Intro stat should not be like drinking water through a fire hose

Kirk Steinhorst
Professor of Statistics
University of Idaho

What is intro stat?
Introductory statistics is …

The syllabus 30 years ago…
- Descriptive statistics
- Probability
- Sampling distributions
- Hypothesis testing
- Confidence Intervals
- One-way ANOVA
- Simple linear regression
- Chi-square tests

The syllabus today…
- Descriptive statistics
- Study design
- Probability
- Sampling distributions
- Inference in sample surveys—point and interval estimation
- Inference in experiments—and hypothesis testing
- One-way ANOVA
- Simple linear regression
- Chi-square tests (as time allows)
The other addition to the modern intro course is statistical computing.

Descriptive statistics in the old days
- Measures of location—mean, median, geometric mean, harmonic mean
- Measures of scale—range, variance, standard deviation
- Graphs—relative frequency diagrams and histograms
A great deal of time was spent teaching students how to do the computations.
Descriptive statistics today

- **Graphs**
  - Quantitative variables—stem-and-leaf plot, dot plot
  - Continuous—histogram, box plot
  - Discrete—bar plot
- **Categorical variables**—pie chart, bar plot

- **Statistics**
  - Measures of location—mean vs. median and why
  - Measures of scale—range, interquartile range, standard deviation (and variance)
  - Measures of position—percentiles, deciles, quartiles, median

Note. For categorical variables, we use proportions as the descriptive statistics.

**Note:**
Resist the temptation to cover two-way frequency tables and scatterplots as part of descriptive statistics. This material is better saved for the chi-square and regression sections later.

Drawing histograms properly

Many modern introductory texts and computer programs confuse frequency graphs, relative frequency graphs, and histograms.

See histogram video

For example…done poorly

And drawn properly…

How does a bar chart differ?
Study design

- Surveys
- Experiments
- Other—observational studies, case studies, random processes, retrospective studies, etc.

Probability in the old days

- Axioms of probability, Venn diagrams, balls and urns
- General addition rule, complements, conditional probability, independence
- Permutations/combinations—counting rules
- Bayes’ rule (optional)
- Binomial and normal distributions

Note. Many examples involved dice, cards, coins and the like.

The difficulty

- Students found the material hard and boring.
- Students did not see the connection between probability and statistics.
- Students were distracted by the counting rules.

Probability today

The coverage of probability should be closely tied to the subsequent coverage of statistics. The key is to present the basic rules of probability by using probability to describe populations and random sampling from populations.

So what do we do?

We cover the same topics but we use realistic statistical examples.
We introduce probability density and mass functions in general.

See A New Approach to Learning Probability in the First Statistics Course

Journal of Statistics Education, V9N3: Keeler

For example

<table>
<thead>
<tr>
<th>College</th>
<th>Ag-Far</th>
<th>Arts-Sci</th>
<th>Bus-Bmgt</th>
<th>Edu</th>
<th>Env</th>
<th>Law</th>
<th>Med</th>
<th>Under</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>700</td>
<td>1200</td>
<td>1500</td>
<td>1500</td>
<td>800</td>
<td>1500</td>
<td>500</td>
<td>1100</td>
<td>5500</td>
</tr>
<tr>
<td>Male</td>
<td>600</td>
<td>2000</td>
<td>1000</td>
<td>1200</td>
<td>200</td>
<td>1000</td>
<td>300</td>
<td>1200</td>
<td>5000</td>
</tr>
<tr>
<td>Totals</td>
<td>1300</td>
<td>3200</td>
<td>2500</td>
<td>2700</td>
<td>1000</td>
<td>2500</td>
<td>800</td>
<td>2300</td>
<td>10500</td>
</tr>
</tbody>
</table>

If I select a student at random, answer the following:

a) Find P (Female or Male).
b) Find P (not Ag-Far).
c) Find P(Female|Bus-Econ).
...and another example

You are the manager of the complaint department for a large mail order company. Your data and experience indicate that the time it takes to handle a single call has pdf.

Sampling distributions

Most modern texts discuss the asymptotic distribution of sample proportions.

