**STT 461 Computations in Probability and Statistics, Spring 2014**

Professor Raoul LePage, C428 Wells Hall, lepage@stt.msu.edu

Office hours Tu, Th 11:30-12:30 C428 WH or by appt.

Class meets MW A318 WH, 3:00-4:20 *revised as needed (red)

Daily attendance.

---

### Course Descriptions: Subject/Course Search Results

**Course:** STT 461 Computations in Probability and Statistics  
**Semester:** Spring of every year  
**Credits:** Total Credits: 3  
**Recommended Background:** CSE 131 and STT 441 and (MTH 309 or MTH 314)  
**Description:** Computer algorithms for evaluation, simulation and visualization. Sampling and prescribed distributions. Robustness and error analysis of procedures used by statistical packages. Graphics for data display, computation of probabilities and percentiles.

**Effective Dates:** FALL 2008 - Open  
[View all versions of this course](#)

---

### Classes Begin

<table>
<thead>
<tr>
<th>Class</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Martin Luther King Day</td>
<td>Monday, 1/6</td>
</tr>
<tr>
<td>University open, classes</td>
<td>Monday, 1/20</td>
</tr>
<tr>
<td>cancelled</td>
<td></td>
</tr>
<tr>
<td>Holiday-University Closed</td>
<td></td>
</tr>
<tr>
<td>End of Tuition Refund</td>
<td></td>
</tr>
<tr>
<td>Dates can be found by clicking</td>
<td></td>
</tr>
<tr>
<td>Holiday-University Closed</td>
<td></td>
</tr>
<tr>
<td>Middle of Semester</td>
<td>Wednesday, 2/26</td>
</tr>
<tr>
<td>Spring Break</td>
<td>Monday, 3/3-Friday, 3/7</td>
</tr>
<tr>
<td>Holiday-University Closed</td>
<td></td>
</tr>
<tr>
<td>Classes End</td>
<td>Friday, 4/25</td>
</tr>
<tr>
<td>Final Exams</td>
<td>Monday, 4/28-Friday, 5/2</td>
</tr>
<tr>
<td>Commencements</td>
<td>Friday, 5/2-Sunday, 5/2</td>
</tr>
</tbody>
</table>

---

**No final exam. You do not have to attend.**
Note for week 1-6-14

Files 461 1-6-14.pdf and 461 1-8-14.pdf have been posted to

www.stt.msu.edu
Academics
Class Pages List
STT 461
Class Notes

or simply


Note: I cannot post to the page until it is put up by the STT department.

Your first HW assignment, due in class Wednesday, 1-8-14 is

Prepare by getting R up and running on your computer or tablet (see below).
Experiment with R, carrying out a rich selection of the activities (e.g. see videos below).
Turn in four pages of printout and your commentary on what you have attempted.
O. Obtain R freely from
   http://cran.us.r-project.org
OO. Some like RStudio which employs a different user interface available free from
   http://www.rstudio.com
OOO. New, using a tablet, even an iPad, to access R from the Cloud.
   http://www.easyr.it/10-reasons-to-use-easyr/

I am looking into EasyR myself and you might want to. It is not from cran.
Let me know if you have experience/information about this type of product, e.g.
   http://www.r-bloggers.com/rstudio-in-the-cloud-for-dummies/

1. Google R tutorial(s) R 1.1 through R 4.6. Launching R 1.1 below will start through all.
   http://www.youtube.com/watch?v=iffR3fWv4xw&list=PLOU2XLYxmslK9qQfztXeybpHvu-TrqAP

2. Use Google search to get answers to your R questions! Documentation is uneven.
3. Here is a link to a fast paced introduction on loading data into R.
   http://www.youtube.com/watch?v=VLtazaiYo-o
4. Here is a follow up to (3) on dealing with ‘outliers’ and missing data in R files.
   http://www.youtube.com/watch?v=ckxEZDNliok

There are some uncorrected mis-speaks in this and (3), for example R1.3 errs in
defining ‘vector’ at around 6 seconds. Report serious errors to me.

