Applied Sampling
SURV 625/SURVMETH 625
Spring/Winter, 2017

CLASS MEETINGS
> 1:10-4:00 PM, Wednesdays, January 4-April 26
> 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
> **Brady T. West**
> Room 4118 ISR (University of Michigan) &
> > Room 1218 Lefrak (University of Maryland)
> Telephone: 734/223-9793 (cell) and 734/647-4615 (University of Michigan)
> Email: bwest@umich.edu
> Web Site: http://www.umich.edu/~bwest

ASSISTANTS
> **Jodi Holbrook** (Video / Technical / Administrative Support)
> Room 4012 ISR (University of Michigan)
> Telephone: 734/764-6595
> Email: hjodi@umich.edu
>
> **Roberto Melipillan** (Grading)
> Room 4132 ISR (University of Michigan)
> Telephone: 734/764-4369
> Email: robmeli@umich.edu

OFFICE HOURS
> By appointment (in person, by telephone, or online / Skype / Google Hangout)

COURSE CONTENT
> Applied Sampling is an applied statistical methods course concerned almost exclusively with the *design* of data collection. Little of the analysis of collected data will be discussed, but rather the course will concentrate on problems of applying sampling methods to human populations.
> The course is presented at a moderately advanced statistical level. While mathematical aspects of sampling theory are not covered, statistical notation and some algebraic derivations will be presented. Thus, a thorough understanding of statistical notation and principles will be needed.
> 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 area sampling. The course will also cover sampling frames, cost models, sampling error

estimation techniques, non-sampling errors, and compensating for missing data.

COURSE LECTURES

This course combines pre-recorded lectures (courtesy of **Dr. James Lepkowski**, a former instructor and developer of the course) and in-class discussions. Pre-recorded lectures are available through the course Canvas web site as are associated lecture slides. **The pre-recorded lectures are to be viewed before scheduled class discussions sessions.** Live discussion sessions will include answers to questions on the pre-recorded lectures, quizzes, homework preparation, review of marked homework, and discussion of the course project.

DISCUSSION SESSIONS

Dr. West will be present each Wednesday at one of two course sites: the Joint Program in Survey Methodology at the University of Maryland (Lefrak Hall) or the Institute for Social Research at the University of Michigan (Perry Building). An interactive audio-video system connects the instructor site to remote sites.

Students may also join discussion sessions via the Bluejeans system, a cloud-based online system used to connect classrooms. Students wishing to join discussions via Bluejeans from a location other than the classrooms will be given instructions on how to connect through a web browser.

Students thus have the option of attending class in-person at one of the three classroom sites or attending via a desktop or laptop from elsewhere. If a student chooses to join class sessions via a desktop or laptop, they are required to have a separate camera (desktop or laptop built-in cameras typically have inadequate video quality) and a separate headset with microphone (desktop and laptop built-in microphones and speakers typically have inadequate audio for this kind of an application). Poor quality desktop or laptop equipment can create echoing, degrading the video or audio of all connected participants, including the regular classrooms.

Each weekly discussion session is scheduled form 1:10 PM till 4:00 PM. There will be a two- or three-item quiz covering content in the assigned pre-recorded lectures at the beginning of almost every session. Dr. West will also present brief lectures reviewing pre-recorded material and provide opportunities for students to ask questions about lectures. Dr. West will also discuss homework exercises and the course project.

HOMEWORK AND COURSE PROJECT

Nine homework assignments and a final project are planned. Each are to be turned in by the beginning of the class session when due (see syllabus below).

Homework assignments correspond with course units and are designed to aid skill development. Assignments will be graded check-plus (100 points), check (90), check-minus (80), and not submitted (0), and will ordinarily be marked and returned before the next class session. Students may request permission to submit homework late via email to the instructor, but the request must be no later than one hour before the homework is due. Permission is not guaranteed, although typically granted. If late submission is granted, there will be an agreed date and time when the late assignment must be submitted; such assignments will be graded using the specified marking system. If homework is submitted late without prior permission, scores will be check-plus (70), check (60), and check-minus (50).

The project is a semester-length exercise leading to a 10-page technical memorandum describing the student's proposed approach to a survey sampling problem set by Dr. West. The

student projects will be assigned to a team of four to five students, and each team will submit a team project. The grading will consist of two parts: 1) a team score (80 points maximum) and a student project score (up to 20 points). Each student will evaluate the contributions of each of the other team members to the final team written project. The student's teammate's evaluation scores will be averaged to generate the individual student project score. Each student must complete an evaluation for the other students on the team; students not submitting evaluations receive zero for the individual component score.

