7 Wonders of the Statistical World

marty@triolastats.com
1. Changing Nature of Statistics

“What can you change in a statistics book?”

“The statistics course has not changed in many years.”
### Printed Tables

**Table A-3  $t$ Distribution: Critical $t$ Values**

<table>
<thead>
<tr>
<th>Degrees of Freedom</th>
<th>Area in One Tail</th>
<th>Area in Two Tails</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.005</td>
<td>0.01</td>
</tr>
<tr>
<td>1</td>
<td>63.657</td>
<td>31.821</td>
</tr>
<tr>
<td>2</td>
<td>9.925</td>
<td>6.965</td>
</tr>
<tr>
<td>3</td>
<td>5.841</td>
<td>4.541</td>
</tr>
<tr>
<td>4</td>
<td>4.604</td>
<td>3.747</td>
</tr>
<tr>
<td>5</td>
<td>4.032</td>
<td>3.365</td>
</tr>
<tr>
<td>6</td>
<td>3.707</td>
<td>3.143</td>
</tr>
</tbody>
</table>
Use z if $n > 30$

- Journals/professionals: Almost never
- Technology now makes $t$ easy
$H_0$ with $\geq$ or $\leq$

- Journals/professionals: almost never
- Now: $H_0$ with = only
Confidence Interval and Hypothesis Test for $\mu$ assuming known $\sigma$

- Unrealistic!
- Journals/professionals: *never*
- Technology: *t* easy
Confidence Interval and Hypothesis Test for two independent means with

1. Known standard deviations

or 2. Equal standard deviations*
Two Independent Means

2 known standard dev’s: *unrealistic*

* Assuming equal standard dev’s: *pooling*
  - OK in placebo/treatment groups
  - Yields higher $df$, so greater power and/or narrower confidence interval
Critical value method of testing hypotheses...

- Almost obsolete in practice
- **But** good to teach hypothesis testing with:
  - $P$-values
  - Confidence Intervals
  - Critical value method
  - Simulations
More prob/stat in K-12:

- Mean
- Median
- Dotplot
- Scatterplot
- Probability
- Addition Rule
- Conditional Probability
2. Sees same symptoms that led to blood clots.
3. Records search using basic statistics:
   — blood clot danger
New: Big Data

1. Prof. mined FDA records for adverse reactions.
2. Found problem pairings: Paxil/pravastatin led to high blood sugar.
New: Big Data

- Google: live traffic maps by analyzing GPS data from smart phones.
- Forecasts of flu epidemics made by analyzing Internet searches of flu symptoms.
Relatively New: *Simulations*

- Probability (birthdays)
- Hypothesis Testing (next)
- Statistical Theory:
  
  2 Means: Don’t assume $\sigma_1 = \sigma_2$
Simulation: Hypothesis Test

“The body temperature of an average, healthy human child or adult is 98.6 degrees Fahrenheit (37 degrees Celsius).”

–USA Today Nov. 9, 2014
(Ebola article)
Simulation: Hypothesis Tests

Claim: \( \mu = 98.6 \) degrees.

Sample Data:

\[
\begin{align*}
  n &= 106 \\
  \text{Mean} &= 98.20 \\
  s &= 0.62
\end{align*}
\]
Simulation

Assume $\mu = 98.6$.

Each student randomly generates $n = 106$ values using:

- Normal distribution
- Mean 98.6
- Standard deviation 0.62
<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>98.7086</td>
<td>98.5790</td>
<td>98.6065</td>
<td>98.5134</td>
<td></td>
</tr>
<tr>
<td>98.6417</td>
<td>98.6801</td>
<td>98.6459</td>
<td>98.5490</td>
<td></td>
</tr>
<tr>
<td>98.5726</td>
<td>98.5512</td>
<td>98.5219</td>
<td>98.6157</td>
<td></td>
</tr>
<tr>
<td>98.7668</td>
<td>98.5900</td>
<td>98.5051</td>
<td>98.5980</td>
<td></td>
</tr>
</tbody>
</table>
New

Bootstrap Resampling
2. Technology
Technology has fostered many changes:

• *P*-values
• Tables, known $\sigma$, . . .
• Cover more topics (via exercises)
• Makes statistics *accessible*!
• Online courses
• ebooks
“Any scientific calculator”

- Fails to give students opportunity to develop technology skills
- Misses $P$-values, graphs, normality assessment, cover more topics
Nothing has had a greater impact on statistics courses, the teaching of statistics, and the subject of statistics itself ... than *technology*.
Steve Jobs’ Garage
Technology for Graphs

• Power to convince
• Power to mislead
• Enable *understanding*
• Explore data
• Check for normality
Graphs: Power to Convince

![Graph](image)
Jai Alai Bribery

Auditor noticed some unusually large bets, when ....

Certain players won games much less often than expected, so ...

