Discovery 2014

A JMP Add-In for Teaching Statistical inference Using Resampling Methods

Mia Stephens and Don McCormack

Mia.Stephens@jmp.com
Don.McCormack@jmp.com

jmp Statistical Discovery.™ From SAS.

sas THE POWER TO KNOW.

Outline

- Challenges in teaching statistics
- Directions in statistical education
- The role of resampling in statistics education
- Randomization (permutation) tests
- **A JMP randomization test add-in**

About Teaching Statistics

- How many of you teach statistics?
  - Academically?
  - In industry?
- Observation: Learning statistics, and particularly statistical inference, is difficult
- Core inferential topics:
  - Confidence interval for a mean
  - Hypothesis testing for one mean
- Do you remember the first time you saw these topics?
  - Did you get it right away?
  - Any obstacles? Challenges?

To Conduct a One Sample *t*-Test

1. State the **null** and the **alternative** hypotheses.
2. Determine the **significance level**, α.
3. Based on α and the sample size, determine the **critical t values** and the **rejection region**.
4. Calculate the **test statistic**, $t = \dfrac{\bar{X} - \mu}{S/\sqrt{n}}$

df\p	0.4	0.25	0.1	0.05	0.025	0.01
1	0.325	1.000	3.078	6.314	12.706	31.821
2	0.289	0.816	1.886	2.920	4.303	6.965
3	0.277	0.765	1.638	2.353	3.182	4.541
4	0.271	0.741	1.533	2.132	2.776	3.747
5	0.267	0.727	1.476	2.015	2.571	3.365
6	0.265	0.718	1.440	1.943	2.447	3.143
7	0.263	0.711	1.415	1.895	2.365	2.998
8	0.262	0.706	1.397	1.860	2.306	2.896
9	0.261	0.703	1.383	1.833	2.262	2.821
10	0.260	0.700	1.372	1.812	2.228	2.764
11	0.260	0.697	1.363	1.796	2.201	2.718
12	0.259	0.695	1.356	1.782	2.179	2.681
13	0.259	0.694	1.350	1.771	2.160	2.650
14	0.258	0.692	1.345	1.761	2.145	2.624
15	0.258	0.691	1.341	1.753	2.131	2.602
16	0.258	0.690	1.337	1.746	2.120	2.583
17	0.257	0.689	1.333	1.740	2.110	2.567
18	0.257	0.688	1.330	1.734	2.101	2.552
19	0.257	0.688	1.328	1.729	2.093	2.539
20	0.257	0.687	1.325	1.725	2.086	2.528
21	0.257	0.686	1.323	1.721	2.080	2.518
22	0.256	0.686	1.321	1.717	2.074	2.508
23	0.256	0.685	1.319	1.714	2.069	2.500
24	0.256	0.685	1.318	1.711	2.064	2.492
25	0.256	0.684	1.316	1.708	2.060	2.485

5. Compare the **test statistic** to the **critical values** and the rejection region to either **reject** or **fail to reject** the null hypothesis, interpret the **p-value**.
6. **Summarize** your conclusions (interpret).

Some Challenges

- Population versus sample
- Descriptive statistics
- Theoretical probability distributions (Normal)
- Sampling distributions (t)
- Greek symbols – mu, sigma, alpha
- Formulas and notation (subscripts)
- Null and alternative hypothesis
- Look up tables
- Confidence level
- Degrees of freedom
- Critical value
- Tail of distribution
- Percentile
- P-values
- Type I error
- Confusing language
- **Interpretation**

An Image Problem?

*"Historically the discipline and methods of statistics have been viewed by many students as a difficult topic which is **unpleasant** to learn."*

*"It is not uncommon for people to recount tales of statistics as the **worst course** they took in college".*

*"Despite attempts…the image of statistics as a **hard** and **dreaded** subject is hard to dislodge."*

Joan Garfield and Dani Ben-Zvi

The Discipline of Statistics Education ('07)

Directions in Statistics Education

Attempts to improve and standardize statistics education:

Common Core State Standards and **GAISE Recommendations**

Common Core: State-led effort to provide a clear and consistent framework to prepare students for college and the workforce.

