

Art of Stat

Resampling on your phone

Bernhard Klingenberg
(bklingen@artofstat.com)

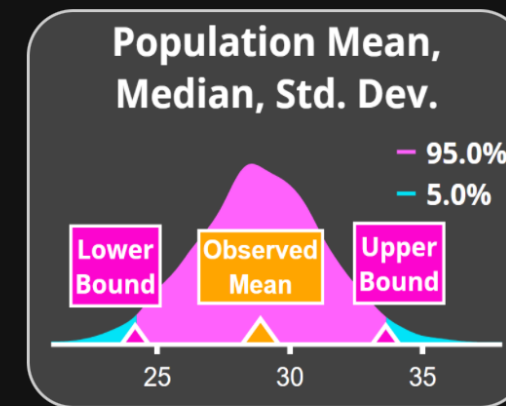
3:00

5G 49

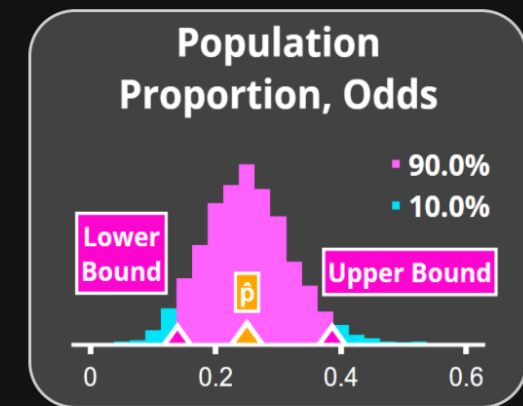
ART OF STAT Resampling



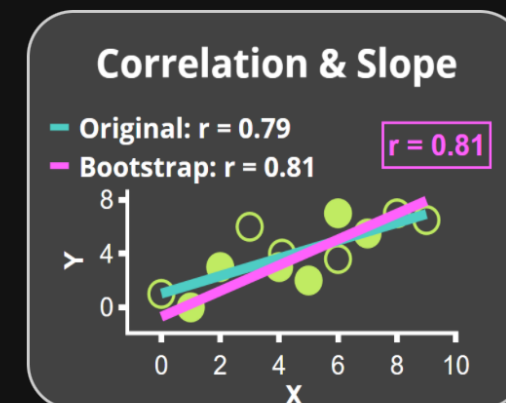
Bootstrap Confidence Intervals:



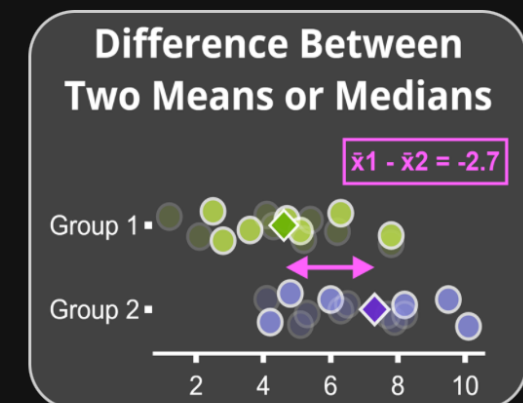
Mean, Median,
Std. Deviation



Proportion & Odds



Correlation & Slope



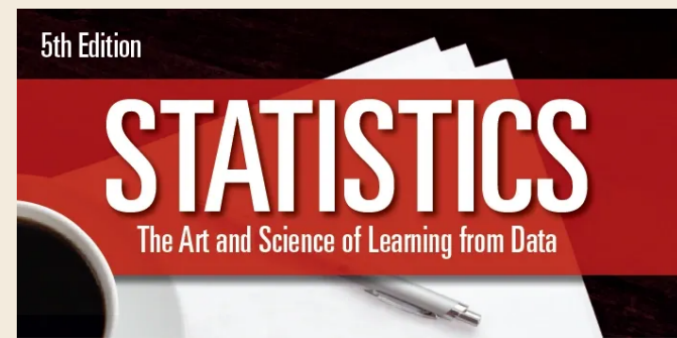
Differences in Means or Medians



Find slides and much more info on mobile apps at ArtofStat.com

Introductory Statistics Textbook & Apps

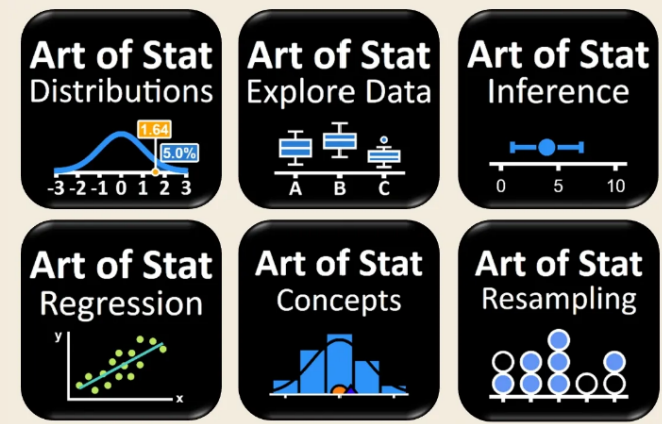
Intro Stats Textbook



Statistics: The Art and Science of Learning from Data, 5th edition.
Authors: Agresti, Franklin, Klingenberg
An introductory statistics textbook for a one or two-semester course.

[Go to Publisher's Website](#) >

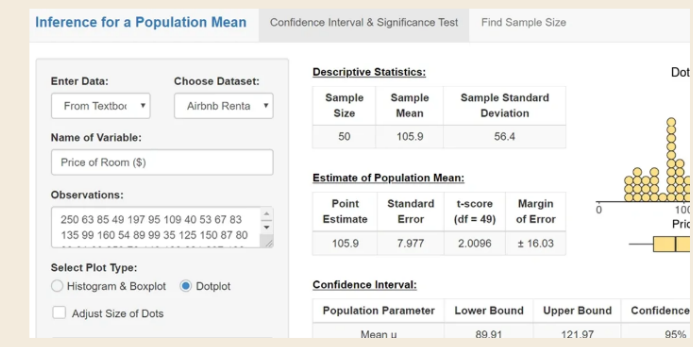
Mobile Apps for your Phone



Mobile apps for exploring and working with distributions and data. All from your phone. Download it now for [iOS](#) or [Android](#).

