SOCI 902: Analysis of Complex Survey Data  
Spring 2020  
Monday, 2:30 PM – 5:00 PM  
707 Oldfather Hall

Instructor: Dr. Kristen Olson  
Office: 703 Oldfather Hall  
Office Phone: (402) 472-6057  
e-mail: kolson5@unl.edu  
Office hours: Monday 11:00 AM-12:00 PM, Tuesday 3:30-4:30 PM, or by appointment

Prerequisites: An introductory class to survey research and statistical analysis. A class on survey sampling is not required. Experience with both linear regression and generalized linear models will make this class much easier.

A. Overview of Course

This course covers advanced topics related to sampling error in surveys. This course has three main components. First, we will review different complex sample designs used to measure populations of humans. Second, we will examine the implications of complex sample designs for analytic statistics such as linear regression coefficients, contingency tables, and logistic regression coefficients and their standard errors. Taylor Series and replication approaches to variance estimation will be covered. Finally, we will look at the effect of failure to measure part of the population on particular items (item nonresponse) or the complete failure to measure a sampled person (unit nonresponse) on analysis, and methods available to “repair” the missing information (imputation and weighting). These problems affect estimates made from surveys and proper adjustments to estimates of sampling error and to the data themselves are often necessary for valid inference. Familiarity with conducting statistical analyses in a syntax-based statistical software program (e.g., SAS, Stata) is highly recommended before taking this course.

B. Goals of the Course

The course assumes that students understand basic analysis techniques under the assumption of a simple random sample. This class assumes that students do not necessarily know the implications of sample design on linear estimators (e.g., means, proportions, totals) and complex estimators (e.g., regression coefficients, contingency tables) or understand the statistical underpinnings of what happens and what to do when things go awry during data collection (e.g., unit and item nonresponse). This course covers common sample designs used in studying populations, analysis of complex survey data under ideal circumstances, and analysis of complex survey data under the presence of nonresponse. The class will focus on sample designs for household surveys. However, the constructs can be applied to surveys of establishments and of special populations.

In particular, this course has five main goals.

1. Identify common sample designs used to measure populations of households or persons.
2. Examine the implications of sample design and nonresponse on complex estimators such as correlation coefficients, regression coefficients, and generalized linear models.
   a. Implications for point estimates
   b. Implications for variance estimates
   c. Use of appropriate analytic methods and software packages to properly account for a complex sample design
3. Explore the conditions under which unit and item nonresponse affects survey estimates.
   a. Implications for linear estimators
   b. Implications for nonlinear estimators
   c. Weighting and imputation methods to adjust for nonresponse in surveys
4. Make statistical writings more accessible and understandable.
5. Communicate findings from analyses of complex survey data in written and verbal presentation.

C. Format of the Course

The course will be comprised of readings, lectures and discussion. Students are expected to come to class with questions on assigned readings.

It is understood by the instructor that students in this course will vary in their statistics background, exposure to mathematical statistics, and other courses in survey sampling. As such, the assigned readings are of two types – standard readings and readings of a more advanced, technical nature. The “Technical Readings” are intended to provide a more advanced discussion of the week’s topic. All students are encouraged to attempt the technical readings; Ph.D. students interested in specializing in quantitative methods or students who are interested in furthering their statistical knowledge beyond the level discussed in the course lectures are strongly encouraged to do the technical readings.

D. Homework Assignments

This class will consist of participation, homework assignments and a final project.

Homework assignments and paper assignments will be given on alternating weeks, starting with the second week of class. Each assignment will be due the following Monday by the beginning of class. Late assignments will be deducted by one letter grade for each day that they are late. Because there will be homework or a paper assignment each week, no extensions will be given on homework assignments except for excused absences. Late assignments will be deducted by 10 points for each day that they are late. Assignments handed in on Monday after class will be considered one day late, on Tuesday will be considered two days late, and so on. All assignments that are not received within five business days (by the end of the day on Friday) will receive a grade of zero. Details on each assignment will be given during class and/or posted on Canvas.

