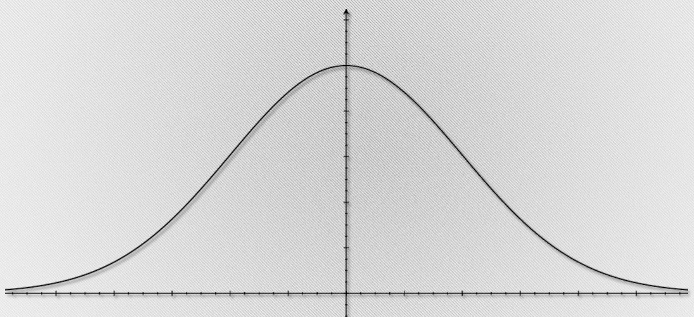


Distribution of Sampling Error

The Distribution of Sample Means

The Logic of Hypothesis Testing

The Distribution of Sample Means

The Logic of Hypothesis Testing

Single Group Analyses	**Single Group Experiments**
T-test, Z-test	Single Sample Design
T-test	Pre-Post Designs
Mixed-effects models	Repeated Measures Designs
Multi-Group Analyses	**Multi-Group Experiments**
T-test	Two Group Design
Analysis of Variance	Multi-Group Design
Regression Analysis	Continuous Predictor Design
Mixed Analyses	**Mixed Experiments**
Mixed-effects models	Group Design with Repeated Measurements
Analysis of Covariance	Group Design also measured on a continuous variable