[Find out more](#) >

Online Web Apps

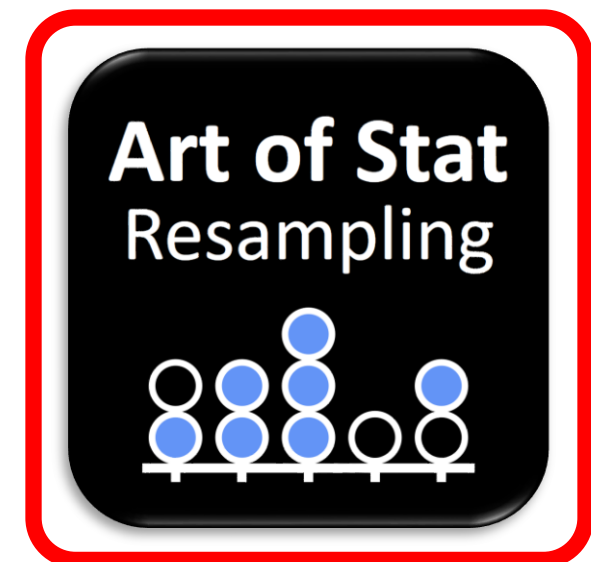
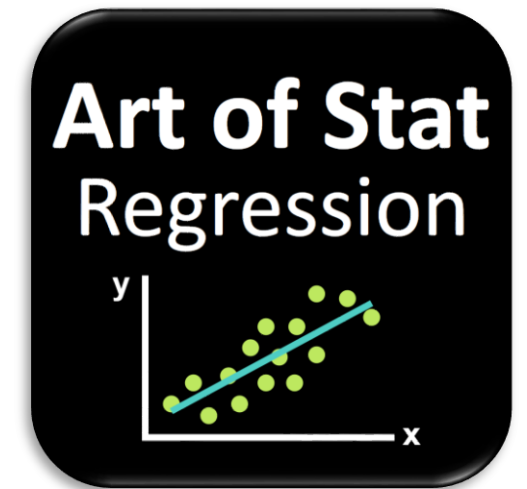
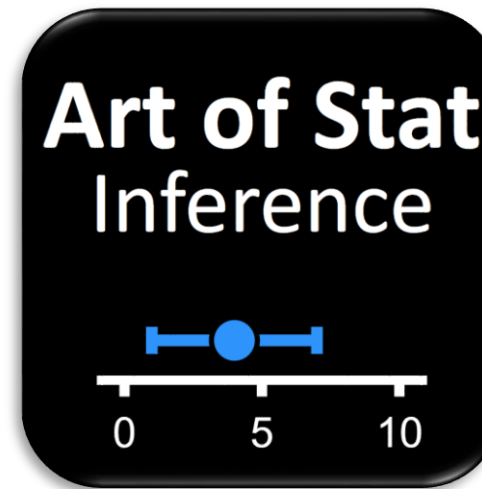
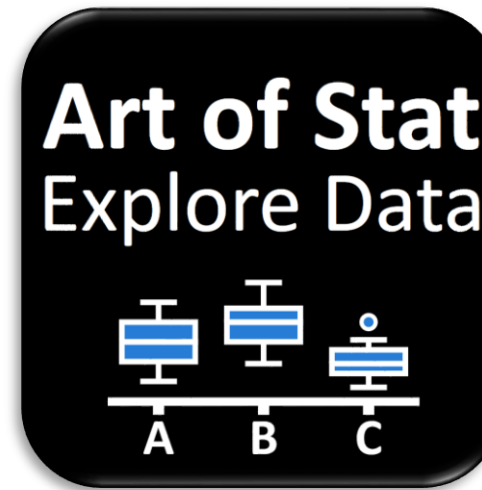
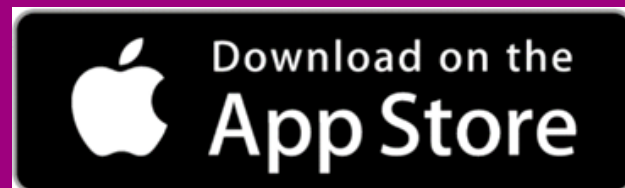


Online **interactive web-apps** for illustrating & learning statistics. Explore statistical concepts or carry out data analysis (descriptive & inferential), right in your browser.

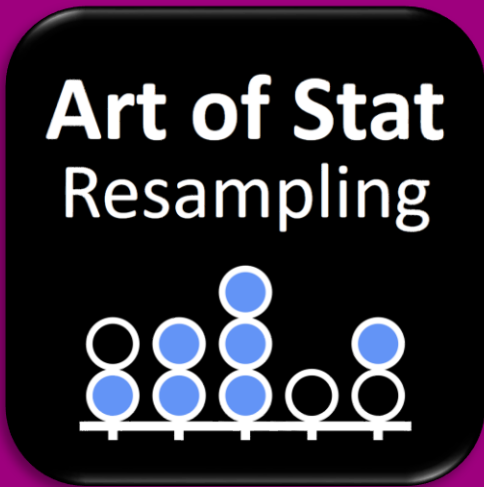
[Go to Web Apps](#) >

The Six Art of Stat Mobile Apps

- Explore Data
- Inference
- Concepts
- Distributions
- Regression
- Resampling



Art of Stat Resampling



- Bootstrap Confidence Intervals

- Permutation Tests

3:00 5G 49

ART OF STAT Resampling

Bootstrap Confidence Intervals:

Population Mean, Median, Std. Dev.

95.0%
5.0%

Lower Bound Observed Mean Upper Bound

25 30 35

Mean, Median, Std. Deviation

Population Proportion, Odds

90.0%
10.0%

Lower Bound Upper Bound

0 0.2 0.4 0.6

Proportion & Odds

Correlation & Slope

Original: $r = 0.79$
Bootstrap: $r = 0.81$

$r = 0.81$

8
4
0

0 2 4 6 8 10

Correlation & Slope

Difference Between Two Means or Medians

$\bar{x}_1 - \bar{x}_2 = -2.7$

Group 1

Group 2

2 4 6 8 10

Differences in Means or Medians

3:00 5G 49

Permutation Tests:

Permutation Test Mean & Median

$\bar{x}_{perm} = 19$
 $\bar{x}_{obs} = 31$

10 20 30 40 50

Permutation Test for the Mean or Median

Permutation Test Two Means or Medians

$\bar{x}_1 - \bar{x}_2 = 1.4$

Group 1

Group 2

1 2 3 4 5 6

Permutation Test Comparing Two Groups

Chi-Square Test (Independence/Homogeneity)

Group 1

Group 2

Group 3

0% 50% 100%

Under Construction

Permutation Chi-Square Test

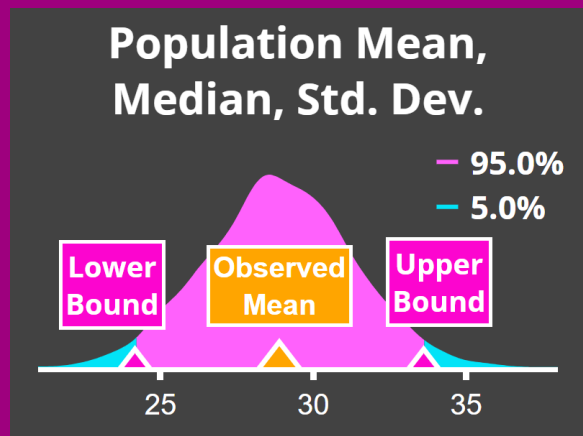
Create Or Edit Your Own Datasets

Data Editor

Obs.	Animal	Longevity

Art of Stat Resampling

Bootstrap Interval for a Population Mean or Median



- Screens:
- Enter Data
 - Bootstrap Distrib.
 - Confidence Interval

10:06

Bootstrap Interval

Enter Data:
Open Example Datasets

Choose from several example datasets. Press the help icon in the upper right corner to find descriptions and use cases:

Select Dataset:
Animals

Animal	Family	Common Name	Gestation (days)
1	Canidae	Maned wolf	64
2	Canidae	Gray fox	57
3	Canidae	Black-backed jackal	60
4	Canidae	Crab-eating fox	55

Showing all 24 rows, and all 8 columns (swipe left).

