**BSTA 651**  
**Biostatistics III: Linear and Generalized Linear models**

**Time:** Tu, Th: 1:30 PM- 2:50 PM  
**Location:** Blockley Hall

**Instructors:** Justine Shults, Ph.D

**Contact information:**  
Justine Shults, Ph.D,  
**Office:** Blockley 626 (and Roberts 4471)  
**Phone:** 215-518-8628  
**Email:** shultsj@email.chop.edu  
**Office Hours:** Tuesdays and Thursdays/ After Class or by appointment

**Teaching Assistant:**  
Victoria Arthur  
**Office:** Blockley 108/109  
**Email:** vlynn@pennmedicine.upenn.edu  
**Office Hours:** 3:00 Thursday or by appointment

**Text-book:** *Linear Models in Statistics, 2nd Edition* by Rencher and Schaalje  

**Additional references:**  

**Note:** You do not need to purchase the references. A good text on matrix algebra (e.g. by Harville, ISBN 978-0-387-22677-4 ) will also be useful to have for future reference.

**Prerequisites:** BSTA 620 and BSTA 630. Both BSTA 621 and BSTA 631 may be taken concurrently, with permission of the instructor.

This course provides an introduction to the theory (primarily) and application of linear and nonlinear models. Topics covered in this course include: (1) multiple linear regression models; (2) analysis of variance models; and (3) generalized linear models.

Weekly (approximately) homework assignments will be given. Please work on these assignments independently, unless indicated otherwise.

In addition, there will be 2 examinations. One midterm exam and one final exam. Your final grades will be based on your work in both the homework problems and the two exams according to the following distribution:
(1) HW (40%); (2) Part I exam (25%); (3) Final (30%); (4) Class-Participation (5%).

**Note regarding software:** The emphasis of this course will be on the theory of linear models, but some applications will also be presented. Examples for Part I of the course will primarily be presented in Stata. If you are asked to do programming for an assignment and would like to use a different software package, please consult the instructor for permission. Please be sure that your results are clearly summarized, and any supporting code is included.

**Spring Break:**

No classes will be held on March 5 or March 7.

**Reading Days:**

Reading days will be held from May 2-3.

**Final Exam:**

Please see the final exam policy that is posted on the following UPenn web-site:

[https://www.registrar.upenn.edu/pdf_main/provost-rules.pdf](https://www.registrar.upenn.edu/pdf_main/provost-rules.pdf)

Please also see the following exam schedule for 2019:

[https://www.registrar.upenn.edu/pdf_main/19A_Final_Exam_Tentative_01102019.pdf](https://www.registrar.upenn.edu/pdf_main/19A_Final_Exam_Tentative_01102019.pdf)

Our scheduled final exam will be held on Monday, May 13 from 9 a.m. to 11 a.m. The final exam will focus primarily on the second half of the course.
### Classes and Topics for Part A (Linear Models):

(please note that this will be updated to include GLM. Please be sure to read the text.)

<table>
<thead>
<tr>
<th>Activity</th>
<th>Date</th>
<th>Reading</th>
<th>Topic</th>
<th>Homework: Assigned</th>
<th>Due</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lecture 1</td>
<td>Th: 1/17/2019</td>
<td>R: Ch 1-2</td>
<td>Introduction and Review</td>
<td>HW1</td>
<td></td>
</tr>
<tr>
<td>Lecture 2</td>
<td>Tu: 1/22/2019</td>
<td>R: Ch 6-7</td>
<td>Linear Regression: simple and multivariable</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lecture 3</td>
<td>Th: 1/24/2019</td>
<td>R: Ch 2,4 and 7.6</td>
<td>Multivariate Normal Distr. &amp; Maximum likelihood Regression</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lecture 4</td>
<td>Tu: 1/29/2019</td>
<td>R: Ch 5</td>
<td>Distributions of Quadratic Forms</td>
<td>HW2</td>
<td>HW 1</td>
</tr>
<tr>
<td>Lecture 5</td>
<td>Th: 1/31/2019</td>
<td>R: Ch 8</td>
<td>Regression: Tests of hypothesis and confidence intervals</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lecture 6</td>
<td>Tu: 2/5/2019</td>
<td>R: Ch 8</td>
<td>General tests of hypothesis</td>
<td>HW 3</td>
<td>HW 2</td>
</tr>
<tr>
<td>Lecture 7</td>
<td>Th: 2/7/2019</td>
<td></td>
<td>Catch up and/or examples</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lecture 8</td>
<td>Tu: 2/12/2019</td>
<td>R: Ch 12</td>
<td>Non-full rank models and estimable functions</td>
<td>HW 4</td>
<td>HW 3</td>
</tr>
<tr>
<td>Lecture 9</td>
<td>Th: 2/14/2019</td>
<td>R: Ch 13</td>
<td>One way analysis of variance model- balanced case</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lecture 10</td>
<td>Tu: 2/19/2019</td>
<td>R: Ch 14</td>
<td>Two way analysis of variance model- balanced case</td>
<td>HW 5</td>
<td>HW 4</td>
</tr>
<tr>
<td>Lecture 11</td>
<td>Th: 2/21/2016</td>
<td>R: Ch 15</td>
<td>Analysis of variance- cell means model for unbalanced data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lecture 12</td>
<td>Tu: 2/26/2019</td>
<td>R: Ch 16</td>
<td>Analysis of Covariance</td>
<td>HW 6</td>
<td>HW 5</td>
</tr>
<tr>
<td>Lecture 13</td>
<td>Th: 2/28/2019</td>
<td>R: Ch 7</td>
<td>Generalized Least Squares</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lecture 14</td>
<td>Tu: 3/12/2019</td>
<td></td>
<td>Misc. &amp; Review</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Midterm</td>
<td>Th: 3/14/2019</td>
<td></td>
<td>Midterm Exam (Closed Book)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Spring Break:** 3/5 – 3/7