Single Group Experiments

Single Sample Design

Population

IQ of All Individuals in the USA

$\mu = 100$

$\sigma = 15$

Population

IQ of All Individuals in the USA

$\mu = 100$

$\sigma = 15$

Treatment

Population

IQ of All Individuals in the USA

$\mu = 100$

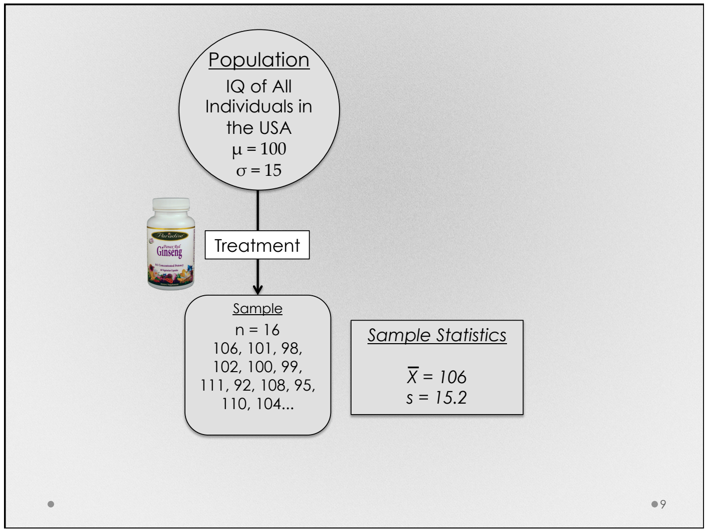
$\sigma = 15$

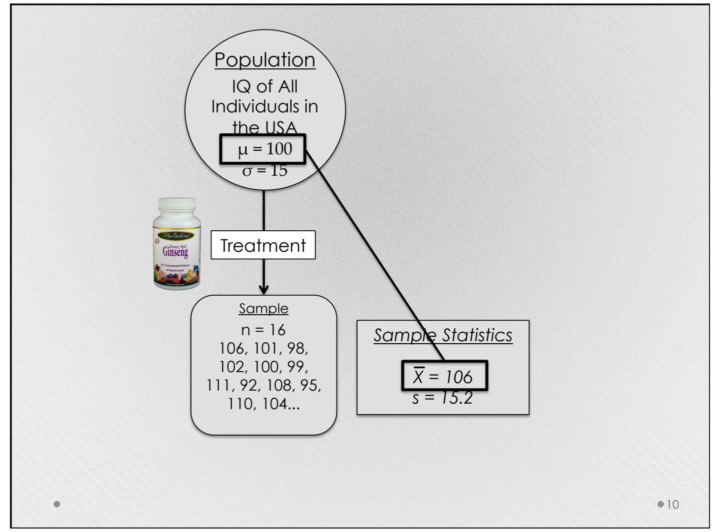
Treatment

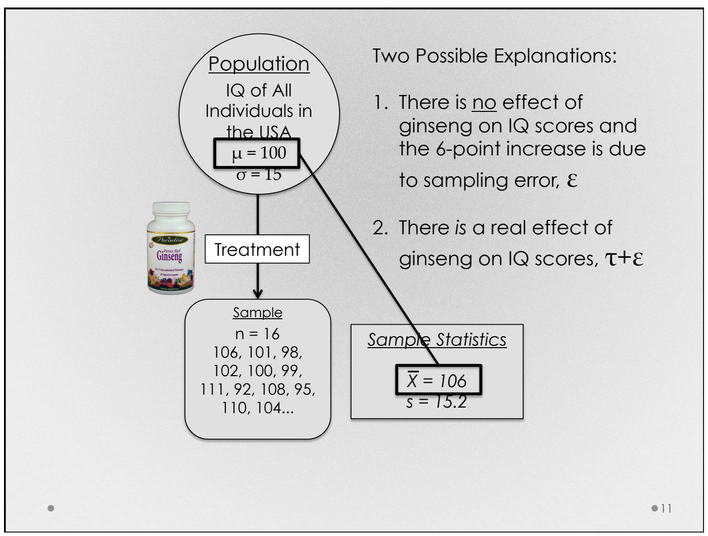
Sample

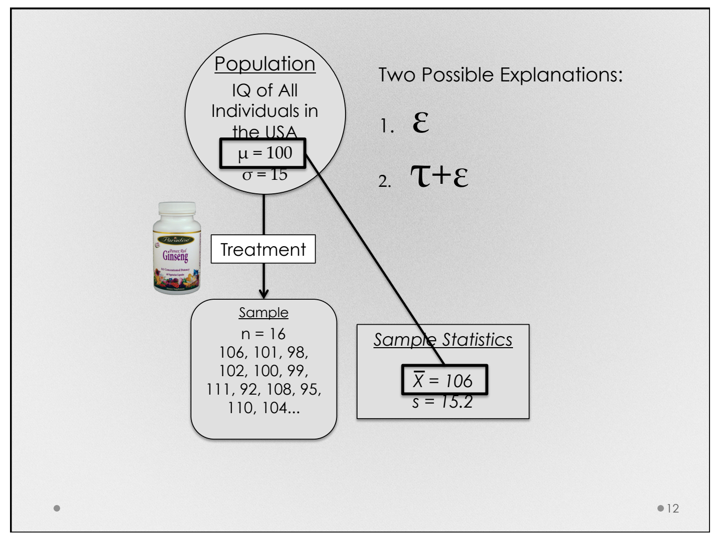
n = 16

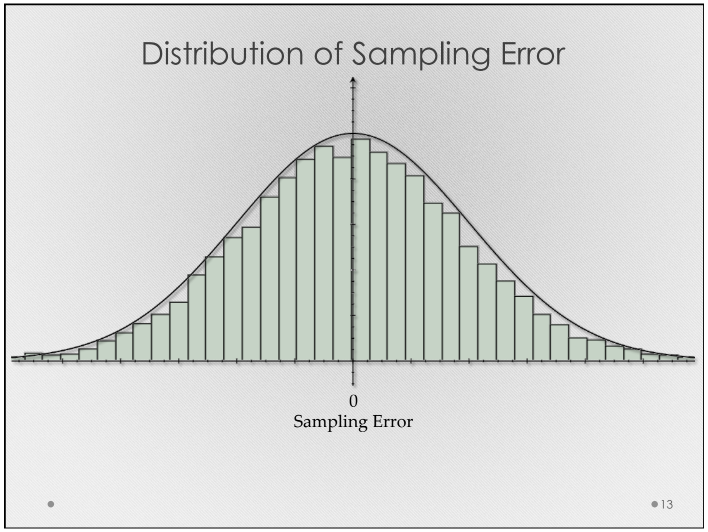
106, 101, 98, 102, 100, 99, 111, 92, 108, 95, 110, 104...