Select Variable:
Longevity (yrs)

Population Parameter of Interest:
Population Mean

Descriptive Statistics:

Statistic	Value
Sample Size (n)	24
Sample Mean (\bar{x})	19.00
Sample Standard Deviation (s)	5.64
Minimum	9
First Quartile (Q1)	15.5
Median	19
Third Quartile (Q3)	22
Maximum	32
Interquartile Range (IQR)	6.5
Range (Max - Min)	23

Adjust Number of Digits

Histogram

ENTER DATA

BOOTSTRAP DISTRIBUTION

CONFIDENCE INTERVAL

10:06

Bootstrap Interval

Descriptive Statistics:

Statistic	Value
Sample Size (n)	24
Sample Mean (\bar{x})	19.00
Sample Standard Deviation (s)	5.64
Minimum	9
First Quartile (Q1)	15.5
Median	19
Third Quartile (Q3)	22
Maximum	32
Interquartile Range (IQR)	6.5
Range (Max - Min)	23

Adjust Number of Digits

Histogram

ENTER DATA

BOOTSTRAP DISTRIBUTION

CONFIDENCE INTERVAL

10:07

Bootstrap Interval

Original Sample: n = 24, \bar{x} = 19.00, s = 5.64

Bootstrap Sample:

Press a Button Below To Generate Bootstrap Samples

Bootstrap Distribution of the Sample Mean
Simulations: 0

Draw 1 Sample

Draw 100 Samples

Draw 1000 Samples

Reset

Zoom In

Normal Sampling Distb.