They also discuss the normality of sample means (directly or via the central limit theorem).

I cover the BIGGER ideas...

Students understand...

Of course, we cover

Survey Inference

Many texts assume that data come from infinite populations under i.i.d. random sampling. Today’s student will see analyses of more surveys in their lifetime than anything else.

For this reason, I introduce point and interval estimation with data from sample surveys.
Point estimation

Point estimation follows quite naturally from the previous discussion of descriptive statistics…

\[ \hat{\mu} = \bar{X} \]
\[ \hat{\sigma} = \text{the sample median} \]
\[ \hat{\theta} = s \]

etc.

In addition...

We note that the sample histogram is the estimate of the pdf of a continuous random variable and the sample bar graph is the estimate of the pmf of a discrete or categorical random variable.

\[ \hat{f}(\cdot) \text{ is the histogram} \]
\[ \hat{p}(\cdot) \text{ is the bar graph} \]

That is why it is important to get these graphs right when you do descriptive statistics.

CIs

One can construct confidence intervals for lots of parameters or differences of parameters.

Restrict the discussion to CIs for single proportions and means. Keep it simple.

I also introduce the fpc—they can handle it 😊

Hypothesis testing

Hypothesis testing has more possibilities for “firehose” teaching behavior.

- Mean or binomial proportion
- z or t
- 1-sample, paired, unpaired
- 1-sided, 2-sided
- Equal or unequal variances

There are at least 22 distinct cases.

Show restraint

- Treat one-sample and paired cases as one.
- Do only t tests for hypotheses involving means.
- In the two-sample, unpaired case, choose either the equal or unequal variance case—I choose equal because it leads to a seamless transition to one-way ANOVA.

One scenario...

- The Z test for a binomial proportion
- The paired t test
- The equal variance unpaired t test

Do both 1- and 2-sided alternatives in each case.

It is a matter of personal choice whether you test means first or proportions.
One-way ANOVA

• The idea
• Side-by-side box plots
• The AOV table
• The F and p-values and their interpretation.

Simple linear regression

• Scatterplots
• Review the equation of a line
• Point and interval estimation (use “The Figure”)
• Test of zero slope
• Calculation and interpretation of correlation

Chi-square tests (as time allows)

• Find a good example of a one-way table of counts. Illustrate the computation of expected values.
• Compute the Pearson chi-square formula for this example.
• If time permits do a two-way table.

Schedule

Intro and descriptive statistics (or study design) ………. 1 week
Study design (or descriptive statistics) ………………. 1 week
Probability using the population approach …………. 3 weeks
Sampling distributions in general …………………… 1 week
Survey inference ……………………………… 2 weeks
Experiments and hypothesis testing ………………. 2 weeks
One-way ANOVA …………………………….. 1 week
Simple linear regression …………………….. 2 weeks
Chi-square tests (as time allows) ………………. 1 week
Tests ……………………………………………. 1 week
Total ………………………………………… 15 weeks
In the beginning...

- Lecture
- Homework
- Tests

We had to teach a lot of mechanics ... and thus we tested a lot of mechanics.

Now...

- Minilectures
- In-class activities
- Homework
- Tests

Assessment has changed because of changes in philosophy and technology.

Changed philosophy

- Authentic assessment
- Departure from traditional lecturing
- Change from teaching mechanics to analyzing real data sets to teaching statistical reasoning (as well)

New technology

Ubiquitous computing has removed the burden of arithmetic ... so after hand calculating a small example or two...

At the present time, most students have access to a computer where they live or in campus labs.

Some students have laptops and some campuses provide wireless connectivity.

In the not so distant future students will carry computing devices.

Assessment

- In-class activities keep the students focused and they help me keep track of what the students are learning and where more effort is needed.
- Homework is still important.
- I use 2 or 3 specially organized hourly tests and a comprehensive final.
In-class activities

I use in-class activities as a learning tool and for assessment.

Students work in small groups after the minilecture.

The activities replace the rest of the traditional lecture providing opportunities for experiential learning.