Homework is submitted electronically via the online Canvas system's Assignment tool as an attachment. Homework solutions may be handwritten or typed, and submitted in a single .pdf format file with name and homework number at the top of the first page and page numbers at the bottom of each page. Files must be submitted in a standard name convention. For homework, the name convention is 'Surname First-Initial HW #.pdf' (for example, 'West B HW 1.pdf'). For the project, the convention is 'Team name project.pdf' (for example, 'Team A project.pdf'). Failure to follow these conventions may lead to deduction of points from the assignment score. The marked submitted homework and project along with a copy of a solution will be returned via the Assignment tool.

Study groups are useful, and encouraged, for preparing answers to homework exercises. Group answers are not acceptable. That is, each student must submit individual homework solutions.

EXAMINATIONS & FINAL GRADE

There will be a two-hour, in-class, cumulative, open book / open notes midterm examination on **Wednesday, February 22**. The two-hour open book / open notes final examination will be held on **Wednesday, April 26**. (If this examination time conflicts with other regularly scheduled examination times for University of Michigan students, students must inform Dr. West 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 quizzes (10%), homework (25%), class project (25%), and examination scores (45%).

CLASS ATTENDANCE

Students are required to attend all live class sessions. A student anticipating missing a class due to illness or other reason must notify Dr. West via email at least 30 minutes before the session the reason for missing the session. University of Maryland (including those at Census) students should examine University attendance policies at http://www.faculty.umd.edu/teach/attend_student.html. University of Michigan students should review http://www.lsa.umich.edu/facstaff/saa/studentclassattendance.

Each student is permitted up to two excused absences for class without penalty. Absences after two excused will cause deduction of points from the course average score. Each student is also permitted to miss up to two quizzes. The quiz contribution to the grade will be based on the average of the eight best quiz scores; missed quizzes will have a score of zero.

COURSE WEB SITE

All registered students have access to a course web site (Canvas) located at the University of Michigan. Guest log-in's will be arranged for University of Maryland (and Census Bureau) students. The Canvas web site contains lecture notes, homework problems and solutions,

readings, course project materials, and a chat room. Materials will be posted on the web site and not distributed as paper copies.

ACADEMIC INTEGRITY

The course follows ethical standards at respective campuses. The University of Maryland nationally recognized Honor Code (http://www.president.umd.edu/policies/iii100a.html) is administered by the Student Honor Council. The Student Honor Council statement *I pledge on my honor that I have not given or received any unauthorized assistance on this assignment/examination* ordinarily should be *handwritten* and signed on the front cover of all papers, projects, or other academic assignments submitted for evaluation in this course. The instructor waives this requirement for Maryland (including Census) students, but not the other provisions of the Honor Code. The student name on a document is the student's pledge that this work is theirs and theirs alone. University of Michigan students are expected to follow by academic integrity policies in the University of Michigan Standard Practice Guide (http://spg.umich.edu/pdf/303.03.pdf). Violation of the Maryland Honor Code pledge or the University of Michigan policy is a serious matter and may lead to a failing mark on a submission or for the entire course.

STUDENT COURSE EVALUATION

Both the University of Maryland (via TESTUDO) and University of Michigan (via CTools) have online student course evaluation systems. All students are strongly urged to use the respective university online course evaluation system at the end of the term. Since course evaluations will only be available to other students when at least 70% of registered students complete the evaluation, it is very important to future students that all enrolled students log on at the end of the semester to complete a course evaluation.

TEXTBOOKS & 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 other texts useful supplemental reading, including *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). Selected assigned readings will also be available on the course web site:

[1] Rust, K, and Rao, J.N.K. "Variance Estimation for Complex Surveys Using Replication Techniques," *Statistical Methods in Medical Research*, Vol. 5, 1996, pp. 283-310.
[2] Kish, L. and Frankel, M. "Inference from complex samples," *Journal of the Royal Statistical Society*, Series B, **36** (1974): 1 - 37.
[3] American Association for Public Opinion Research. *Standard Definitions:Final Disposition of Case Codes and Outcome Rates for Surveys*. AAPOR, 2016.
[4] Kalton, G. and Kasprzyk, D. "The treatment of missing survey data," *Survey Methodology*, **12** (1986): 1 - 16.

COMPETENCY

The student will learn and will be tested on the following competencies in this course:

1. Learn the meaning and application of expressions for sampling variance of means and proportions under a variety of finite population sampling techniques.
2. Understand the properties of and how to apply various sampling techniques, including simple random sampling, stratification and stratified random sampling, cluster sampling, systematic sampling, two-stage sampling, multistage sampling, probability proportionate to size sampling, and stratified multistage sampling.
3. Understand sampling variance estimation techniques for means and proportions for each of the sampling techniques in the course, including procedures for nonlinear statistics using the Taylor series expansion technique, as well as the balanced repeated replication and jackknife repeated replication techniques.
4. Learn how nonresponse can affect survey estimates, and what techniques can be used to reduce nonresponse and compensate through weights and imputation.