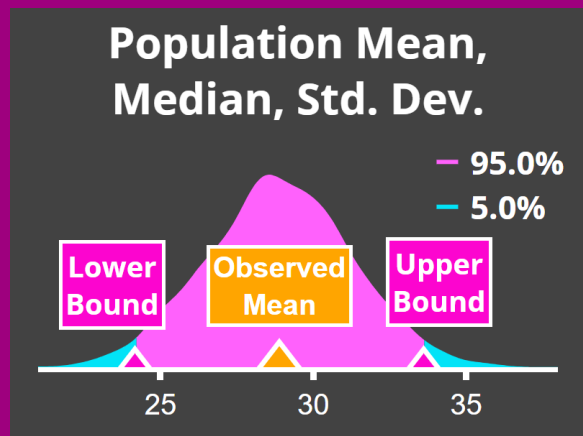
Last Bootstrap Sample Generated

BOOTSTRAP DISTRIBUTION

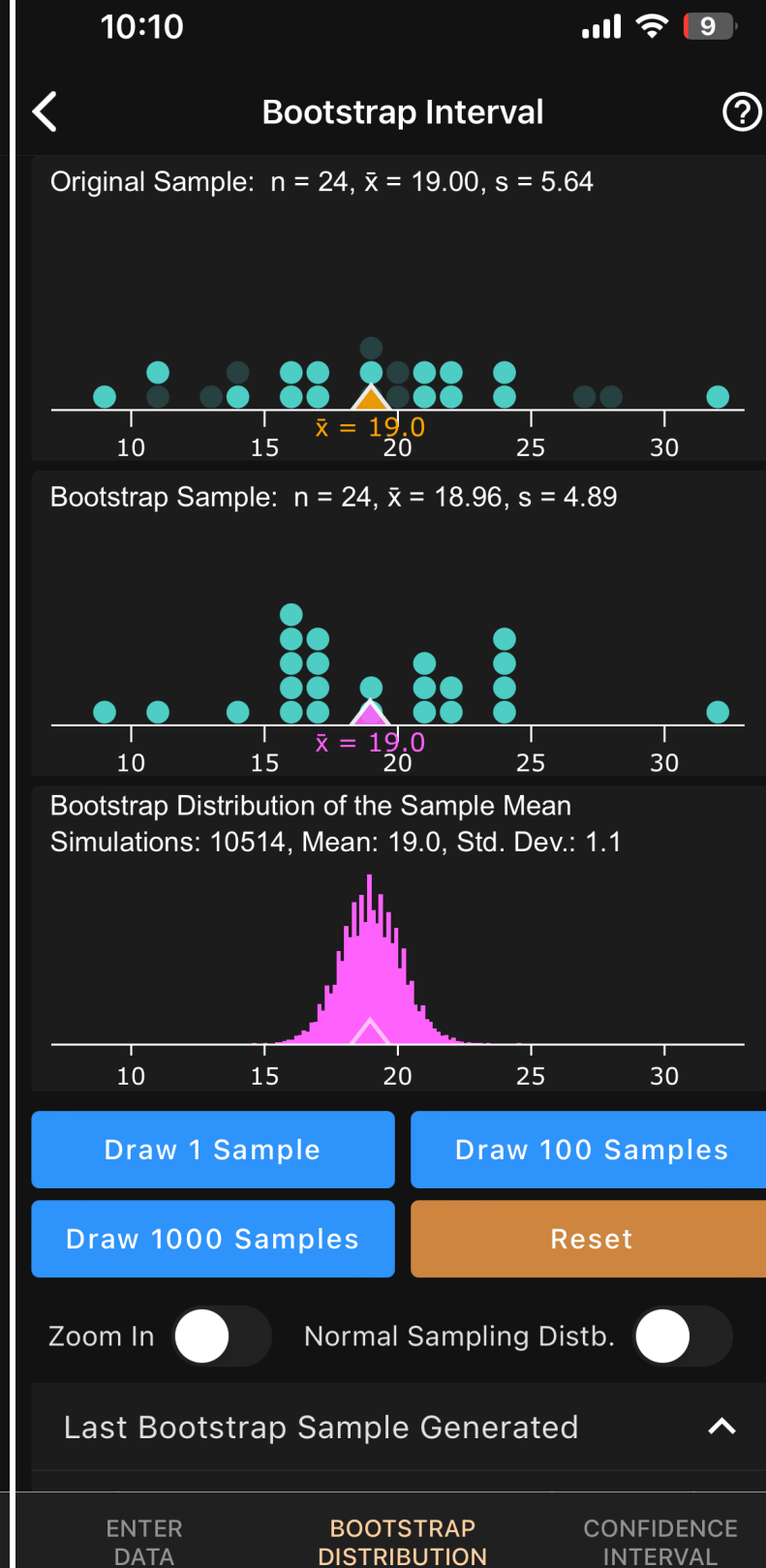
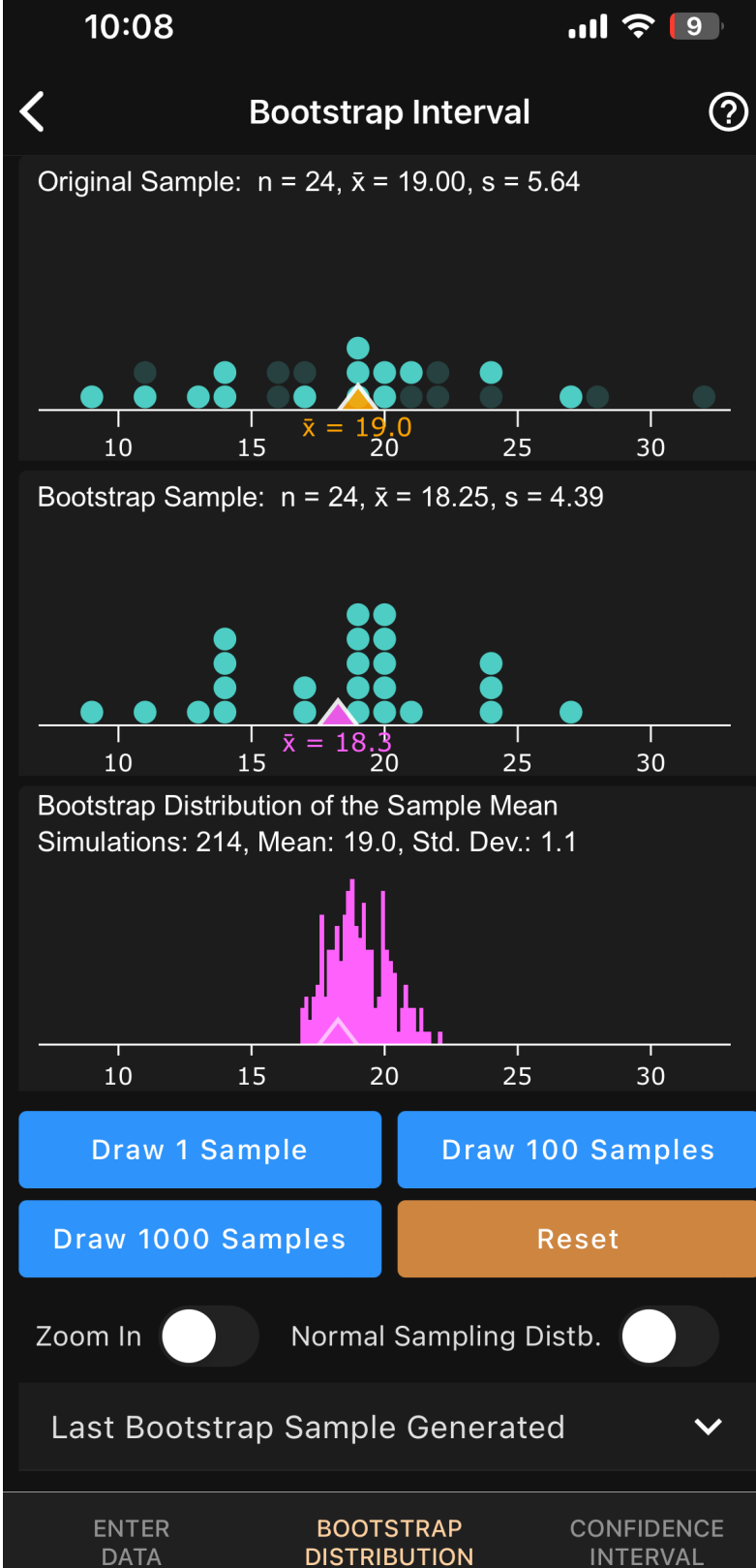
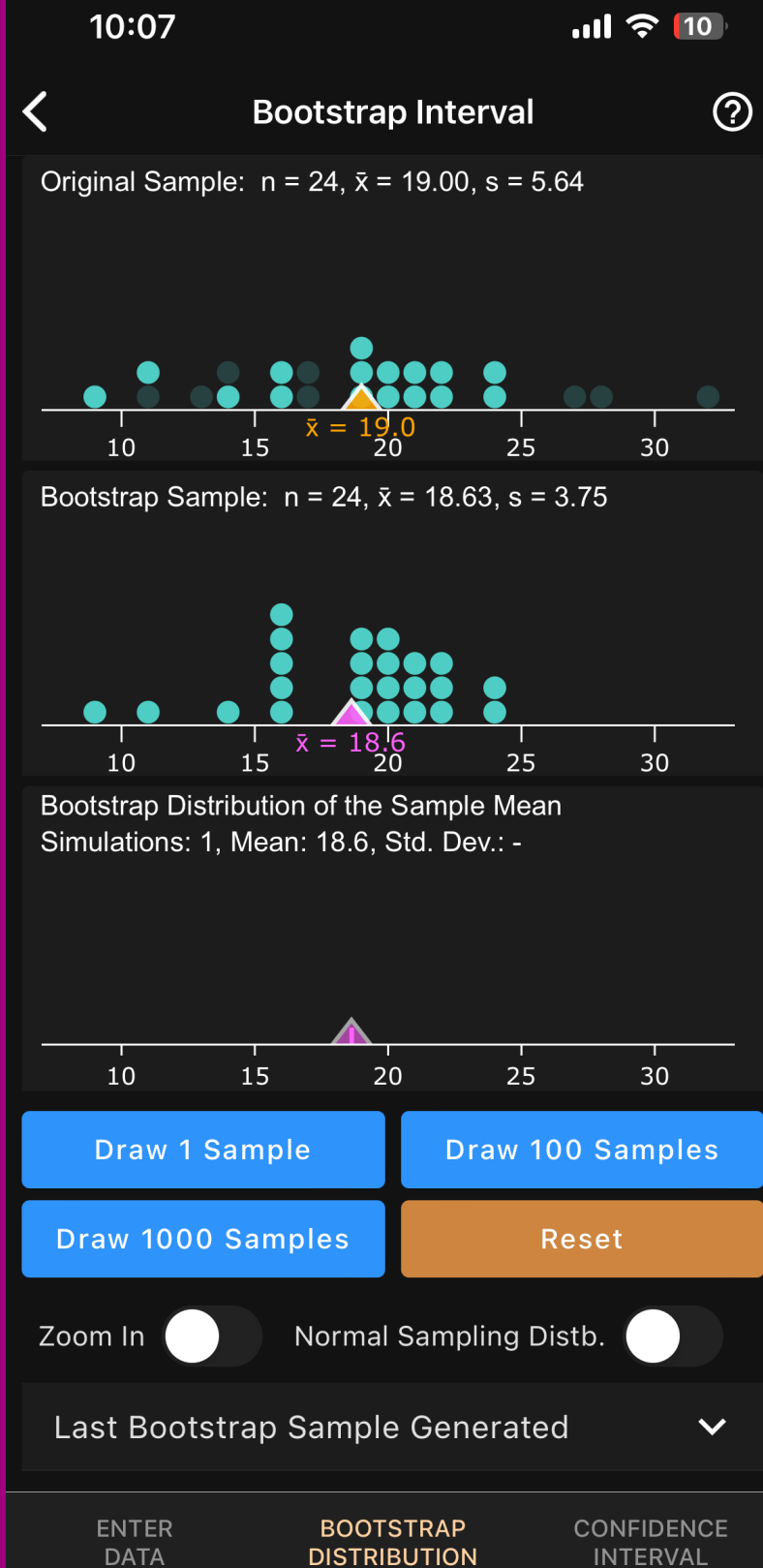
CONFIDENCE INTERVAL

Art of Stat Resampling

Bootstrap Interval for a Population Mean or Median

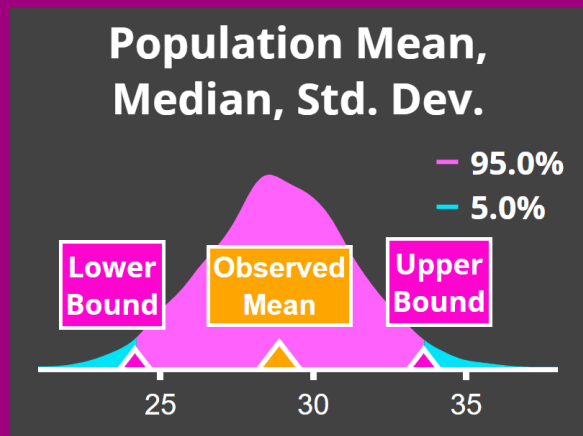


- Screens:
- Enter Data
 - **Bootstrap Distrib.**
 - Confidence Interval



Art of Stat Resampling

Bootstrap Interval for a Population Mean or Median



10:10

Bootstrap Interval

Distribution of the Sample Mean

95% Confidence Interval: (16.8, 21.2)

Observed Mean Smooth Distb.

Summary Statistics Normal Distb.

