

Integrating Computing into the Statistics Curricula Instructors Workshop

Deb Nolan, UC Berkeley
Duncan Temple Lang, UC Davis
Mark Hansen, UCLA
Roger Peng, Johns Hopkins University

Berkeley, July 08

Context

Second workshop – focus on instructors who are developing & teaching computing courses

First workshop (May 07) – discussion by stat. computing researchers of topics and how to teach statistical computing

Third workshop – 2009 –
same as this with different set of instructors?
or an intensive workshop to teach computing to “New Researchers”?

Larger Context

NSF DUE funding
Division of Undergraduate Education

Goals

Develop model syllabi and curricula for “computing”
within statistics programs
primary focus undergraduate

Develop teaching resources

lecture notes
book
homeworks, labs, code/software
community resources (wiki, mailing list, ...)

Motivation

Science and statistics are changing a great deal,
becoming multi- & inter-disciplinary.

Generation, collection and dissemination of data very
different, and very large.

Web portals

databases

data technologies (XML, SOAP, ...)

Data analysis and “modeling” only one part of
statisticians work.

Computing changes what statistical methods we can use and what we do each day.

So our curricula should bend to embrace some of these changes.

Teach modern statistical methods within a computing class to ensure students see “new” material.

Opportunity to create a new breed of statistician that is involved in more aspects of data acquisition, management, analysis, dissemination, etc.

And can appreciate trends and forces in computing & technology

Why are data technologies fun?

Students get to do things, and be creative.

A sense of achievement of analyzing data that are topical in an informal manner.

And with modern statistical methods that we teach heuristically.

Expose them to research and give them a sense of statistics in action.

Learn practical tools that they can use in real settings to connect statistics to their other interests & work.

Mathematics is an essential tool with which we study statistical methods.

Computing is an essential tool with which we study and “do” data analysis.

Yet, very few classes in the kind of computing we do in statistics, and very few computing prerequisites.

Our students are missing a vital dimension to be effective in industry, policy, inter-disciplinary research, statistical research, etc.

Scientific Computing is becoming an essential tool for statisticians.

Computing – broad categories

Computational Statistics – algorithms for stat. methods.

Programming

Data Technologies – databases, XML, Web, ...

Advanced Computing – large data, low-level languages, parallel computing, object-oriented programming, ...

Visualization – composition, software, dynamic & interactive, GIS, ...

What are the sub-topics,
how do they overlap,
how do they depend on each other,
which ones are appropriate for different students,
within an existing class or a separate computing class,
quarter or semester or multi-course sequence,
or intensive boot-camp for graduates?

How do we teach the somewhat foreign subject matter
and different medium?

Grading & evaluation?

Goals for the week

2 primary goals

discuss how to design, introduce & teach computing
topics within the statistics curriculum,

teach some of the topics that may not be familiar
(and teach some of the more abstract and deeper
aspects of familiar topics)

Challenging to get the right level across this group, so
please speak up and help establish the right level of
detail, abstraction, discussion, ...

Schedule

Start 8.30

8.30 - 10.00: topic

10.00 - 10.30: coffee

10.30 - 11.45: topic

1.00 - 1.45: topic

1.45 - 3.15: lab

3.45 - 5.00: discussion

No lab on the last day, unless.....

Topics

Mon	Tue	Wed	Thu
R basics	Programming	Regular Expressions	Computational Statistics
Graphics	Databases	Command line & Shell	XML
Input/Output †	Comparison of Languages	open discussion	Advanced Computing
Lab	Lab	Lab	Wrap up
Designing Curricula	Elementary Computing & Administration	Designing homeworks	