Directions in Statistics Education

Example of Core Standards in **Probability and Statistics**:

Grade 7

- Use ***random sampling*** to ***draw inferences*** about a population.
- Draw informal comparative ***inferences about two populations***.
- Investigate chance processes and develop, use, and evaluate *probability models*.

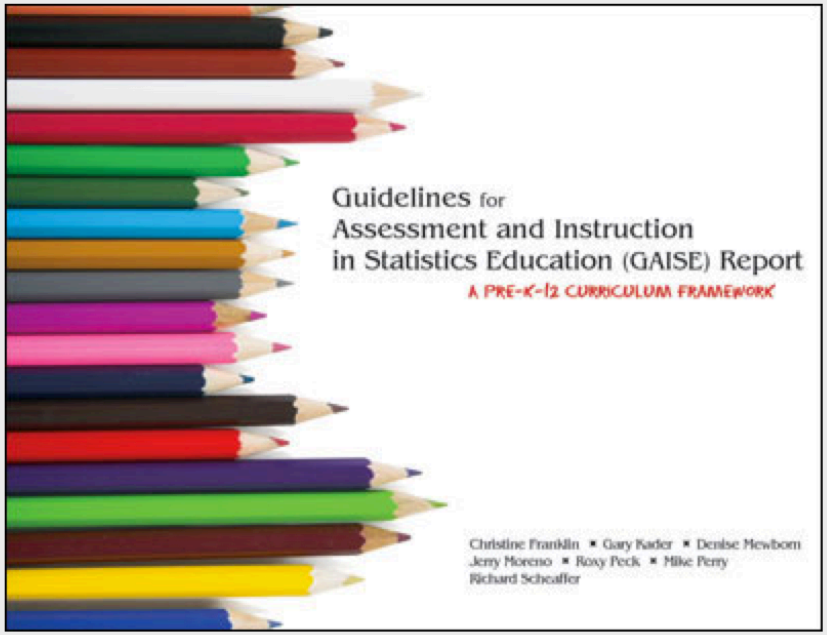
Note: The Common Core Standards do not tell teachers specifically HOW to teach.

Directions in Statistics Education

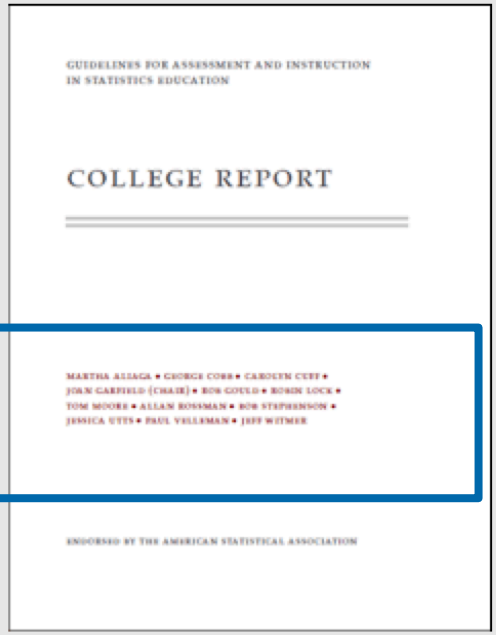
The HOW is provided by the GAISE report:

Guidelines for Assessment and Instruction in Statistics Education

PreK-12 Report



College Report



Directions in Statistics Education

- **George Cobb
- Martha Aliaga
- Carolyn Cuff
- **Joan Garfield**
- **Rob Gould**
- **Robin Lock**
- Tom Moore
- **Allan Rossman**
- Bob Stephenson
- Jessica Utts
- Paul Velleman
- Jeff Witmer

Many of these thought leaders teach and lead research on ***resample methods*** to address the challenges of teaching statistical inference.

The Role of Resampling Methods in Stat Ed

"...What we teach is largely the technical machinery of numerical approximations based on the normal distribution and its many subsidiary cogs. This machinery was once necessary, because the ***conceptually simpler alternative based on permutations*** *was computationally beyond our reach. Before computers statisticians had no choice.* ***These days we have no excuse****."*

George Cobb, JISE (07)

What Do we Mean by "Resampling" Methods?