Inference for the Population Mean

Statistic	Bootstrap Percentile Interval
Sample Mean	19.0
Standard Error	1.1
95% Lower Bound	16.8
95% Upper Bound	21.2

Confidence Level: Interval Type:

ENTER DATA BOOTSTRAP DISTRIBUTION CONFIDENCE INTERVAL

10:12

Bootstrap Interval

95% Confidence Interval: (16.8, 21.2)

Observed Mean Smooth Distb.

Summary Statistics Normal Distb.

Summary of Bootstrap Distribution

Statistic	Value
Number of Simulations	10514
Unique Values	176
Mean of Distribution	19.0
Standard Deviation	1.1
Minimum	14.6
Lower Percentile (2.5%)	16.8
First Quartile (25%)	18.2
Median	19.0
Third Quartile (75%)	19.8
Upper Percentile (97.5%)	21.2
Maximum	24.6

ENTER DATA BOOTSTRAP DISTRIBUTION CONFIDENCE INTERVAL

10:12

Bootstrap Interval

Observed Mean Smooth Distb.

Summary Statistics Normal Distb.

Inference for the Population Mean

Statistic	Bootstrap Percentile Interval	Student t Interval
Sample Mean	19.0	19.0
Standard Error	1.1*	1.2**
95% Lower Bound	16.8	16.6
95% Upper Bound	21.2	21.4

* Standard error based on bootstrap distribution.
** Standard error based on formula.

Confidence Level: 95 Interval Type: Two-Sided

Show Student t-Interval

Method of Constructing Bootstrap Interval: Bootstrap Percentile Method

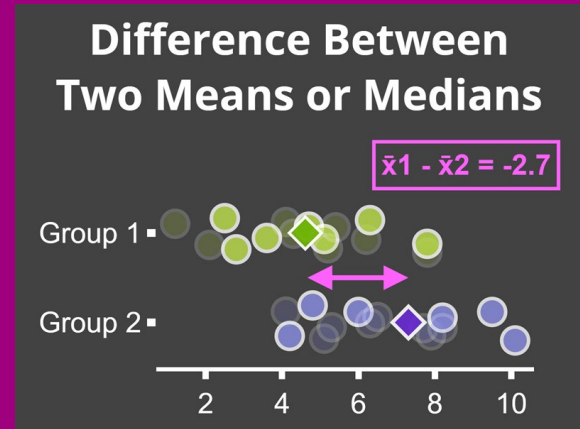
ENTER DATA BOOTSTRAP DISTRIBUTION CONFIDENCE INTERVAL

Screens:

- Enter Data
- Bootstrap Distrib.
- Confidence Interval

Art of Stat Resampling

Bootstrap Interval for the Difference Between Two Population Means



- Screens:
- Enter Data
 - Bootstrap Distrib.
 - Confidence Interval

10:21

Bootstrap Two Samples

Enter Data:
Open Sample Datasets

Choose from several example datasets. Press the help icon in the upper right corner to find descriptions and use cases:

Select Dataset:
Class Survey

Student	Frosh	Chocolate	Textbook Costs (\$)
1	Yes	milk	550
2	Yes	milk	190
3	No	milk	300
4	Yes	milk	70
5	Yes	dark	0

Showing all 70 rows, and all 8 columns (swipe left).

Group Variable: Frosh
Response Variable: Textbook Costs (\$)

Select a grouping variable that indicates group membership, and select a response variable that holds the observations.

Group 1: Yes
Group 2: No

ENTER DATA BOOTSTRAP DISTRIBUTION CONFIDENCE INTERVAL

10:23

Bootstrap Two Samples

Population Parameter of Interest:
Difference Between Means

Descriptive Statistics for Textbook Costs (\$):

Statistic	Yes	No
Sample Size (n)	36	34
Sample Mean (\bar{x})	246.75	191.62
Sample Standard Deviation (s)	229.96	195.40
Standard Error (se) of the Sample Mean	38.33	33.51
Minimum	0	0
First Quartile (Q1)	73.75	0
Median	200	125
Third Quartile (Q3)	300	342.5
Maximum	1000	650
Interquartile Range (IQR)	226.25	342.5
Range (Max - Min)	1000	650

Difference in Sample Means: $\bar{x}_1 - \bar{x}_2 = 55.1$

ENTER DATA BOOTSTRAP DISTRIBUTION CONFIDENCE INTERVAL

10:23

Bootstrap Two Samples

Side-by-Side Boxplots

Textbook Costs (\$)

Show Sample Means on Boxplot

Show Observations on Boxplot

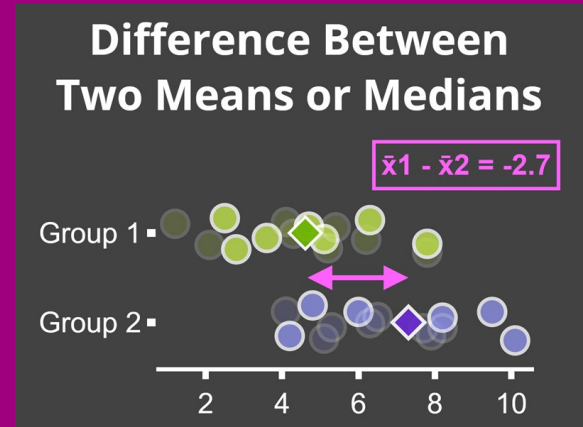
Histograms

Textbook Costs (\$) (Yes)

ENTER DATA BOOTSTRAP DISTRIBUTION CONFIDENCE INTERVAL

Art of Stat Resampling

Bootstrap Interval for the Difference Between Two Population Means



- Screens:
- Enter Data
 - **Bootstrap Distrib.**
 - Confidence Interval

10:24

Bootstrap Two Samples

Original Data

Yes: $n_1 = 36, \bar{x}_1 = 246.75$

No: $n_2 = 34, \bar{x}_2 = 191.62$

$\bar{x}_1 - \bar{x}_2 = 55.13$

Textbook Costs (\$)

Bootstrap Distribution

Simulations: 0

Draw 1 Sample

Draw 100 Samples

Draw 1000 Samples

Reset

Zoom In

Normal Approximation

Jitter Points in Group 1 and Group 2

ENTER DATA

BOOTSTRAP DISTRIBUTION

CONFIDENCE INTERVAL

10:25

Bootstrap Two Samples

Bootstrapped Data

Yes: $n_1 = 36, \bar{x}_1 = 277.03$

No: $n_2 = 34, \bar{x}_2 = 163.82$

$\bar{x}_1 - \bar{x}_2 = 113.20$

Textbook Costs (\$)

Bootstrap Distribution

Simulations: 1, Mean: 113.20, Std. Dev.: -

Draw 1 Sample

Draw 100 Samples

Draw 1000 Samples

Reset

Zoom In

Normal Approximation

Jitter Points in Group 1 and Group 2

ENTER DATA

BOOTSTRAP DISTRIBUTION

CONFIDENCE INTERVAL

10:26

Bootstrap Two Samples

200	250	70	145	95
500	500	0	800	10
240	245	245	300	160
0	500	240	95	245
250	268	500	300	75
300	160	250	145	0
160	1000	1000	35	500
190				

Last Bootstrap Sample Generated (Group 2)

380	380	350	350	0
0	0	0	0	380
200	50	60	400	300
240	0	0	650	20
200	0	0	120	0
0	400	130	0	60
400	0	0	500	

Generated Differences in Means (1 Simulations)