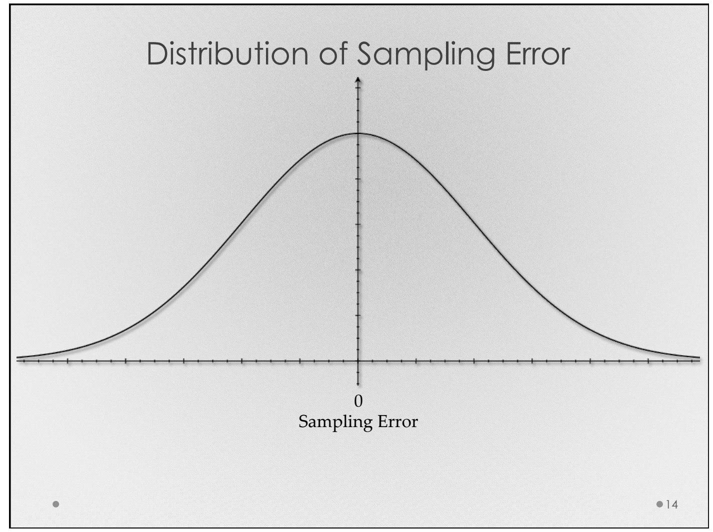

Population
IQ of All Individuals in the USA
μ = 100
σ = 15
Ginseng
Treatment
Sample
n = 16
106, 101, 98, 102, 100, 99, 111, 92, 108, 95, 110, 104...
Sample Statistics
X̄ = 106
s = 15.2
9

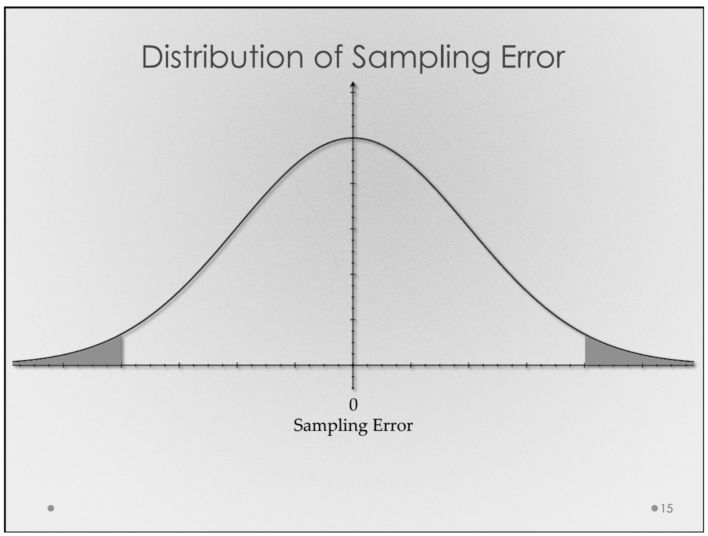

Population
IQ of All Individuals in the USA
μ = 100
σ = 15
Ginseng
Treatment
Sample
n = 16
106, 101, 98, 102, 100, 99, 111, 92, 108, 95, 110, 104...
Sample Statistics
X̄ = 106
s = 15.2
10

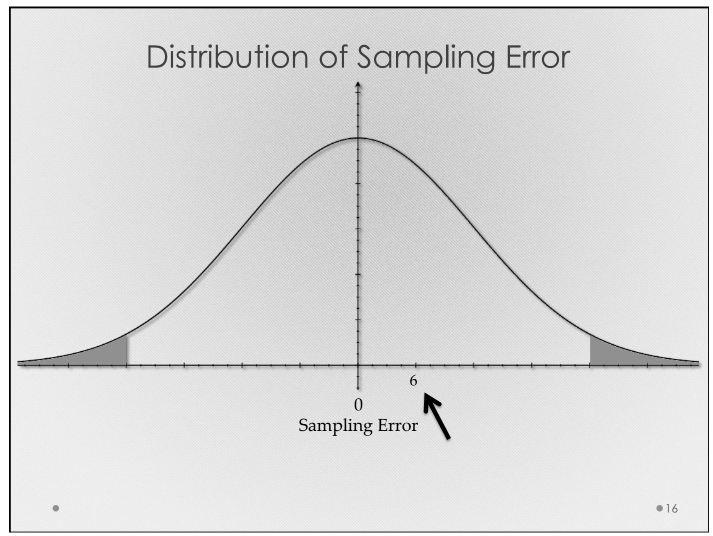

Population
IQ of All Individuals in the USA
μ = 100
σ = 15
Ginseng
Treatment
Sample
n = 16
106, 101, 98, 102, 100, 99, 111, 92, 108, 95, 110, 104...
Sample Statistics
X̄ = 106
s = 15.2
Two Possible Explanations:
1. There is no effect of ginseng on IQ scores and the 6-point increase is due to sampling error, ε
2. There is a real effect of ginseng on IQ scores, τ+ε
11


Population
IQ of All Individuals in the USA
μ = 100
σ = 15
Ginseng
Treatment
Sample
n = 16
106, 101, 98, 102, 100, 99, 111, 92, 108, 95, 110, 104...
Sample Statistics
X̄ = 106
s = 15.2
Two Possible Explanations:
1. ε
2. τ+ε
12