SYLLABUS

The syllabus below presents approximate dates of lecture topics. Students must complete assigned readings prior to the Wednesday lecture. Each homework assignment is to be submitted via the Canvas web site Assignment tool by the start of the class session when due.

Week Date(s)	Time	Activity	Location	Subject (Reading)

1 Jan 4	1:10 PM	Lecture [a]	Classroom[b]	Unit 1: Preliminaries. *(Live)*
				    1. Course description (Syllabus, Canvas Syllabus tool).
				    2. Course perspectives (Kish 1.0-1.7).

1 Jan 4 to Jan 11	Anytime	Lecture	MediaSite/ At home	Unit 2: Elements. *(Pre-recorded lectures)*
				    3. Simple random sampling (Kish 2.1-2.6).
				    4. Three frame problems (Kish 2.7, 11.1-11.2).

2 Jan 11	1:00 PM	Homework	Canvas	HW 1 & Project released.
	1:10 PM	Discussion	Classroom or Bluejeans	Student connection tests (Bluejeans). Quiz 1. Preview HW 1. Review lectures 1, 2, 3, & 4. Preview lectures 5 & 6.

2 Jan 11 to Jan 18	Anytime	Lecture	MediaSite/ At home	Unit 2: Elements (*Pre-recorded lectures*)
				    5. Frames: Clusters (Kish 11.3).
				 Unit 3: Clusters (*Pre-recorded lectures*)
				    6. Cluster sampling (Kish 5.1-5.2).

3 Jan 18	1:00 PM	Homework	Canvas	**Submit HW 1** via Assignment tool. Homework 2 released.
	1:10 PM	Discussion	Classroom or Bluejeans	Quiz 2. Preview HW 2. Review lectures 5 & 6. Preview lectures 7 & 8. Project introduction.

3 Jan 18 to Jan 25	Anytime	Lecture	MediaSite/ At home	Unit 3: Clusters (*Pre-recorded lectures*)
				    7. Two-stage sampling (Kish 5.3-5.4)
				    8. Subsample size (Kish 8.3)

Week Date(s)	Time	Activity	Location	Subject (Reading)
4 Jan 25	1:00 PM	Homework	Canvas	**Submit HW 2** via Assignment tool. HW 3 released.
	1:10 PM	Discussion	Classroom or Bluejeans	Quiz 3. Review HW 1 solution & preview HW 3. Review lectures 7 & 8. Preview lectures 9 & 10. Project teams & target sample sizes.

4 Jan 25 to Feb 1	Anytime	Lecture	MediaSite/ At home	Unit 4: Stratification (*Pre-recorded lectures)*
				    9. Stratified random sampling (Kish 3.1-3.3)
				    10. Allocations (Kish 3.4-3.5)

5 Feb 1	1:00 PM	Homework	Canvas	**Submit HW 3** via Assignment tool. HW 4 released.
	1:10 PM	Discussion	Classroom[c] or Bluejeans	Quiz 4. Review HW 2 solution & preview HW 4. Lecture 9 & 10 questions. Preview lectures 11 & 12. Project subsample size.

5 Feb 1 to Feb 8	Anytime	Lecture	MediaSite/ At home	Unit 4: Stratification (*Pre-recorded lectures)*
				    11. Stratification topics (Kish 3.6, 4.5A)
				Unit 5: Systematic selection (*Pre-recorded lectures)*
				    12. Systematic selection (Kish 4.1-4.3)

6 Feb 8	1:00 PM	Homework	Canvas	**Submit HW 4** via Assignment tool. HW 5 released.
	1:10 PM	Discussion	Classroom or Bluejeans	Quiz 5. Review HW 3 solution & preview HW 5. Review lectures 11 & 12. Preview lectures 13 & 14. Project achieved precision.