113.20

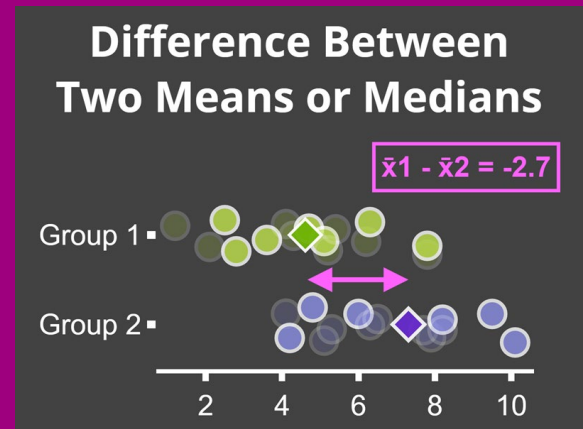
ENTER DATA

BOOTSTRAP DISTRIBUTION

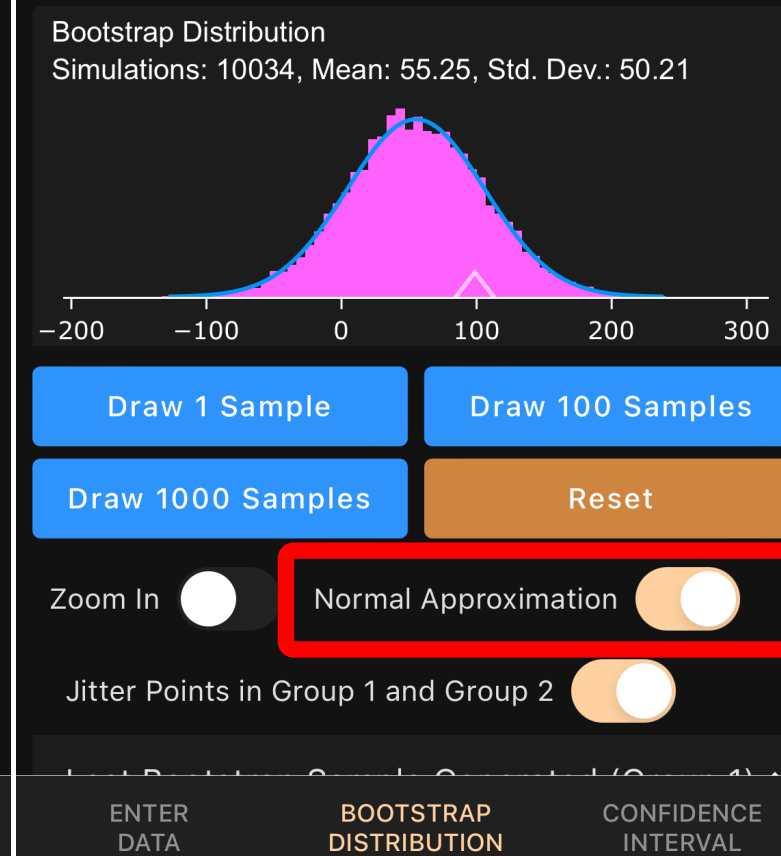
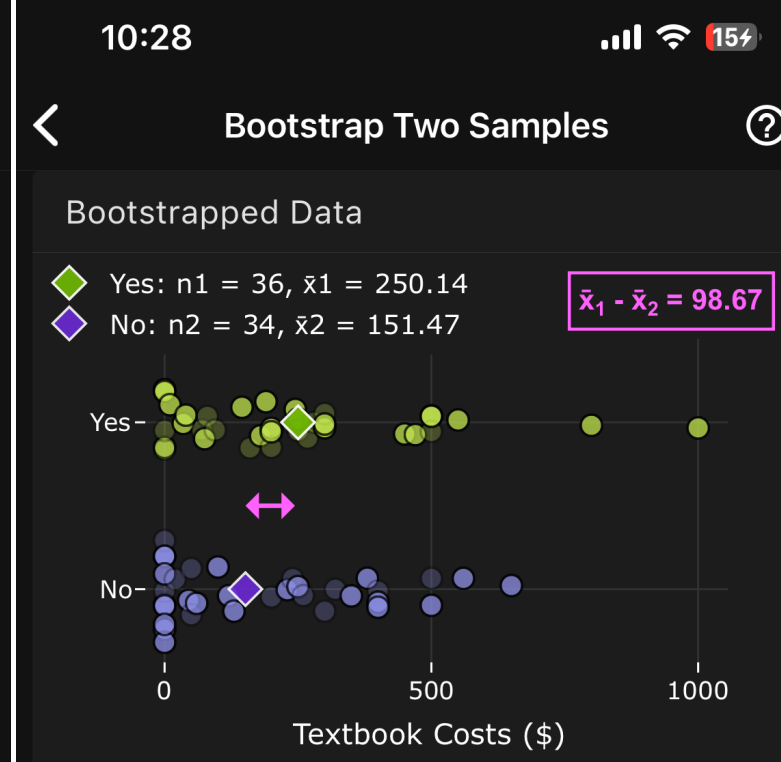
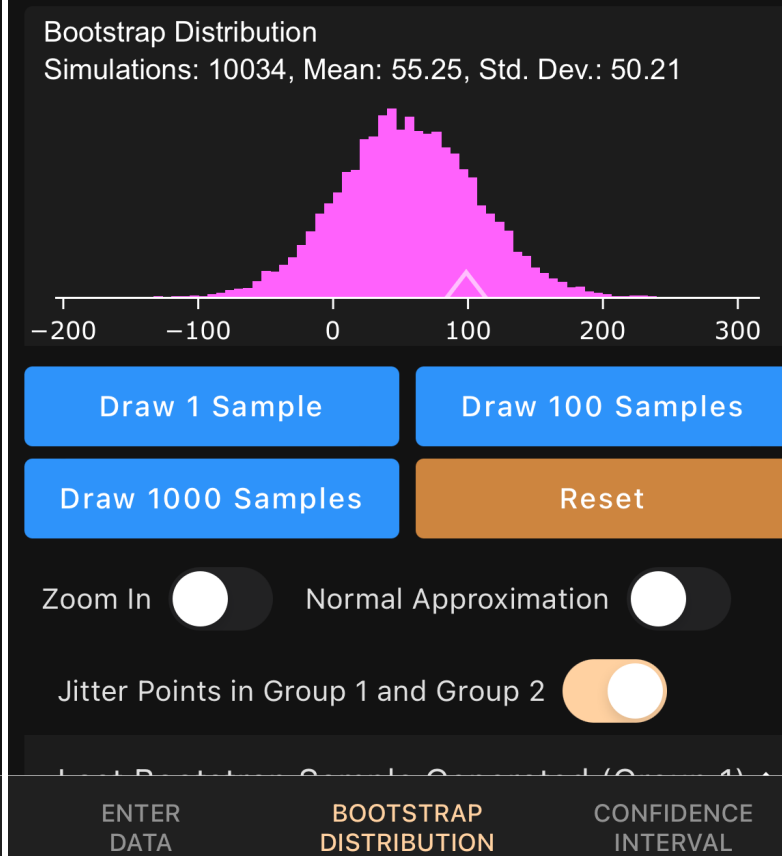
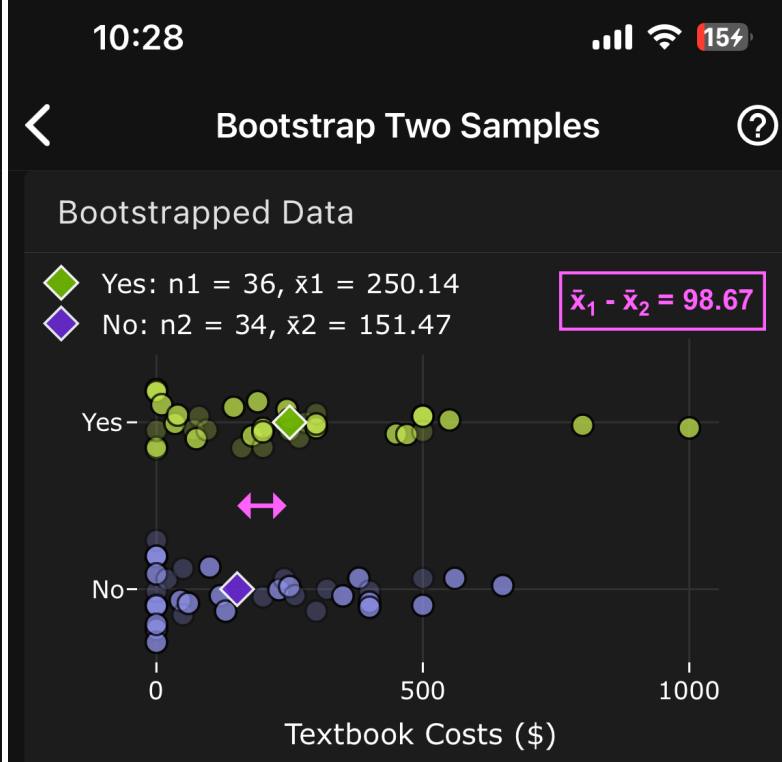
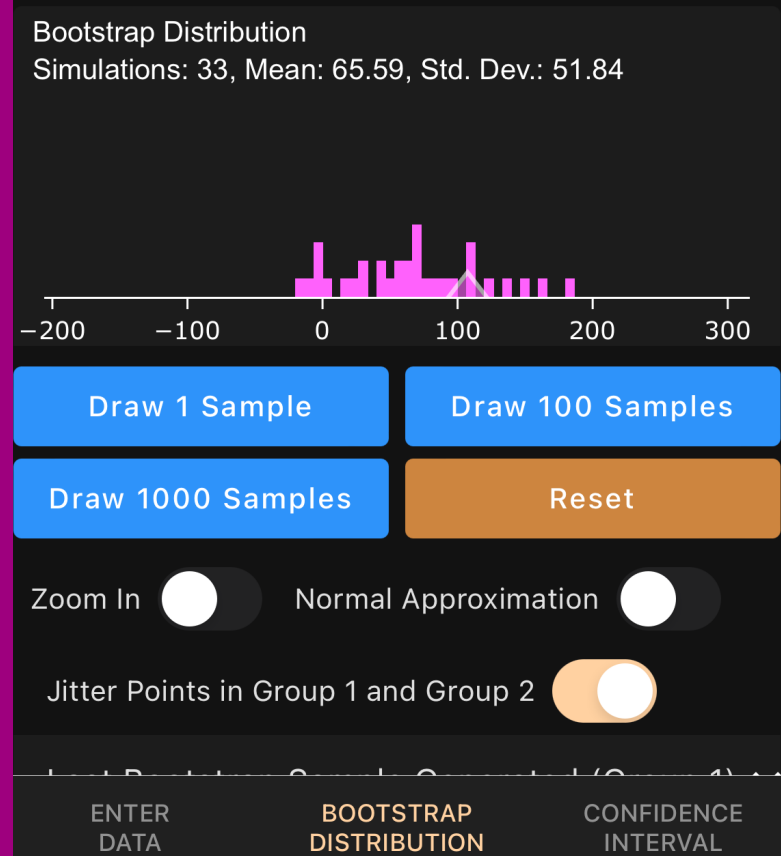
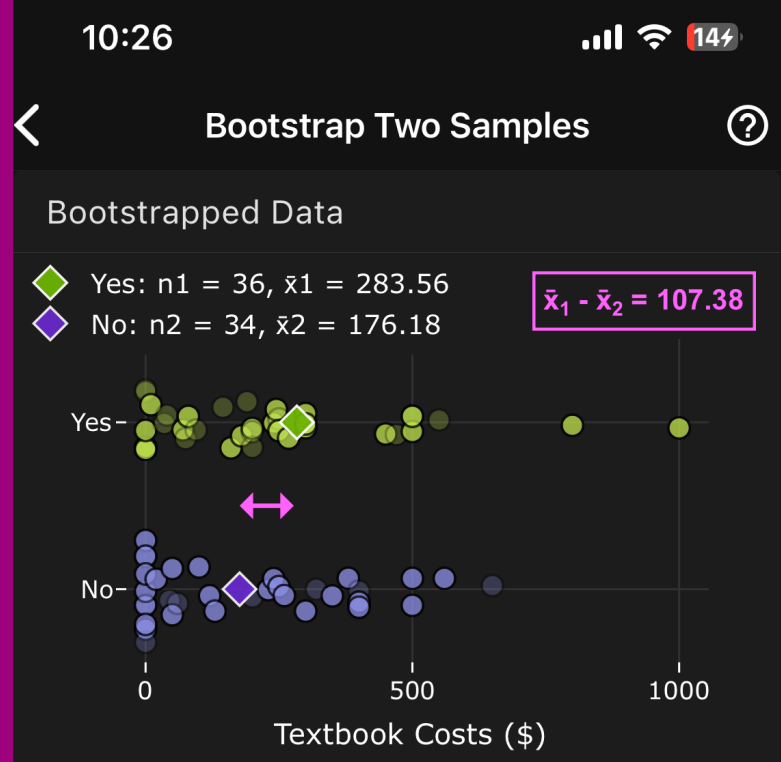
CONFIDENCE INTERVAL