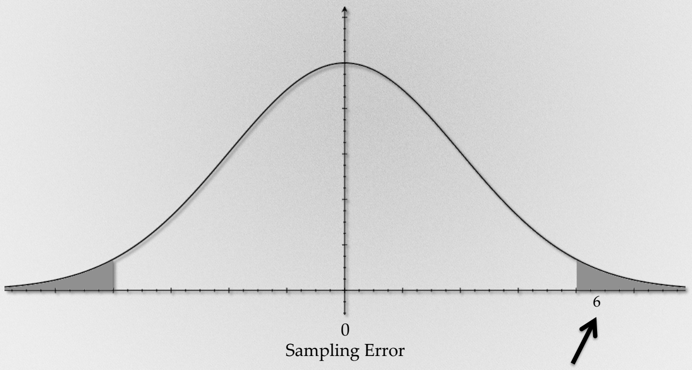

Distribution of Sampling Error
0
Sampling Error
13


Distribution of Sampling Error
0
Sampling Error
14


Distribution of Sampling Error
0
Sampling Error
15

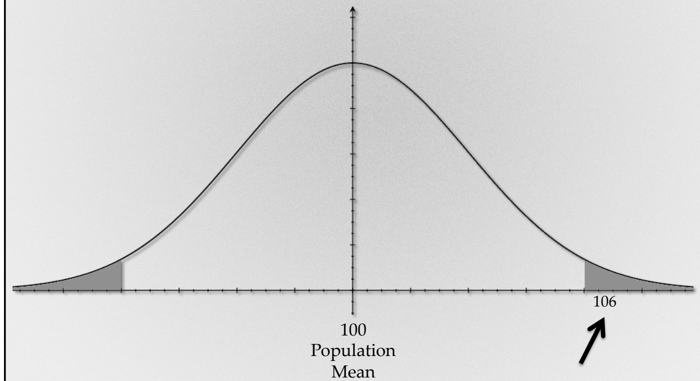

Distribution of Sampling Error
6
0
Sampling Error
16

Distribution of Sampling Error



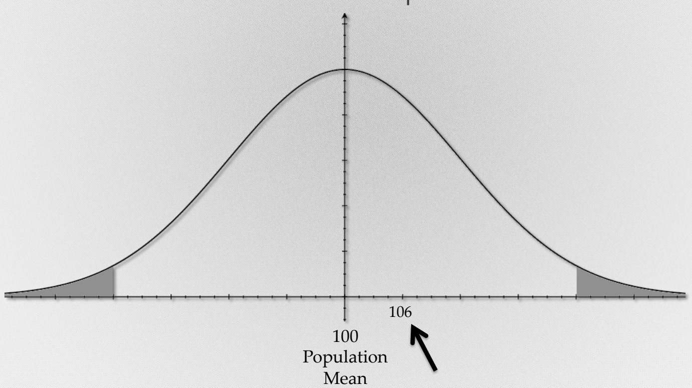
17

Distribution of Sample Means



18

Distribution of Sample Means



19

The Distribution of Sample Means

The collection of sample means
for all the possible random
samples of a particular size (n)

20

Sampling Distribution

A distribution of statistics obtained by computing a statistic for every possible random sample of size n from a population

The Distribution of Sample Means

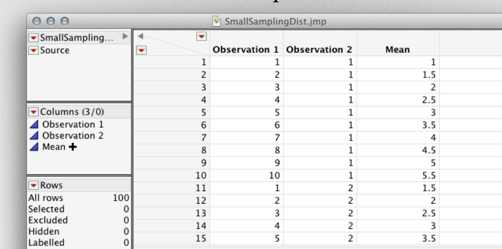
Sampling Distribution of $\bar{X}$

Constructing a *real* Sampling Distribution

1. Choose a population
2. Enumerate *every possible* sample of size n
3. Calculate the mean for *each* sample of size n
4. Plot of histogram of the *means* calculated

Full Sampling Distribution for Small Population

Sampling Distribution for
Small Population



	Observation 1	Observation 2	Mean
1	1	1	1
2	2	1	1.5
3	3	1	2
4	4	1	2.5
5	5	1	3
6	6	1	3.5
7	7	1	4
8	8	1	4.5
9	9	1	5
10	10	1	5.5
11	1	2	1.5
12	2	2	2
13	3	2	2.5
14	4	2	3
15	5	2	3.5

Approximating a Sampling Distribution

1. Choose a population
2. <u>Take a random sample of size n</u>
3. Calculate the mean for that *particular* sample
4. Record the mean to a table
5. <u>Go back to #2 and repeat 100,000 times</u>

