Mel and Enid Zuckerman College of Public Health  
University of Arizona

SYLLABUS  
BIOS 576C Applied Biostatistics Analysis  
FALL 2020  
3 units

**Time:**  Monday and Wednesday 10:30 am – 11:45 am  
**Location:** online, zoom link is on D2L homepage. Note that classes will be recorded and will be available on D2L.  
**Instructor and Contact Information:** Dr Melanie Bell, Drachman A240, 1295 Martin Ave, Tucson AZ 85724, 626-2795 MelanieBell@email.arizona.edu  
**Teaching Assistant (5 hr/wk):** Phil Jones, pajones@email.arizona.edu  
**Instructor Availability:** Office hours by appointment; please email to request. Weekday email will usually be responded to within 24 hours.  

**Catalog Description:** Integrate methods in biostatistics (e.g., BIOS 576A, B) and epidemiology (e.g., EPID 573A) to develop analytical skills in a health research project setting.

**Course Description:** This course provides opportunities to apply and extend statistical concepts learned in previous courses to analyze data from epidemiological, clinical and experimental studies.

**Course Prerequisites:**  
Biostatistics:  
EPID/CPH/BIOS 576A (Biostatistics in Public Health)  
EPID/CPH/BIOS 576B (Biostatistics for Research)  

Epidemiology:  
EPID/CPH 573A (Basic Principles in Epidemiology)  
Or permission of instructor.

**Course Objectives and Expected Learning Outcomes:**

- **Learning Outcomes**  
1. To gain practical experience in statistical analysis, including identifying the appropriate scientific question, determining the appropriate method(s) of analysis, developing an analysis plan, and performing an appropriate statistical analysis.  
2. To extend basic statistical modeling approaches to address complex survey data, analyzing data from repeated measures experiments, and assessment of missing data.  
3. To develop effective oral and written presentations of statistical results; teamwork and collaboration; and ability to interact with a variety of collaborators.

- **Learning Outcomes (Competencies Obtained):** Upon completion of this course students should be able to:  
1. Undertake appropriate statistical analyses that align with the scientific question  
2. Perform sensitivity analysis to assess robustness of results to analytical assumptions  
3. Manage real data; use a data dictionary; restructure and/or reformat data in order to perform analyses  
4. Engage in collaborative research with scientific mentors and peers  
5. Deliver oral and written communication skills for disseminating scientific and statistical findings

**Biostatistics MPH Competencies Covered:** At the end of the course, you should be able to:

<table>
<thead>
<tr>
<th>Competency</th>
<th>Assessed</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Ability to identify appropriate statistical tools to address specific scientific questions.</td>
<td>Projects1-3, Final project</td>
</tr>
</tbody>
</table>
B. Demonstrate excellent presentation skills and the ability to explain statistical concepts and findings to a general scientific audience.  

C. Demonstrate skills in data management to handle a variety of practical problems in data format and structure.  

<table>
<thead>
<tr>
<th>Competency</th>
<th>Assessed</th>
</tr>
</thead>
<tbody>
<tr>
<td>D. Demonstrate the ability to skillfully engage in statistical collaboration with mentors, colleagues, and clients.</td>
<td>Projects1-3, Final project</td>
</tr>
<tr>
<td>E. Suggest preferred methodological alternatives to commonly used statistical methods when assumptions are not met.</td>
<td>Projects1-3, Final project</td>
</tr>
</tbody>
</table>

**Biostatistics MS Competencies Covered:** At the end of the course, you should be able to:

**Course Notes:** A webpage has been created for this class using the Desire 2 Learn (D2L) interface. The course website contains the syllabus, class notes, and project assignments.

**Recommended Texts:**


**Additional resources:**


**Required or Special Materials :**

**Computing:** You will be required to use a code based statistical software package for this class (i.e., no point and click programs). You may use SAS, Stata or R. I will be giving some software support, although part of the objectives of this class is to improve your abilities in statistical software, which includes figuring out how to do things on your own. I am familiar with each of the packages, with my expertise following the order they are listed.

R is freeware. Stata is available in various sizes for students at reduced prices. SAS has been $95/year available from the bookstore, but there is now a student edition, SAS University Edition. [http://www.sas.com/en_us/software/university-edition.html](http://www.sas.com/en_us/software/university-edition.html). The University edition is not all of SAS, but it has nearly everything that most of us use: Base SAS, STAT, IML, Studio and Access.

Since you will be using data, I will need copies of your **CITI biomedical research training certificate**. Please upload to D2L.

**Course Requirements:**

This class will be project based, with three group projects and one final project.

