequal sized clusters, ratio means. Midterm.</i>	Kish 6.1-6.5	--	Optimum b*	Perry
9	Mar 3	<i>Stratified cluster sampling. PPS & PPES selection.</i>	Kish 7.1-7.5	--	--	Perry
10	9	Chat room (6:00-7:00 PM): HW 4.				
	10	<i>Area sampling.</i>	Kish 9.1-9.5	HW 4	Stratification, allocation	Lefrak
11	17	<i>Maryland Spring break. No class session.</i>	--	--	--	--
12	23	Chat room (6:00-7:00 PM): HW 5.				
	24	<i>Telephone sampling.</i>	[1]	HW 5	Sampling rate, PPeS	Lefrak
13	31	<i>Variance estimation.</i>	[2], [3]	--	1+L, estimation	Perry
14	Apr 6	Chat room (6:00-7:00 PM): HW 6.				
	7	<i>Variance estimation. Nonsampling error.</i>	[4]	HW 6	Final questions	Lefrak
15	13	Chat room (6:00-7:00 PM): Course project.				
	14	<i>Nonresponse, weighting, and imputation.</i>	Kish 13.1-13.6	Project	--	Perry
16	27	Chat room (6:00-7:00 PM): Final exam.				
	28	<i>Final exam (1:00-4:00 PM).</i>	--	--	--	Perry

^a Readings are from the textbooks by Kish, or the number of a specified paper.