Normal, n = 2

Sample Means Samples of Size 2 from a Normal Population

normal n of 2.jmp

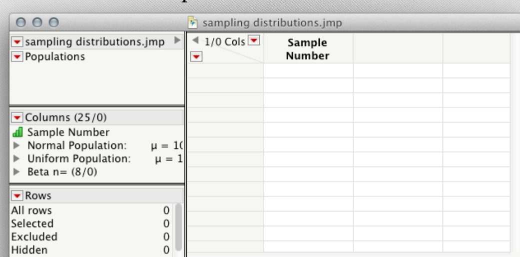
Columns (3/0)
Observation 1
Observation 2
Mean

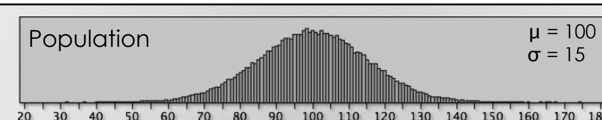
Rows
All rows 100,000
Selected 0
Excluded 0
Hidden 0
Labelled 0

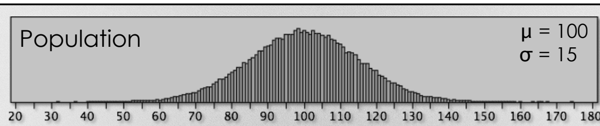
	Observation 1	Observation 2	Mean
1	103	79	90.94
2	126	103	114.37
3	86	93	89.28
4	123	115	119.03
5	125	110	117.96
6	89	90	89.56
7	102	119	110.88
8	89	98	93.52
9	135	96	115.29
10	87	122	104.43
11	121	94	107.67
12	112	105	108.49
13	70	87	78.32
14	111	115	112.86
15	99	129	113.52

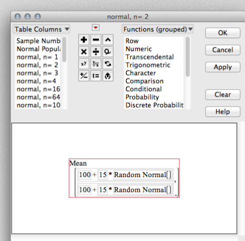
Sampling Distributions, 3 Populations

Sample Means for Random Samples of Different Sizes

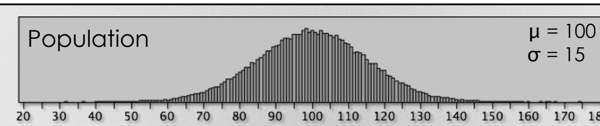


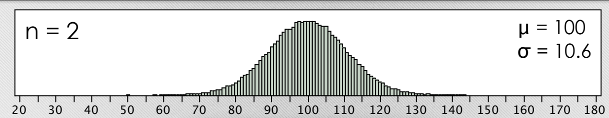


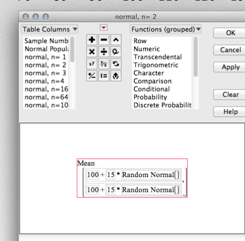




29







30

FROM BEFORE

Normal, n = 2

Sample Means Samples of Size 2
from a Normal Population

normal n of 2.jmp

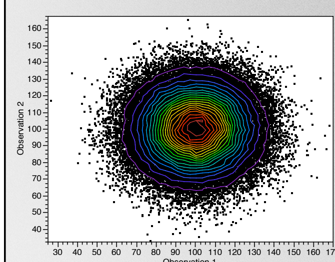
Columns (3/0): Observation 1, Observation 2, Mean

Rows	
All rows	100,000
Selected	0
Excluded	0
Hidden	0
Labelled	0

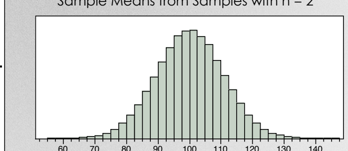
	Observation 1	Observation 2	Mean
1	103	79	90.94
2	126	103	114.37
3	86	93	89.28
4	123	115	119.03
5	125	110	117.96
6	89	90	89.56
7	102	119	110.88
8	89	98	93.52
9	135	96	115.29
10	87	122	104.43
11	121	94	107.67
12	112	105	108.49
13	70	87	78.32
14	111	115	112.86
15	99	129	113.52