Two general types of resample methods are being used by statistics educators:

- **Bootstrapping**: to introduce confidence intervals.

  Involves randomly sampling with replacement from the original sample.

- ****Permutation (or Randomization) Tests**: to introduce hypothesis tests.

  Generally involves randomizing or "shuffling" the labels for data points.

  Note: not randomization exact tests.

The Role of Resampling Methods in Stat Ed

Why introduce inference via resampling methods?

- **Few prerequisites** – understanding of basic concepts (Dot plot, descriptive statistics, random sampling, parameter versus statistic).
- **Delays the more theoretical discussions** of normal and t-based inference and distributional assumptions.
- **Easier**: A **more direct** progression, from the sample and sample statistics to inferences about the population.
- Approach is **general** - applies to other statistics and more complex inferential situations.

Many traditional challenges go away, and the conceptual understanding is established early on.

What is a Randomization Test?

Randomization (permutation) tests:

- Simulate the distribution of sample statistics (or test statistics) one would observe *under the null hypothesis.*
- Compare the observed statistic to this simulated distribution.
- Provides an *empirical estimate* of the p-value.

Delay the background theory and **mechanics** to focus on understanding the concept.

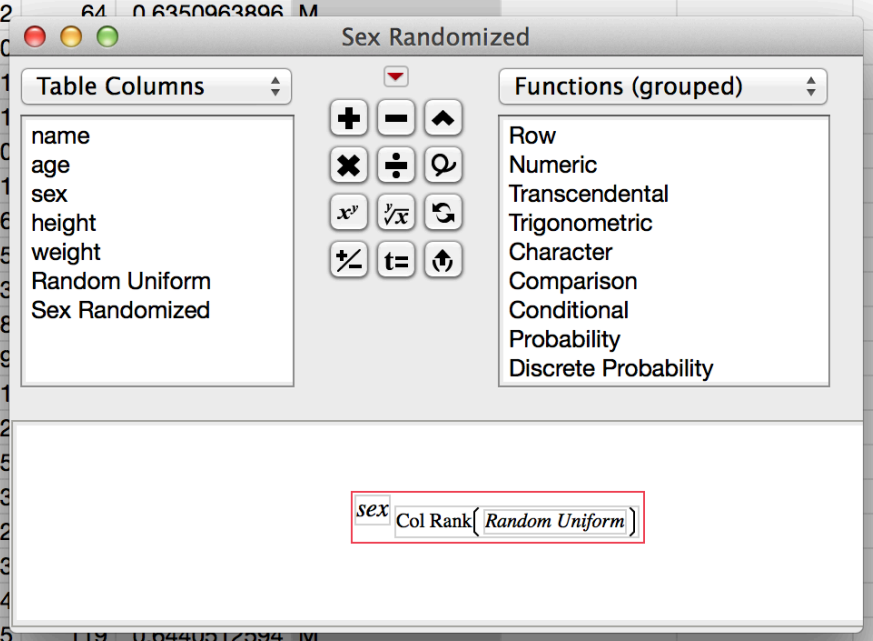
Same framework can be used to introduce different types of hypothesis tests.

What is a Randomization Test?

The basic idea – comparing two means:

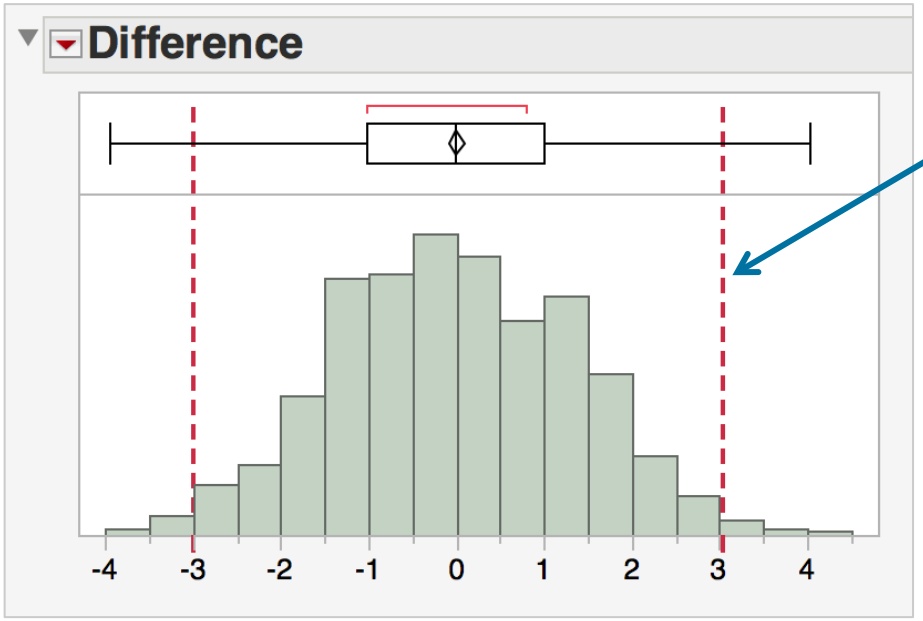
1. Shuffle the levels of the factor.

	name	age	sex	height	weight	Random Uniform	Sex Randomized
1	KATIE	12	F	59	95	0.7727489718	M
2	LOUISE	12	F	61	123	0.1160456382	M
3	JANE	12	F	55	74	0.7356067428	M
4	JACLYN	12	F	66	145	0.1970257347	F
5	LILLIE	12	F	52	64	0.6350963896	M
6	TIM	12	M	60			
7	JAMES	12	M	61			
8	ROBERT	12	M	51			
9	BARBARA	13	F	60			
10	ALICE	13	F	61			
11	SUSAN	13	F	56			
12	JOHN	13	M	65			
13	JOE	13	M	63			
14	MICHAEL	13	M	58			
15	DAVID	13	M	59			
16	JUDY	14	F	61			
17	ELIZABETH	14	F	62			
18	LESLIE	14	F	65			
19	CAROL	14	F	63			
20	PATTY	14	F	62			
21	FREDERICK	14	M	63			
22	ALFRED	14	M	64			
23	HENRY	14	M	65	119	0.6440512594	M



What is a Randomization Test?

2. Compute the averages for each level and the difference between the averages.
3. Repeat many times (say, 1000 or 10,000).
4. Graph the differences.



Question: How extreme is our observed difference?

Provides a conceptual understanding of a *p*-value, an "empirical *p*-value".

Does Resampling Address Our Challenges?

- **Population versus sample**
- **Descriptive statistics**
- Theoretical probability distributions (Normal)
- Sampling distributions (t)
- Greek symbols – mu, sigma, alpha
- Formulas and notation (subscripts)
- Null and alternative hypothesis
- Look up tables
- Confidence level
- Degrees of freedom
- Critical value
- **Tail of distribution**
- **Percentile**
- **P-values**
- Type I error
- Confusing language
- **Interpretation**

A JMP Randomization Test Add-In

Randomization Test Add-In (under development)

Current *test options*:

- Single mean
- Two means (unpaired)
- One proportion

Randomization methods available:

- *Resampling* with replacement
- *Shuffling* group membership
- *Shifting* to the null with replacement

Add-In Organization

- Because of size, code was split into multiple files with similar routines:
  - Main (driver for functions in other files)
  - Dialog Box
  - Statistics
  - Graphics
  - Data
  - Randomization Tests
- Encapsulation was used to hide file specific routines from global routines.

  Namespaces were used to define what was global.

  Names Default to Here(1) made file specific functions private.