The homework assignments will focus on application of the methods discussed in class to actual data sets. They are designed to acquaint you with the methods and software for each topic of the course.

Homework assignments must be typed. Students are required to use equation editors for accurate notation. All work and final answers must be neat and clearly labeled. All answers in which more than one answer is provided or the final answer is ambiguous will be marked as incorrect. All assignments must be handed in via hard copy to the instructor at the beginning of class, unless otherwise noted. Homework assignments will be graded on a three-point scale, +, , and -, corresponding to 100, 90 and 80 points, respectively. Although study groups are permitted to facilitate understanding the material, all students are required to turn in their own homework assignments. Copying another student’s homework or handing in another student’s homework as your own (including with minor changes) is considered cheating, per the UNL Student Code of Conduct (http://stuafs.unl.edu/dos/code#rules).

This class requires students to write statistical findings for a scientific audience. Students who need writing assistance are encouraged to visit the UNL writing center: http://www.unl.edu/writing/

E. Academic Honesty

Cheating and plagiarism will not be tolerated. Both cheating and plagiarism are violations of UNL’s Student Code of Conduct. Any assignment, exam, or project in which cheating or plagiarism or any other form of academic misconduct is identified will result in immediate failure of the assignment and, depending on the scope of the assignment, may result in immediate failure of the class. These acts of cheating, plagiarism, or any other
violations of academic integrity will be reported to the Sociology graduate chair, Sociology department chair, and to the Office of the Dean of Students, as detailed in the UNL Sociology Graduate Handbook. The instructor reserves the right to use Turnitin or other plagiarism detection tools to help in assessing the risk of plagiarism. For information on plagiarism and what constitutes acceptable and unacceptable citations, please see the Graduate Studies website at [http://www.unl.edu/gradstudies/current](http://www.unl.edu/gradstudies/current), [https://www.unl.edu/gradstudies/current/integrity](https://www.unl.edu/gradstudies/current/integrity) and [http://www.unl.edu/gradstudies/current/integrity#plagiarism](http://www.unl.edu/gradstudies/current/integrity#plagiarism)

Class notes will be placed on Canvas before each class period. Students are responsible for bringing an electronic or printed copy of the class notes to each class. The instructor will not make copies of the notes.

**F. Paper Assignments and Final Project**

The final project will be a piece of independent empirical research conducted using a data set of your choice, applying the methods discussed here. The data must have a complex sample design. The final paper may be substantive or methodological in nature. Students are strongly encouraged to meet with Dr. Olson to discuss their topics and datasets.

To assist in writing the paper, there are intermediate deadlines throughout the term for portions of the paper. Paper assignments will be graded using letter grades (e.g., A+, A, A-). The following scores will be assigned to each letter grade:

<table>
<thead>
<tr>
<th>Letter Grade</th>
<th>Numeric Score</th>
<th>Letter Grade</th>
<th>Numeric Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>A+</td>
<td>100</td>
<td>C</td>
<td>71</td>
</tr>
<tr>
<td>A</td>
<td>95</td>
<td>C-</td>
<td>68</td>
</tr>
<tr>
<td>A-</td>
<td>91</td>
<td>D+</td>
<td>65</td>
</tr>
<tr>
<td>B+</td>
<td>88</td>
<td>D</td>
<td>61</td>
</tr>
<tr>
<td>B</td>
<td>85</td>
<td>D-</td>
<td>58</td>
</tr>
<tr>
<td>B-</td>
<td>78</td>
<td>F</td>
<td>0</td>
</tr>
<tr>
<td>C+</td>
<td>75</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The final paper will be between 15 and 20 pages in length (excluding references), double spaced, Times New Roman 12 pt. font with 1 inch margins. All students will conduct a brief oral presentation of their papers in class. Your final presentation grade will be determined by evaluations from Dr. Olson and your classmates. More information on the paper assignments and the presentations will be distributed throughout the semester.