6 Feb 8 to Feb15	Anytime	Lecture	MediaSite/ At home	Unit 6: Unequal sized clusters (*Pre-recorded lectures)*
				    13. Unequal sized cluster sampling (Kish 6.1-6.3)
				    14. Stratified unequal sized clusters (Kish 6.4)

Week Date(s)	Time	Activity	Location	Subject (Reading)
7 Feb 15	1:00 PM	Homework	Canvas	**Submit HW 5** via Assignment tool. HW 6 released.
	1:10 PM	Discussion	Classroom or Bluejeans	Quiz 6. Review HW 4 & 5 solutions. Review lectures 13 & 14. Preview lectures 15, 16, & 17. Review. Project sampling rates.
7 Feb 15 to Feb 22	4:00 PM	Anytime	MediaSite/ At home	Unit 6: Unequal sized clusters (*Pre-recorded lectures*) 15. Complex designs (Kish 6.5)
8 Feb 22	1:10 PM	**Midterm exam**	Classroom[b]	

9 Mar 1 to Mar 8	Anytime	Lecture	MediaSite/ At home	Unit 6: Unequal sized clusters (*Pre-recorded lectures*)
				    16. PPS (Kish 7.1-7.3)
				    17. PPeS (Kish 7.4-7.5)

10 Mar 8	1:10 PM	Discussion	Classroom or Bluejeans	Quiz 7. Review HW 5 solution & preview HW 6. Review lectures 15, 16 & 17. Preview lectures 18, 19, & 20. Project domains, PPeS.

10 Mar 8 to Mar 15	Anytime	Lecture	MediaSite/ At home	Unit 7: Area sampling (*Pre-recorded lectures*)
				    18. Two-stage (Kish 9.1-9.6, 9.8)
				    19. Three-stage (Kish 10.1)
				    20. Multistage (Kish 10.2-10.3)

11 Mar 15	1:00 PM	Homework	Canvas	**Submit HW 6** via Assignment tool. HW 7 released.
	1:10 PM	Discussion	Classroom or Bluejeans	Preview HW 7. Review lectures 18, 19, & 20. Preview lectures 21, 22, & 23. Project 2nd stage selection & subclasses.

11 Mar 15 to Mar 22	Anytime	Lecture	MediaSite/ At home	Unit 8: Variance estimation (*Pre-recorded lectures*)
				    21. Forming computing units (Kish 8.6)
				    22. Taylor series expansion ([1] Rust & Rao)

Week Date(s)	Time	Activity	Location	Subject (Reading)
				23. Replicated sampling (Kish 4.4)
12 Mar 22	1:00 PM	Homework	Canvas	**Submit HW 7** via Assignment tool. HW 8 released.
	1:10 PM	Discussion	Classroom or Bluejeans	Quiz 8. Preview HW 8. Review lectures 21, 22, & 23. Preview lectures 24, 25, & 26. Project 2nd stage selection & subclasses.

12 Mar 22 to Mar 29	Anytime	Lecture	MediaSite/ At home	Unit 8: Variance estimation (*Pre-recorded lectures)*
				    24. Balanced repeated replication ([1] Rust & Rao)
				    25. Jackknife repeated replication ([2] Kish & Frankel)
				    26. Estimation & presentation (Kish 14.1, 14.2)

13 Mar 29	1:00 PM	Homework	Canvas	**Submit HW 8** via Assignment tool. HW 9 released.
	1:10 PM	Discussion	Classroom or Bluejeans	Quiz 9. Preview HW 9. Review lectures 24, 25, & 26. Preview lectures 27, 28, & 29. Project: describing estimation procedures.

13 Mar 29 to Apr 5	Anytime	Lecture	MediaSite/ At home	Unit 9: Non-sampling error (*Pre-recorded lectures)*
				    27. Total survey error (Kish 13.1-13.3)
				    28. Response error
				    29. Nonresponse (Kish 13.4-13.6; [3] AAPOR)

14 Apr 5	1:00 PM	Homework	Canvas	**Submit HW 9** via Assignment tool.
	1:10 PM	Discussion	Classroom[c] or Bluejeans	Quiz 10. Review HW 8 solution. Review lectures 27, 28, & 29. Preview lectures 30 & 31.
14 Apr 5 to Apr 12	Anytime	Lecture	MediaSite/ At home	Unit 9: Non-sampling error (*Pre-recorded lectures)* 30. Weighting ([4] Kalton & Kasprzyk)

Week Date(s)	Time	Activity	Location	Subject (Reading)
				31. Imputation
15 Apr 12	1:10 PM	Discussion	Classroom or Bluejeans	Quiz 11. Review HW 9 solution. Review lectures 30 & 31. Final exam questions.
16 Apr 19	1:00 PM	Homework	Canvas	**Submit Project & Team Evaluations** via Assignment tool.
	1:10 PM	Discussion	Classroom or Bluejeans	**Optional course review.**
17 Apr 26	1:10 PM	**Final exam**	Classroom[b]	

[a]Recorded and available for viewing after 24-48 hours on MediaSite.
[b]1208 Lefrak/T10 Census/G300 ISR-Perry.
[c]Instructor at JPSM, 1208 Lefrak Hall, University of Maryland.