Add-In Organization

Main

Run Script | Debug Script | Balance | Import as Data

Definition of global namespace

```
1  //MAIN
2
3  Clear Globals();
4
5  ns = New Namespace( "JMP_RT" );
6  Include("Globals.JSL");
7  Include("Data Sets.JSL");
8  Include("Randomization Tests Methods.JSL");
9  Include("Data Methods.JSL");
10 Include("Graphs Methods.JSL");
11 Include("Statistics Methods.JSL");
12 Include("Dialog Box.JSL");
13
14
```

Main Driver

Keeps routines private unless scoped otherwise

```
1  /* ----------------------------------------
2     DATA ACCESS ROUTINES
3     ----------------------------------------
4     Names Default to Here(1);
5     dataStruct = Associative Array(
6        {
7           {"Table Ref",     Empty()},
8           {"Table Name",    ""},
9           {"Table Path",    Empty()},
10          {"N Obs",         0},
11          {"Y Names",       Empty()},
12          {"X Names",       Empty()},
13          {"Y Data",        Empty()},
14          {"X Data",        Empty()},
15          {"Y Orders",      Empty()},
16          {"X Orders",      Empty()},
17          {"Format",         1}
18       }
19    );
20
21    getNObs          = Expr(dataStruct["N
22    getYNames        = Expr(dataStruct["Y
23    getYData         = Expr(dataStruct["Y
24    getYOrders       = Expr(dataStruct["Y
25    getXNames        = Expr(dataStruct["X
26    getXdata         = Expr(dataStruct["X
27    getXOrders       = Expr(dataStruct["X
```

Data (private) file

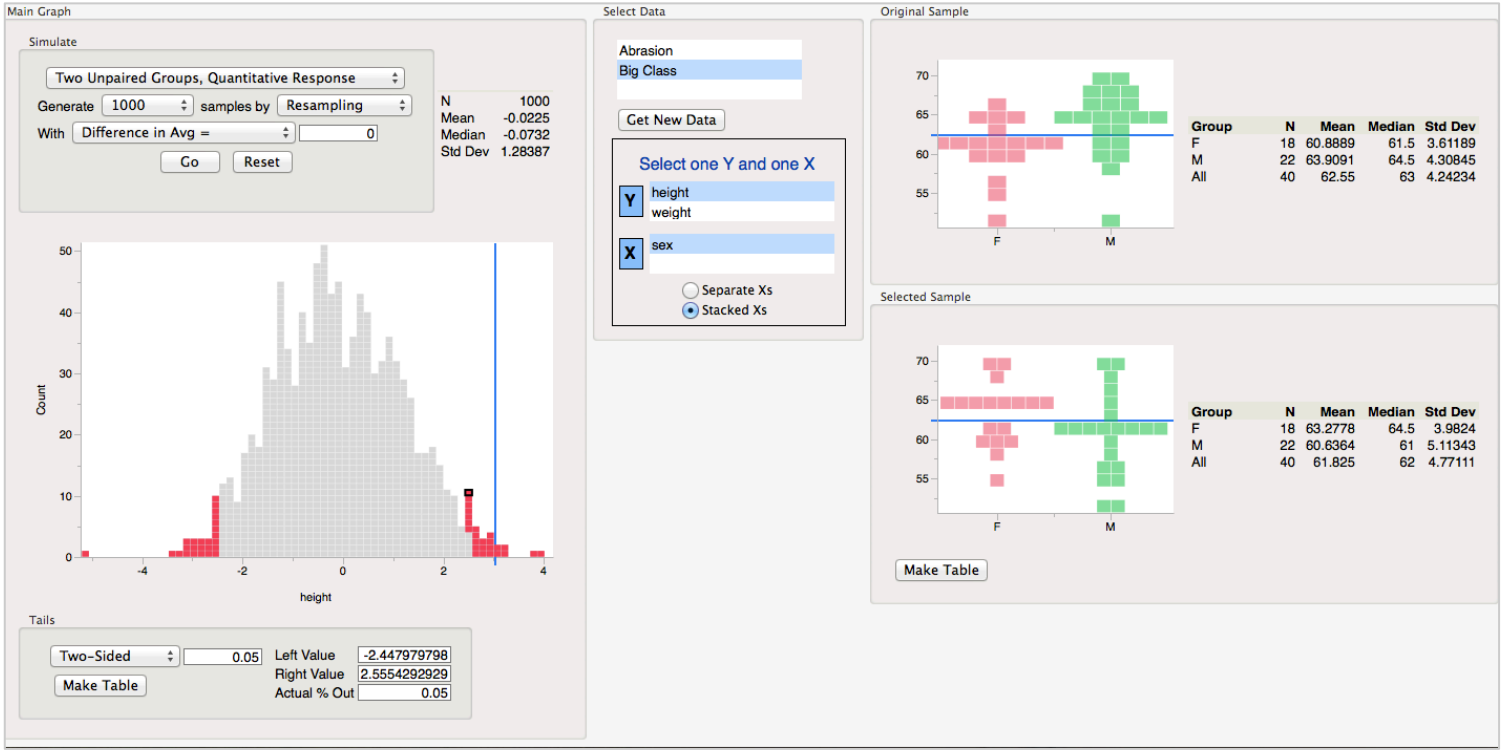
Public routines

```
1     Include("Data.JSL");
2
3  /* ..............
4     PUBLIC METHODS
5     PUBLIC METHODS
6     PUBLIC METHODS
7     .............. */
8     JMP_RT:getActiveTblName = Expr(getActive
9     JMP_RT:getActiveTblRef  = Expr(getActiveT
10
11    JMP_RT:getNObs      = Expr(getNObs);
12    JMP_RT:getYNames    = Expr(getYNames);
13    JMP_RT:getXNames    = Expr(getXNames);
14    JMP_RT:getYData     = Expr(getYData);
15    JMP_RT:getXData     = Expr(getXData);
```