Also before the first class 1-6-14, A WH, 3:00-4:20 look through Chapters 1
(Introduction to R) and 2 (Probability and Statistics Review) of the required textbook Statistical
Computing with R by Maria Rizzo.

It is not assumed that you know all of what you see in Chapters 1-2. We will begin with
Chapter 3 (Methods for Generating Random Variables), citing and reviewing material from
Chapters 1 and 2 on an as-needed basis.

Since R is new to many of you week 1 will place greater emphasis on the basics of installing
R and using R to accomplish basic tasks. It is ok to use your computer in class. Do get to
know your classmates right away as relationships form, ask questions and share advice. You
are part of a culture of educated people.
Except during exams, consult freely on matters of methodology but of course do not let
that morph-over to copying the work of others.

The standard is that you pull your own weight on HW just as you do not copy on exams.
Gate keeping responsibilities reside with you allowing others to ride on your work as well as
permitting yourself to ride. Each is prohibited. On the other hand discuss R with anyone who
can help you gain skill. This part of 461 material needs to be talked over with others, like
learning a language. After all, R is a language.

Have a look at on-line resources, video lectures and the like. Poke around the R website
above.

In the job market the favored programming languages or other computer software skills vary
with employers/situations. Give serious thought to establishing credentials with several pack-
ages such as SAS (you can self teach and take exams to certify in SAS), and so forth. In job
interviews it really helps to have such skills in your resume.

You should all have R skills simply because everyone is picking it up. There are a lot of
serious statisticians, scientists, business people, engineers of all stripes behind R. It is rela-
tively sophisticated and free! This will probably win over much of the commercial and scientific
space since some of the popular competing statistical computing packages can run to a thou-
sand dollars per year.

This course aims to enlarge your framework for statistical thinking, whether Bayesian,
frequentist or other, adding tools for exploratory computing as well as carrying out sophisti-
cated analyses. You will learn to appreciate the advantages of routinely passing some of your
ideas through computer implementations, verifications, tinkerings and presentations, much as
you might consult with and borrow from an increasingly insightful adviser, especially one
capable of doing a lot of serious analysis for you and to whom you will have uninterruptible
career-long access. Do keep in mind that R is itself morphing and will ultimately be absorbed in
something of greater power.

By undertaking this interchange on a continuing basis you will incrementally enlarge and
explore your own core understandings of probability and statistics. As by-product you will
develop means to formulate and carry out your own independent investigations into the actual
performance of incompletely understood or mistrusted aspects of research and practice, even
if they be complex.

Your course grade will be determined

50% exams grade average, 50% homework and attendance grade average.

Homework: will be due about every other week. Any R code must be fully annotated and shown to work to earn points.

There will be three, in-class, without-notes exams. Tentatively:

Exam 1 W 1-29-14, Exam 2 W 2-26-14, Exam 3 W 4-16-14.

There will be no final exam but I will be present in the classroom at the scheduled final exam period Wed, April 30, 3-5 to answer your questions on any aspects of 461 or statistics. You need not attend.

Topics are selected from the textbook and outside resources. Many major topics will be touched upon including

- Randomization, simulation, Monte-Carlo, Importance Sampling,
- MCMC, EM, likelihood based inference, Bayesian Inference,
- probabilities believed to apply vs imposed vs without foundation,
- Empirical Bayes, Bootstrap, data compression and sparsity,
- on-line learning, boosting, linearization, support vector machines.

When preparing for exams come to each exam "owning" some of the topics and expecting to see the basic skills and understandings tested. You should make every effort to identify the larger points and important skills. I sometimes award > 4.0 for scores far above the class as a whole. Relatively few students achieve that. If you want any given course grade do not lay back on HW unless you are truly pluperfect on exams.