Art of Stat Resampling

Bootstrap Interval for the Difference Between Two Population Means

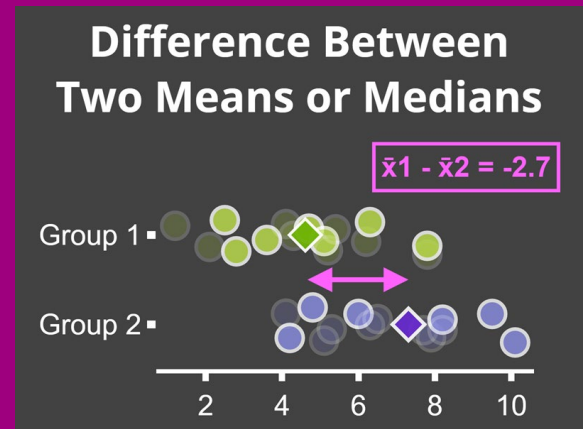


- Screens:
- Enter Data
 - **Bootstrap Distrib.**
 - Confidence Interval



Art of Stat Resampling

Bootstrap Interval for the Difference Between Two Population Means



- Screens:
- Enter Data
 - Bootstrap Distrib.
 - Confidence Interval

10:29

Bootstrap Two Samples

Distribution of the Difference in Means

95% Confidence Interval: (-43.4, 155.1)

Observed Difference Smooth Distb.

Summary Statistics Normal Distb.

Inference for the Difference in Population Means

Statistic	Bootstrap Percentile Interval
Difference of Sample Means	55.13
Standard Error	50.21
95% Lower Bound	-43.39
95% Upper Bound	155.06

ENTER DATA | BOOTSTRAP DISTRIBUTION | **CONFIDENCE INTERVAL**

10:29

Bootstrap Two Samples

Distribution of the Difference in Means

95% Confidence Interval: (-43.4, 155.1)

Observed Difference Smooth Distb.