**Individual project**

Each student must acquire data to use in an individual project. These projects will include an initial consult with the investigator you will be working with and ideally, the instructor; a detailed written statistical analysis plan (SAP), which must be read and approved by the investigator; an oral presentation; and a written report. The initial consult will be about an hour, and will be used to discuss the proposed project. The oral presentation can be no longer than 10-15 minutes (time to be determined later in the semester), after which there will be 5 minutes for questions. The final written project should be no longer than 2500 words, and have no more than 4 tables and 2 figures. Students should consider writing this as a submit-able manuscript (and submit it!). See guidance at the end of the document, and the tips on D2L.

**Group projects**

There will be three group projects. You will be assigned to a group of about 4 students that you will work with on the project. Every person will be in a group that gives one presentation: either a statistical analysis plan (with preliminary descriptive statistics) or a progress report. For each project, you will develop a statistical analysis plan, perform an appropriate data analysis, and prepare a written report. Each **student** will submit interim analysis, descriptive statistics and primary analysis, for each of the projects before the report is due. The descriptive statistics should include a shell table (i.e., empty), stratified appropriately. You should then include the output from your statistical software that you will use to populate that table. The primary analysis will be an attempt at the primary analysis, but does not need to be the final analysis that is included in the group report. Each **group** will submit a written
report of the final analysis for each project (due by 5 pm on the respective due date). Please turn in assignments as a Word document. See guidance at the end of the document, and the tips on D2L.

Grading/Student Evaluation:
Group projects (60%): Of this 60%, each of the 3 projects is worth 30% (25% for the group reports, 5% for interim analysis (descriptive statistics, primary analysis) for each student). The presentation is worth 10%. Students will be asked to assess the contributions of their group members, and this will be taken into consideration in each individual’s grade, with up to 5 percentage points being marked off the final grade.

Final project (40%): Of this 40%, the consultation will be worth 5%, the analysis plan 10%, descriptive statistics 5%, oral presentation 10% and the written report 70%.

Attendance & in-class: although this semester attendance is not being counted as part of the grade, I urge you to attend. Much of this class is based on practical hands-on analysis, and most students find that they learn a lot during these sessions.

Late work will be accepted up to 2 days after the due date, but at a 50% markdown in score. Please turn your work in on time.

<table>
<thead>
<tr>
<th>Assessment Methods</th>
<th>Individual or group</th>
<th>Due Dates</th>
<th>Points</th>
<th>Competency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project 1 Descriptive statistics</td>
<td>Individual</td>
<td>9/10</td>
<td>1.5</td>
<td>A</td>
</tr>
<tr>
<td>Project 1 Primary analysis</td>
<td>Individual</td>
<td>9/17</td>
<td>1.5</td>
<td>A</td>
</tr>
<tr>
<td>Project 1 Report</td>
<td>Group</td>
<td>9/25</td>
<td>15</td>
<td>A-E</td>
</tr>
<tr>
<td>Final project consultation</td>
<td>Individual</td>
<td>10/2</td>
<td>2</td>
<td>D</td>
</tr>
<tr>
<td>Project 2 Descriptive statistics</td>
<td>Individual</td>
<td>10/8</td>
<td>1.5</td>
<td>A</td>
</tr>
<tr>
<td>Project 2 Primary analysis</td>
<td>Individual</td>
<td>10/15</td>
<td>1.5</td>
<td>A</td>
</tr>
<tr>
<td>Final project statistical analysis plan</td>
<td>Individual</td>
<td>10/16</td>
<td>4</td>
<td>A, E</td>
</tr>
<tr>
<td>Project 2 Report</td>
<td>Group</td>
<td>10/23</td>
<td>15</td>
<td>A-E</td>
</tr>
<tr>
<td>Final project descriptive statistics</td>
<td>Individual</td>
<td>10/30</td>
<td>2</td>
<td>A</td>
</tr>
<tr>
<td>Project 3 Descriptive statistics</td>
<td>Individual</td>
<td>11/5</td>
<td>1.5</td>
<td>A</td>
</tr>
<tr>
<td>Project 3 Primary analysis</td>
<td>Individual</td>
<td>11/12</td>
<td>1.5</td>
<td>A</td>
</tr>
<tr>
<td>Project 3 Report</td>
<td>Group</td>
<td>11/20</td>
<td>15</td>
<td>A-E</td>
</tr>
<tr>
<td>Group project oral report</td>
<td>Group</td>
<td>TBA</td>
<td>6</td>
<td>A</td>
</tr>
<tr>
<td>Final project oral report</td>
<td>Individual</td>
<td>TBA</td>
<td>4</td>
<td>B</td>
</tr>
<tr>
<td>Final project written report</td>
<td>Individual</td>
<td>12/15</td>
<td>28</td>
<td>A-E</td>
</tr>
</tbody>
</table>
Tentative Course Schedule. See D2L for updates (content>schedule)