**G. Participation**

Participation in class will make the course a better experience for everyone. Class attendance is expected. If you choose to or are not able to attend class, you are expected to get notes from a fellow student in the class. In the case of inclement weather, we will hold class at the scheduled time unless the University is closed. Students are required to have completed the readings before the class meeting time. Attendance alone is not sufficient to earn a full participation grade.

The class schedule below is subject to change. Changes will be announced in class and/or on the class web page. Not knowing about syllabus changes, including changes in assignments, because of class absence or for not checking the class Canvas website are not legitimate excuses for failure to complete the course requirements.
H. Grading

Grades will be constituted as follows:

<table>
<thead>
<tr>
<th>Assignment</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homework assignments</td>
<td>25%</td>
</tr>
<tr>
<td>Paper assignments</td>
<td>25%</td>
</tr>
<tr>
<td>Final Presentation</td>
<td>20%</td>
</tr>
<tr>
<td>Final Paper</td>
<td>20%</td>
</tr>
<tr>
<td>Participation</td>
<td>10%</td>
</tr>
<tr>
<td>Total</td>
<td>100%</td>
</tr>
</tbody>
</table>

Final grades will be assigned as follows:

<table>
<thead>
<tr>
<th>Grade</th>
<th>Weighted Percentage</th>
<th>Grade</th>
<th>Weighted Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>A+</td>
<td>100-97</td>
<td>C</td>
<td>72.9-70</td>
</tr>
<tr>
<td>A</td>
<td>96.9-93</td>
<td>C-</td>
<td>69.9-67</td>
</tr>
<tr>
<td>A-</td>
<td>92.9-90</td>
<td>D+</td>
<td>66.9-63</td>
</tr>
<tr>
<td>B+</td>
<td>89.9-87</td>
<td>D</td>
<td>62.9-60</td>
</tr>
<tr>
<td>B</td>
<td>86.9-80</td>
<td>D-</td>
<td>59.9-55</td>
</tr>
<tr>
<td>B-</td>
<td>79.9-77</td>
<td>F</td>
<td>0.0-54.9</td>
</tr>
<tr>
<td>C+</td>
<td>76.9-73</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

I will not give extra credit and will not change grade distributions to ‘curve’ the class. Students may take the course pass/no pass. If taking the class pass/no pass, students must earn a B or better grade to earn a “pass.”

I have a 24/7 grade policy – please wait 24 hours after receiving a grade to talk to me about any concerns that you may have, and do so within 7 days of receiving the grade.

Grade appeals must be made in writing. The procedure for grade appeals is the following: (1) Provide the instructor with your grade appeal in writing, including documentation to support the claim, within two weeks of the grade being given. (2) The instructor will regrade your assignment. Grades may go up or down on regrading. (3) If you still would like to pursue a grade appeal after the regrading, provide written documentation to the Graduate Chair of the Sociology Department. More information about grade appeals can be found here: [http://cas.unl.edu/grading-appeals](http://cas.unl.edu/grading-appeals) and [http://www.unl.edu/gradstudies/bulletin/graduate-grade-appeals](http://www.unl.edu/gradstudies/bulletin/graduate-grade-appeals)

I. Technology and other distractions policy

The use of cell phones is not allowed in class. Laptops or tablets for class purposes are permitted, but not for use of non-class related websites (this includes checking email during class). Turn off all cell phones before class starts. Turn off all cell phones before class starts. One accidental ring of a cell phone will be permitted per semester; after this, the student’s participation grade will have 3 points deducted for each ring, text, or other use of the phone or electronic device in class. All newspapers, magazines, or any other material other than that necessary for this class also must be put away when entering the classroom. Any student who is seen reading a newspaper, magazine, or anything not related to this class will receive an automatic 3 point deduction from his/her participation grade for each use of this material.