Summary Statistics Normal Distb.

Inference for the Difference in Population Means

Statistic	Bootstrap Percentile Interval
Difference of Sample Means	55.13
Standard Error	50.21
95% Lower Bound	-43.39
95% Upper Bound	155.06

ENTER DATA | BOOTSTRAP DISTRIBUTION | CONFIDENCE INTERVAL

10:30

Bootstrap Two Samples

Difference in Population Means $\mu_1 - \mu_2$

Observed Difference Smooth Distb.

Summary Statistics Normal Distb.

Inference for the Difference in Population Means

Statistic	Bootstrap Percentile Interval	Student t Interval
Difference of Sample Means	55.13	55.13
Standard Error	50.21*	50.91**
95% Lower Bound	-43.39	-46.48
95% Upper Bound	155.06	156.74

* Standard error based on bootstrap distribution.
** Standard error based on formula.

Confidence Level: 95 Interval Type: Two-Sided

Show Student t-Interval

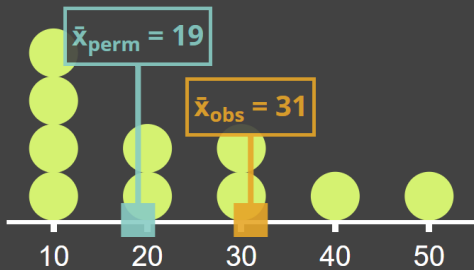
Method of Constructing Bootstrap Interval: Bootstrap Percentile Method

ENTER DATA | BOOTSTRAP DISTRIBUTION | CONFIDENCE INTERVAL

Art of Stat Resampling

Permutation Test for a Population Mean

Permutation Test Mean & Median



12:01

Permutation One Sample

Original Sample: $n = 14, \bar{x} = 3.36, s = 1.65$

$\bar{x}_{orig} = 3.4$

Permuted Sample assuming $H_0: \mu = 4$ is true

Press a Button Below To Generate Samples under H_0

Null Distribution of the Mean
Permutations: 0

1 Permutation | 100 Permutations

1000 Permutations | Reset

Null Value | Observed Test Statistic

Zoom In | Normal Sampling Distb.

ENTER DATA | **PERMUTATION DISTRIBUTION** | PERMUTATION TEST

12:01

Permutation One Sample

Original Sample: $n = 14, \bar{x} = 3.36, s = 1.65$

$\bar{x}_{orig} = 3.4$

Permuted Sample assuming $H_0: \mu = 4$ is true
 $n = 14, \bar{x} = 4.64, s = 1.65$

$\bar{x}_{perm} = 4.6$

Null Distribution of the Mean
Permutations: 1, Mean: 4.6, Std. Dev.: -

1 Permutation | 100 Permutations

1000 Permutations | Reset

Null Value | Observed Test Statistic

Zoom In | Normal Sampling Distb.

ENTER DATA | PERMUTATION DISTRIBUTION | PERMUTATION TEST

12:02

Permutation One Sample

Original Sample: $n = 14, \bar{x} = 3.36, s = 1.65$

$\bar{x}_{orig} = 3.4$

Permuted Sample assuming $H_0: \mu = 4$ is true
 $n = 14, \bar{x} = 4.36, s = 1.74$

$\bar{x}_{perm} = 4.4$

Null Distribution of the Mean
Permutations: 10001, Mean: 4.0, Std. Dev.: 0.5

1 Permutation | 100 Permutations

1000 Permutations | Reset

Null Value | **Observed Test Statistic**

Zoom In | Normal Sampling Distb.

ENTER DATA | PERMUTATION DISTRIBUTION | PERMUTATION TEST

12

Screens:

- Enter Data

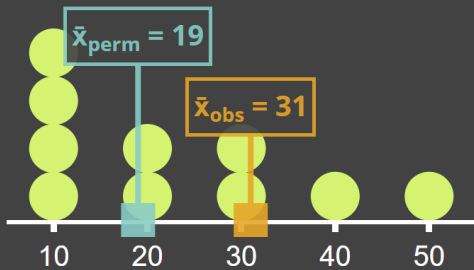
- **Permutation Distrib.**

- Permutation Test

Art of Stat Resampling

Permutation Test for a Population Mean

Permutation Test Mean & Median



12:02

Permutation One Sample

Distribution of the Test Statistic under H0

Observed Statistic Smooth Distb.

Summary Statistics Normal Distb.

Show Student t-Test

Hypothesis Test for the Population Mean

Null Hypothesis	$H_0: \mu = 4$
Alternative Hypothesis	$H_a: \mu \neq 4$
Observed Test Statistic	$\bar{x} = 3.36$
Number of Permutations	10001
Permutations with Test Statistic As or More Extreme than Observed	2144
Permutation P-value	0.2144

Alternative Hypothesis: Two-Sided

ENTER DATA PERMUTATION DISTRIBUTION **PERMUTATION TEST**

12:03

Permutation One Sample

Summary Statistics Normal Distb.

Show Student t-Test

Summary of Permutation Distribution

Statistic	Value
Number of Permutations	10001
Unique Values	22
Mean of Distribution	4.00
Standard Deviation	0.46
Minimum	2.50
First Quartile (25%)	3.64
Median	4.07
Third Quartile (75%)	4.36
Maximum	5.50
Permutations with Value of Test Statistic ≤ 3.36	1104 (11.04%)
Permutations with Value of Test Statistic ≥ 4.64	1040 (10.40%)
Permutations with Value of Test Statistic as or More Extreme than Observed	2144 (21.44%)

Hypothesis Test for the Population Mean

ENTER DATA PERMUTATION DISTRIBUTION **PERMUTATION TEST**

12:04

Permutation One Sample

Show Student t-Test

Hypothesis Test for the Population Mean

Null Hypothesis	$H_0: \mu = 4$
Alternative Hypothesis	$H_a: \mu \neq 4$
Observed Test Statistic	$\bar{x} = 3.36$
Number of Permutations	10001
Permutations with Test Statistic As or More Extreme than Observed	2144
Permutation P-value	0.2144
t-Test Statistic	-1.46
Degrees of Freedom (df)	13
P-value from t-Distribution (df = 13)	0.1676

Alternative Hypothesis: Two-Sided

Generated Means (10001 Permutations, Sorted)

2.5	2.5	2.6	2.6	2.6
2.6	2.6	2.6	2.6	2.8

ENTER DATA PERMUTATION DISTRIBUTION **PERMUTATION TEST**

Screens:

- Enter Data
- Permutation Distrib.
- **Permutation Test**

Thank You!

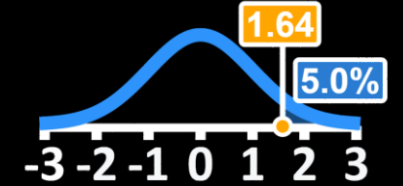
Any Questions?

Slides available at
ArtofStat.com

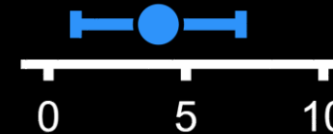
Art of Stat
Explore Data



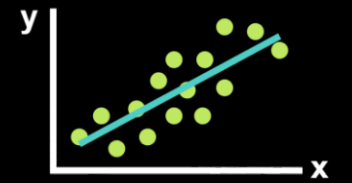
Art of Stat
Distributions



Art of Stat
Inference



Art of Stat
Regression



Art of Stat
Concepts



Art of Stat
Resampling