31

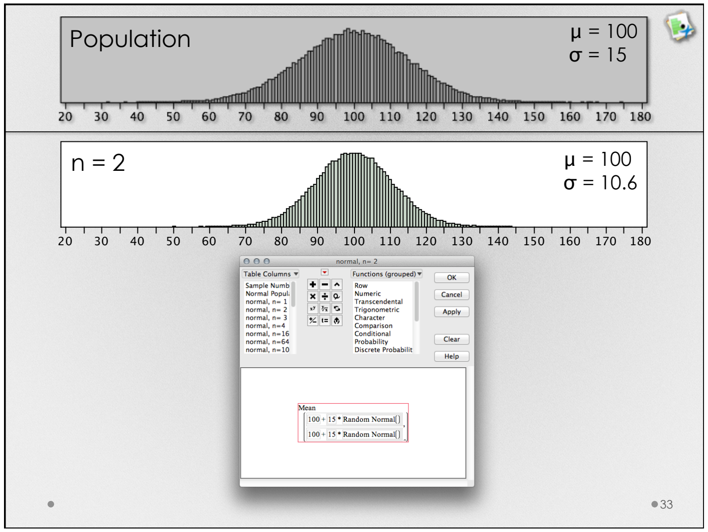
FROM BEFORE

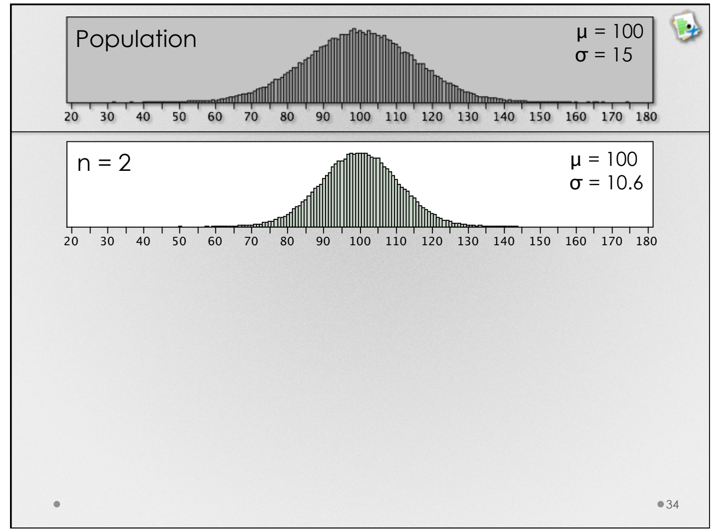


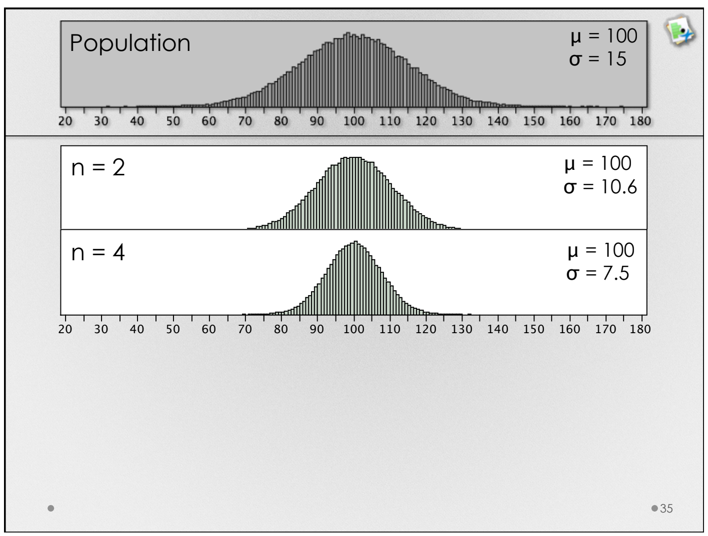
Sample Means from Samples with n = 2

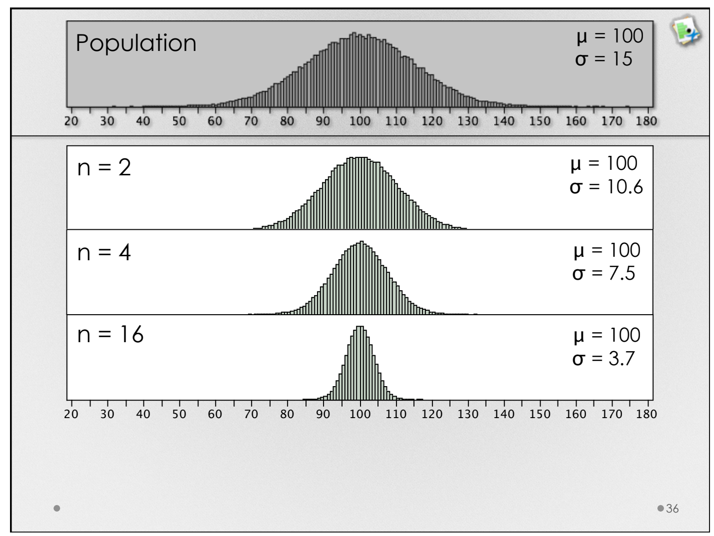


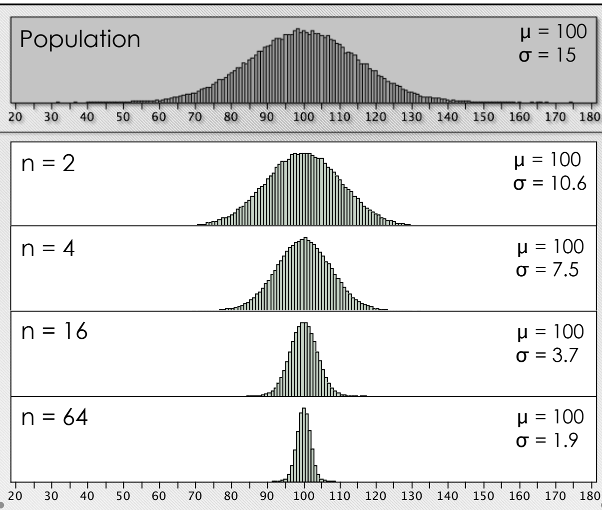
32


Population
μ = 100
σ = 15
20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180
n = 2
μ = 100
σ = 10.6
20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180
33

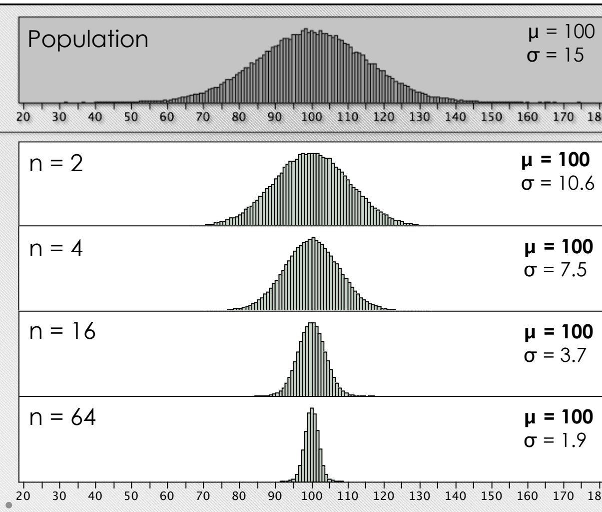

Population
μ = 100
σ = 15
20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180
n = 2
μ = 100
σ = 10.6
20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180
34


Population