These activities are usually done in groups. This encourages “students learning from students” and makes the grading easier.

I do not manage groups, but perhaps I should.

Some students are not prepared to do an activity on material just covered in the minilecture.

For example…

Activity 18

The groups are asked to verify that the standard error is 0.126 by giving the formula and working through the numbers.

They have about 20 minutes to complete this activity.

Conceptual activity

Concrete activity

I usually do the “random rectangles” activity from the Scheaffer et al. book. I give them the rectangle drawing and tell them that the figures represent the numbers of potential customers around the campus for a new bike repair service they and a friend are about to startup.

I first ask them to make an “ocular” estimate of the total number of potential customers.
I then ask them to pick 5 “representative” living units and compute the average and multiply it by 100 thus producing an estimate of the total.

I next have them pick 5 two digit numbers from a random number table and use them to select 5 living groups at random. They compute the average of these and multiply by 100.

We repeat the random selection with 10 random numbers.

At each stage, I build a dot plot of the results on the board. In almost every case, the estimates get better as we go. Students are convinced that random sampling works.

Students turn the answers in for activity credit and I enter all the data into a computer package and produce graphical and tabular summaries which I present at the next class meeting.

Homework?
Just assign homework, but don't grade it?
Assign and grade homework
CyberStats or iLrn or …
Semi automated computer-based homework systems provide a feasible way to assign and grade homework in large classes. Should the students have a limited number of tries? Should students be able to see help notes after X tries? How about algorithmic homework systems?

**Testing**

Testing presents one of the greatest challenges in the current course climate. How can it be authentic? What about technology? Can students use any calculator or computing device? How do you test concepts?

**My solution**

Over the last 10 years I have developed a solution that addresses most of these issues—I call it the takehome/in-class test system.

Students are given a takehome exam on Monday (say) and on Friday they bring the takehome exam to class and use it to answer the in-class exam. They do not turn in the takehome exam.

**Benefits**

- Students can use computer resources when preparing the takehome answers.
- Students can work together on the takehome.
- The takehome removes some of the fear associated with statistics exams.
- Students are individually responsible for knowing the material.
Notice that...

- Question 1 on the in-class is a confidence builder. If they did the takehome and understood it...
- There are a mixture of mechanical and conceptual questions.
- One way to find out if they just copied the takehome answers from a friend is to give a new version of the data and have them rework a problem (see problem 4).

Most students liked this method of testing...

- **Overall, how would you rate the performance of the instructor in teaching this course?** 4
- **Comment:** You did a good job of integrating technology into the teaching of this course. The take home portion of the test is probably the most helpful part of this course. I also like the h/w assignments.
- **Overall, how would you rate the quality of this course?**
- **Comment:** Some concepts in this course seem like ones that I'll never again use. What is the usefulness of learning things that we'll never again use?

But not all students liked this testing system...

**Comment:** He often would use the online portion of the course in the classroom, but on exams we went permitted to use it, so it made it an obstacle to know how do problems by hand on exams with out the aid of the computer.

- **Overall, how would you rate the quality of this course?** 4
- **Comment:** I feel that it should be all online or all off line. It is difficult to learn to compute a problem online then on exams not to use a computer. The link between the two is week.

Last but not least

The university does an end of course evaluation (poorly). While this is of some use, I do my own.

I ask only two questions...

- **Comment on one thing that we should keep doing.**
- **Comment on one thing that we can do better.**
In summary

- In-class activities (100 points using best 70%)
- Homework (100 points)
- “Progressive supper” tests (3x100 + 150 points)
  Total is 650 points.

- Course evaluations

Do better at...
- Working problems not in detail by hand—see CyberTags
- Read...
- Exploit technical difficulties
- Explaining equations and symbols
- Helping us understand material
- Figuring out how to use CyberTags and/or Lab...
- Starting ...
- Too much new material, too fast
- Handing things
- test review, all review sessions
- Talk about homework after it is due
- Test
- Last two or three homework to construct “gunning”
- Make homework exactly like the book
- Having back assignments
- Class discussion
- T.A.
- More review and repetition
- More S.A.C. time