Statistician asked to analyze the data
Figure 1. $\bar{X}$ Chart for Matinee Sessions.
Figure 2. *R* Chart for Matinee Sessions.
Statistician Asked Prosecutors:
“What happened on Valentine’s Day?”

Astounded Prosecutors:
“Arrest was made on Valentine’s Day!”
“I lost the election, but something is wrong!

HELP!”
Democrat Line

Row B Proportions of Votes

- Town: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22
- Row B Proportions: 0.4, 0.5, 0.6
Conservative Line

Row C Proportions of Votes

![Graph showing Row C Proportions of Votes over towns from 1 to 22.](image)
Liberal Party Line

Row D Proportions of Votes

![Graph showing Row D Proportions of Votes with data points and town numbers from 1 to 22.](image)
Row H

Row H Proportions of Votes

Whoa!
Opposition Attorney:

“If you add the 3% margin of error to the 5% from the 95% confidence level, you get 8%, which is too high.”

Statistical illiteracy in law that greatly damaged the defense.
CNN: Remove feeding tube?

---

**Sample:** Interviews conducted by telephone March 18-20, 2005, with 909 adults in the United States.
Graphs: Power to Mislead

OBAMACARE ENROLLMENT

AS OF MARCH 27

6,000,000

MARCH 31 GOAL

7,066,000

SOURCE: HHS

media matters.org

In second place with $26.5 mil, while "MUPP DOW FUT 16.325.00"
Fox news apologizes
What do you conclude?

**Technology:**

\[ r = 0.816 \]

\[ P\text{-value} = 0.002 \]
$r = 0.816; \ P$-value $= 0.002$
Normal Quantile Plots

Alcohol use (sec) in video games (*JAMA*):
84  14  583  50  0  57  207  43  178  0  2  57

![Normal Quantile Plot](image.png)
4. Probability

Nonintuitive!

Birthday Problem: 0.569
Nashville: Cumberland River
Probability

Foundation of inferential statistics
Teaching Probability

• Basic definition:
  Probability values

• Addition Rule

• Multiplication Rule
Probability

Redundancy

• 2 alarm clocks
• Planes: electrical; hydraulic
Teaching Probability

Permutations & Combinations: No

• Do *demo* for lottery
5. Variation

Variation

...the single word that best summarizes statistics!
“There are never in nature two beings which are exactly alike.”

-Gottfried Wilhelm Leibniz
Without variation . . .
Variation: Facial Recognition
**Single** bank waiting line (minutes):

6.5  6.6  6.7  6.8  7.1  7.3  7.4  7.7  7.7  7.7

→→→

\[ s = 0.5 \]

**Multiple** lines (minutes):

4.2  5.4  5.8  6.2  6.7  7.7  7.7  8.5  9.3  10.0

↓  ↓  ↓  ↓

\[ s = 1.8 \]
"See" Variation

<table>
<thead>
<tr>
<th>Wait Time (seconds)</th>
<th>Single Line</th>
<th>Multiple Lines</th>
</tr>
</thead>
<tbody>
<tr>
<td>80</td>
<td>85</td>
<td>90</td>
</tr>
</tbody>
</table>
Teaching Variation

1. Motivate “variation” using bank lines.
2. Walk class through manual calculation of $s$.
3. Develop technology skill for finding $s$.
4. Focus on interpreting $s$ using

**Range rule of thumb:**

“Usual” results: mean $\pm 2$ standard deviation
Assessment

• Is an IQ of 135 unusual? (Given mean = 100 and st. dev. = 15)

• Estimate $s$ for heights of students in class:

$$s \approx \text{range}/4$$
6. Normal Distribution

Incredible fundamental law of nature:

*The Central Limit Theorem*

- Heights, IQs, Lotteries, Incomes, Reaction Times, ...
Normality Assessment

Obsolete: “Assume normal distribution.”

Test for normality:

• Histograms
• Normal quantile plots: Tech
• Hypothesis tests
Normal Quantile Plot

Technology: Easy

Normal if ...
1. Points close to straight line
2. No nonlinear pattern
Not Normal: Pattern

![Not Normal: Pattern Graph]
Not Normal: *Bootstrap*

Alcohol (sec) in video games (*JAMA*):
84 14 583 50 0 57 207 43 178 0 2 57
The seventh wonder of the statistical world is ...
Statistics Students!!!
Beyond teaching statistics, we can help our students **grow** through developing ...
Important Life Skills

• Critical thinking
• Technology
• Collaboration
• Speaking
*Projects/Activities:*

The #1 way to improve your statistics course!
Projects

Capstone Project:
• Groups of 3 or 4
• Oral presentation by all
• Computer printout
• Brief written report
marty@triolastats.com
www.triolastats.com

facebook.com/TriolaStats
linkedin.com/company/triola-stats