<table>
<thead>
<tr>
<th>Week</th>
<th>Date</th>
<th>Topic</th>
<th>Due</th>
<th>Suggested reading</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>M: 8/24</td>
<td>Introduction to class</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>W: 8/26</td>
<td>Data cleaning &amp; exploratory analysis group work</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>M: 8/31</td>
<td>Analysis plans, reporting</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>W: 9/2</td>
<td>Logistic regression</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>M: 9/7</td>
<td>LABOR DAY, NO CLASS</td>
<td>9/10 Project 1 descriptive stats</td>
<td></td>
</tr>
<tr>
<td></td>
<td>W: 9/9</td>
<td>Logistic regression group work</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>M: 9/14</td>
<td>Predictor selection</td>
<td>9/17 Project 1 primary analysis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>W: 9/16</td>
<td>Group 1 – Analysis plan presentation (project 1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>M: 9/21</td>
<td>STROBE review with example</td>
<td>9/25 Project 1 report</td>
<td></td>
</tr>
<tr>
<td></td>
<td>W: 9/23</td>
<td>Group 2 – Progress report presentation (project 1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>M: 9/28</td>
<td>Complex surveys</td>
<td>10/2 Final project consultation</td>
<td>Vittinghoff, Ch12</td>
</tr>
<tr>
<td></td>
<td>W: 9/30</td>
<td>Complex surveys group work</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>M: 10/5</td>
<td>Complex surveys</td>
<td>10/8 Project 2 descriptive stats</td>
<td></td>
</tr>
<tr>
<td></td>
<td>W: 10/7</td>
<td>Group 3 – Analysis plan presentation (project 2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>M: 10/12</td>
<td>TBD</td>
<td>10/15 Proj 2 primary analysis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>W: 10/14</td>
<td>Group 4 – Analysis plan presentation (project 2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>M: 10/19</td>
<td>Longitudinal analysis</td>
<td>10/23 Project 2 report</td>
<td></td>
</tr>
<tr>
<td></td>
<td>W: 10/21</td>
<td>SAP peer review</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>M: 10/26</td>
<td>Longitudinal group work</td>
<td>10/30 Final project descriptive stats</td>
<td>Bell 2015</td>
</tr>
<tr>
<td></td>
<td>W: 10/28</td>
<td>Missing data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>M: 11/2</td>
<td>Missing data</td>
<td>11/5 Project 3 descriptive stats</td>
<td></td>
</tr>
<tr>
<td></td>
<td>W: 11/4</td>
<td>Multiple imputation group work</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>M: 11/9</td>
<td>Group 5 – Analysis plan presentation (project 3)</td>
<td>11/12 Project 3 analysis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>W: 11/11</td>
<td>NO CLASS, Veterans’ Day</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>M: 11/16</td>
<td>Group 6 – Progress report presentation (project 3)</td>
<td>11/20 Project 3 report</td>
<td>JAMA statistical guidelines*</td>
</tr>
<tr>
<td></td>
<td>W: 11/18</td>
<td>Review</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>M: 11/23</td>
<td>Review</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>W: 11/25</td>
<td>Day before Thanksgiving, no class</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>M: 11/30</td>
<td>Final presentations</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>W: 12/2</td>
<td>Final presentations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>M: 12/7</td>
<td>Final presentations</td>
<td>12/15 Final report</td>
<td></td>
</tr>
<tr>
<td></td>
<td>W: 12/9</td>
<td>Final presentations</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*https://jamanetwork.com/journals/jama/pages/instructions-for-authors#SecStatisticalMethodsandDataPresentation
*https://www.nejm.org/author-center/new-manuscripts

Communications: You are responsible for reading emails sent to your UA account from your instructor and the announcements that are placed on the course web site. Information about readings, news events, your grades, assignments and other course related topics will be communicated to you with these electronic methods. The official policy can be found at: https://www.registrar.arizona.edu/personal-information/official-student-email-policy-use-email-official-correspondence-students

UA Smoking and Tobacco Policy:
The purpose of this Policy is to establish the University of Arizona’s (University) commitment to protect the health of University faculty, staff, students, and visitors on campuses and in its vehicles. The official policy can be found at: http://policy.arizona.edu/ethics-and-conduct/smoking-and-tobacco-policy

University Course Policies: (please see the following URL):
https://academicaffairs.arizona.edu/syllabus-policies

Syllabus Changes: Information contained in the course syllabus, other than the grade and absence policies, may be subject to change with reasonable advance notice, as deemed appropriate.