μ = 100
σ = 15
20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180
n = 2
μ = 100
σ = 10.6
n = 4
μ = 100
σ = 7.5
20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180
35


Population
μ = 100
σ = 15
20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180
n = 2
μ = 100
σ = 10.6
n = 4
μ = 100
σ = 7.5
n = 16
μ = 100
σ = 3.7
20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180
36



37



38

Characteristics of the Sampling Distribution of $\bar{X}$

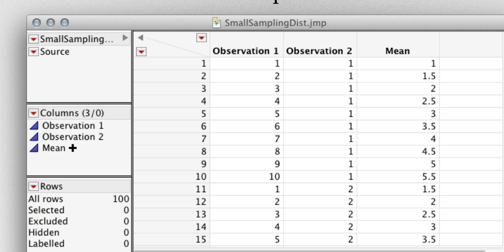
- Sampling distribution always has a mean = μ

39

FROM BEFORE

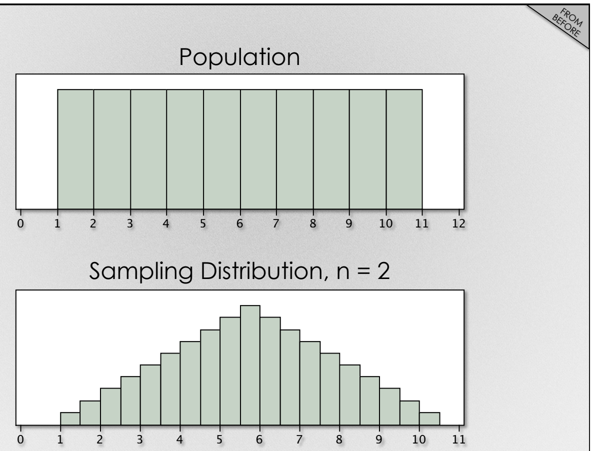
Full Sampling Distribution for Small Population