All students will be expected to use statistical analysis software appropriate for complex sample designs. During some weeks, specific software packages (e.g., WesVar, IVEWare) will be required. During other weeks, students may use the survey analysis procedures in SAS or Stata. We will meet in a computer laboratory so that all
students can use this type of software. If students have laptop computers, then students may bring these computers during hands-on ‘lab’ time. More information on this laboratory time will be distributed throughout the semester.

J. Office hours and e-mail

Office hours will be held Monday 11:00 AM-12:00 PM, Tuesday 3:30-4:30 PM at Dr. Olson’s office, 703 Oldfather Hall. Appointments for in-person meetings with Dr. Olson can be scheduled by contacting her via e-mail (kolson5@unl.edu; preferred) or before class.

All e-mails from Dr. Olson to the class will be conducted through Canvas. It is your responsibility to ensure that the e-mail address in Canvas is up-to-date. It is also your responsibility to ensure that other class members who need your e-mail address have your preferred e-mail address.

K. Accommodations for students with disabilities

The University strives to make all learning experiences as accessible as possible. If you anticipate or experience barriers based on your disability (including mental health, chronic or temporary medical conditions), please let me know immediately so that we can discuss options privately. To establish reasonable accommodations, I may request that you register with Services for Students with Disabilities (SSD). If you are eligible for services and register with their office, make arrangements with me as soon as possible to discuss your accommodations so they can be implemented in a timely manner. SSD contact information: 232 Canfield Admin. Bldg.; 402-472-3787
**Required Readings and Due Dates.** Schedule and assignments subject to change. Updates will be posted on Canvas and/or announced in class.

**Required Texts**

- Book website: [http://www.isr.umich.edu/src/smp/asda/](http://www.isr.umich.edu/src/smp/asda/)


**Recommended Texts**

- This book is out of print. However, it remains one of the main references for the technical details behind analyzing complex survey data. It assumes knowledge about sampling. Many of the technical readings come from this book.

- This book is the key reference for all things variance estimation related. It assumes knowledge of sampling.

**January 13 – Introduction, Overview of Sample Designs (Self-assessment)**

HWB: Chapter 1


**January 20 – MLK Day, No class**

**January 27 – Review of Complex Sample Designs, continued (Homework assignment 1)**

HWB: Chapter 2


Feb. 3 – Weighting, Part 1 (Paper assignment 1)


Feb. 10 – Weighting, part 2 (Homework assignment 2)

Statistical Theory:


How-to in Stata:


Read one of the following two applications papers:


**Feb. 17 – Design-based inference, SECUS, and Variance Estimation (Paper assignment 2)**

HWB: Chapter 3-4


**Feb. 24 – Descriptive and Bivariate Statistics: Continuous Data (Homework assignment 3)**

HWB: Chapter 5


**March 2 – Descriptive and Bivariate Statistics: Categorical Data (Paper assignment 3)**

HWB: Chapter 6


**March 9 – Linear Regression, interaction effects and diagnostics (Homework assignment 4)**
HWB: Chapter 7


March 16 – Logistic and Multinomial Logistic Regression (Paper assignment 4)

HWB: Chapter 8


March 23 – No Class, Spring Break

March 30 – Ordinal and Count Models; Event History Models (Homework assignment 5)

HWB: Chapter 9, 10


April 6 – Imputation, Part 1 (Paper assignment 5)


Read one of the two following articles:


April 13 – Imputation, Part 2 (Homework assignment 6)

HWB, Chp. 12

Overview articles – Read Johnson and Young and one of Schafer/Graham or Raghunathan:


Sequential Regression / Chained Equations – Read one of the following:


Applications:


April 20 – Advanced Topics: Longitudinal Analysis, Bayesian analysis, Multilevel models, Structural Equation Models, Propensity Score Methods, Special Cases of Missingness (Paper assignment 6, Final presentation guidelines distributed)

HWB: Chp. 11, 13


April 27 – Presentation of Final Projects

Monday, May 4, 5:00 PM – Final Project Due