Data Methods (public) file

Add-In Interface

- Compact organization to fit in a small space
- Minimize moving parts
- Create run time widgets



Graphic Challenges

- Dot plots are hard to make in JSL.

  Block plots were created to provide a comparable alternative.

- Graphs had to be interactive without relying on JMP's built-in hyperlinking.

  Mousetrap routines were used.

- Graph modification was needed to add dynamic graphical elements based on user input.

  Add/Remove Graphics Script used to facilitate this.

Other Challenges

- Data Structures

  Needed a Struct like composite data structure to make information passing between modules and user customization easier.

  Associative Arrays were used to do this.

- Dealing with randomization of large lists.

  Convert categorical lists values to ordinal levels so that randomization could be done using matrices.

Next Steps

- Add tests for:
  - Paired t-Test
  - Simple linear regression
  - Two-way tables
  - ANOVA
- Package as a JMP Add-In
- Testing internally and in the classroom
- Make available via file exchange

References

- Aliaga et al. (2012). *Guidelines for Assessment and Instruction in Statistics Education: College Report.* American Statistical Association. http//www.amstat.org/education/gaise/.
- Cobb, G. (2007), *The Introductory Statistics Course: A Ptolemaic Curriculum*. Technology Innovations in Statistics Education, Volume 1 (1), Article 1. http://escholarship.org/uc/item/6hb3k0nz#page-1.
- Cobb, G. (2013), *Comment: Technology and the Future of Statistics Education*. Technology Innovations in Statistical Education, Volume 7 (3), Page 6. http://escholarship.org/uc/item/1j7116jx#page-6
- Efron, B. (1979), *Bootstrap Methods: Another look at the jackknife*. The Annals of Statistics, Volume 7(1).
- Good, P. (2013), *Introduction to Statistics Through Resampling Methods and R*, 2E, Wiley.
- JMP®, Version 11 Pro, SAS Institute Inc., Cary, NC, 1989–2013.
- Lock et al. (2013), *Statistics, Unlocking the Power of Data*. Wiley.
- Mills, J. D. (2002), *Using computer simulation methods to teach statistics: a review of the literature*. Journal of Statistics Education (Online), Volume 10 (1). www.amstat.org/publications/jse/v10n1/mills.html
- Tintle, et al, *Introduction to Statistical Investigations*. Wiley, Forthcoming.
- Wood, M. (2005). *The Role of Simulation Approaches in Statistics*. Journal of Statistics Education, Volume 13 (3). www.amstat.org/publications/jse/v13n3/wood.html.

Questions?

Thank you!