Plagiarism: What counts as plagiarism?
- Copying and pasting information from a web site or another source, and then revising it so that it sounds like your original idea.
- Doing an assignment/essay/take home test with a friend and then handing in separate assignments that contain the same ideas, language, phrases, etc.
- Quoting a passage without quotation marks or citations, so that it looks like your own.
- Paraphrasing a passage without citing it, so that it looks like your own.
- Hiring another person to do your work for you, or purchasing a paper through any of the on- or off-line sources.
BIOS 576C, Report guidance

As a rule, you may follow JAMA submission guidelines, except as noted below. Do not include abstract or key points. https://jamanetwork.com/journals/jama/pages/instructions-for-authors#SecManuscriptStyle

Please turn in a single word document, single-spaced, with tables and figures embedded into the manuscript (rather than at the end of the paper). The report should be no longer than 2500 words, and have no more than 4 tables and 2 figures. You must use referencing software (e.g., Endnote, Mendeley). Include relevant code in an appendix, which is not counted in the word total.

Reports should include the following:

1. Introduction/background (400 words or less)
   This section will be shorter than in a regular research manuscript, as I want you to be more focused on methods and results. However, you should (briefly) include the following elements: Why does this matter? What else has been done on this topic? What gap will your report fill? Specific aims for the paper.
2. Methods. See the SAP guidelines below
3. Results
   a. Descriptive statistics on study sample
   b. Primary analysis results
   c. Secondary results (if any)
   d. Sensitivity results (for the primary analysis)
   e. Assessment of assumptions
4. Discussion
   a. Brief overview of main results
   b. Interpretation, explanations and how do your results compare to others from the literature?
   c. Strengths and limitations
   d. Take home message/implications for public health
5. Appendix
   a. Relevant code

REMEMBER THAT MORE IS NOT BETTER AND IS OFTEN WORSE!

Grading rubrics for projects

Introduction (2.5%)
Study design, setting, participants (5%)
Measures (10%)
Statistical methods (20%)
Sensitivity analysis (10%)
Results, text (10%)
Results, tables and figures (10%)
Discussion (2.5%)
Writing: organization, clarity, grammar (30%)
First, become familiar with your data. What variables do you have in your dataset? How and when are they measured? How much missing data do you have? Perform some descriptive analyses. What do the distributions look like?

For efficiency, you should write the statistical methods section that will go into your report. Use reporting guidelines, such as the STROBE statement, to help you. Find a partner to give you peer review. **Remember that if someone else had your data, they should be able to replicate your results.**

**Your SAP should include**
- Background and objective(s). May include hypotheses.
  - Study design, setting and participants
  - Include eligibility criteria, sampling methods
- Measures
  - Primary outcome
  - Exposure of interest/hypothesis variable
  - Secondary outcome(s)
  - Any covariates or confounders you will include, and why
  - If they are a scale, what is the possible range?
  - If you have a categorical variable, what are the levels of that variable?
  - Are you going to categorize continuous variables? If so, how, and why?
  - If you are transforming any of the variables, state how and why.
  - What are the data sources? (E.g., questionnaire, lab, hospital records, etc).
  - What is the assessment schedule? (E.g., baseline, 3 and 6 months).
- Statistical methods
  Match your methods to your objectives and hypotheses. Be explicit about all analyses planned.
  - Descriptive statistics (Table 1)
    - Will you be stratifying by anything (like your exposure or outcome?)
  - Primary analysis: What models will be used and which objective are you addressing with each model?
  - Effect modification: if you plan to test this, how and why (your “why” should be touched on in your background/objectives).
  - How many regression models are you fitting?
  - How will missing data be handled?
  - Secondary analysis
  - Sensitivity analyses to the primary analysis. How robust are your results to assumptions made in the primary analysis?
- Shell tables
  Your tables and statistical methods section should be congruent. Be thorough in your description/table legend, and include details, if needed, in table footnotes. Sample sizes should be included.
  - Table 1 should describe your sample and may be stratified on exposure or outcome
    - Primarily baseline variables
    - Most variables used in the analyses should be described here.
    - Other variables may also be included, in order to describe your study sample
  - Other tables should include the results from your analyses, including the measure(s) of effect and 95% confidence intervals.