Sampling Distribution for Small Population



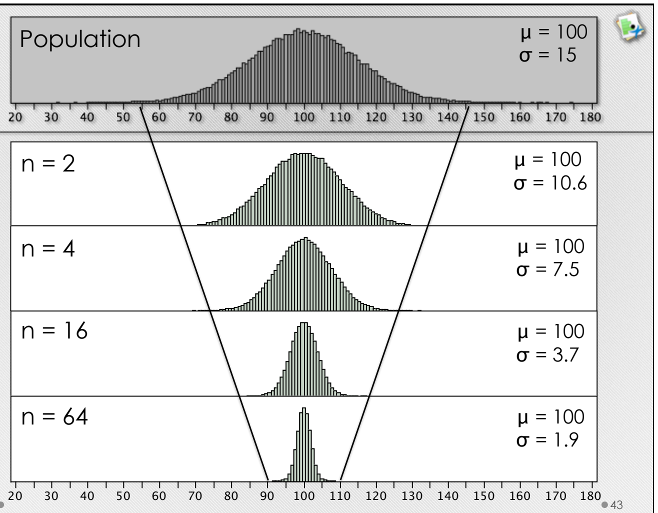
	Observation 1	Observation 2	Mean
1	1	1	1
2	2	1	1.5
3	3	1	2
4	4	1	2.5
5	5	1	3
6	6	1	3.5
7	7	1	4
8	8	1	4.5
9	9	1	5
10	10	1	5.5
11	1	2	1.5
12	2	2	2
13	3	2	2.5
14	4	2	3
15	5	2	3.5

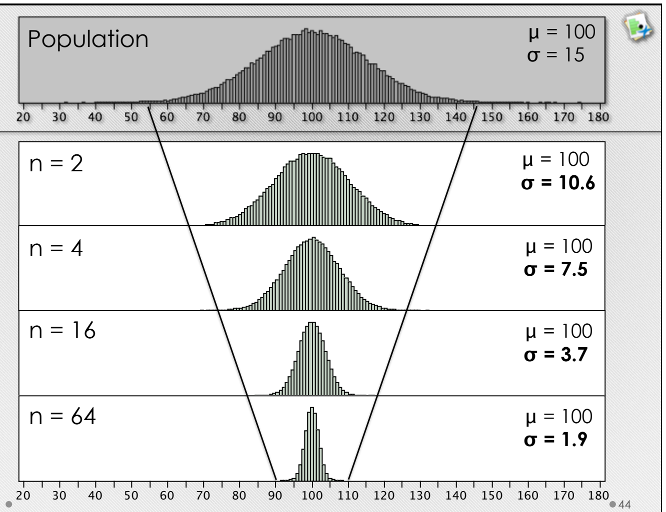
40



Characteristics of the Sampling Distribution of $\overline{X}$

- Sampling distribution <u>always</u> has a mean = μ





Characteristics of the Sampling Distribution of $\overline{X}$

- Sampling distribution always has a mean = μ
- As sample size (n) increases, the standard deviation of the sampling distribution decreases by the square root of the sample size (n)

$$\frac{\sigma}{\sqrt{n}}$$

45

Characteristics of the Sampling Distribution of $\overline{X}$

- Sampling distribution always has a mean = μ
- As sample size (n) increases, the standard deviation of the sampling distribution decreases by the square root of the sample size (n)

$$\frac{\sigma}{\sqrt{n}}$$

46

Standard Deviation of the Sampling Distribution of Sample Means

47

Standard (Sampling) Error

Standard Deviation of a Sampling Distribution of Sample Means

48

Standard (Sampling) Error

$$\sigma_{\overline{X}}$$

49

Standard (Sampling) Error

The Standard Deviation of a Sampling Distribution of Sample Means

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \sqrt{\frac{\sigma^2}{n}}$$

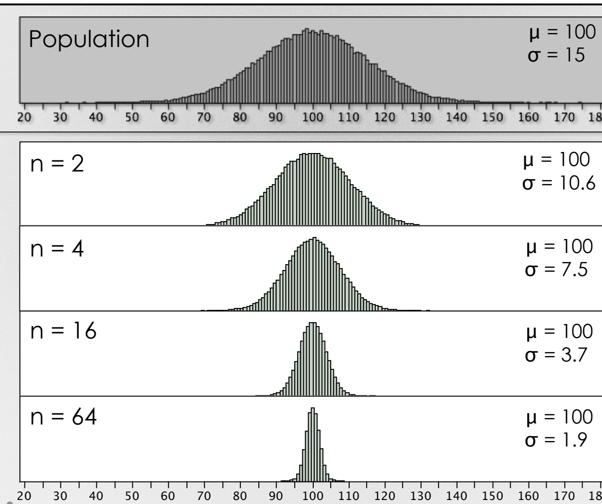
50

Characteristics of the Sampling Distribution of $\overline{X}$

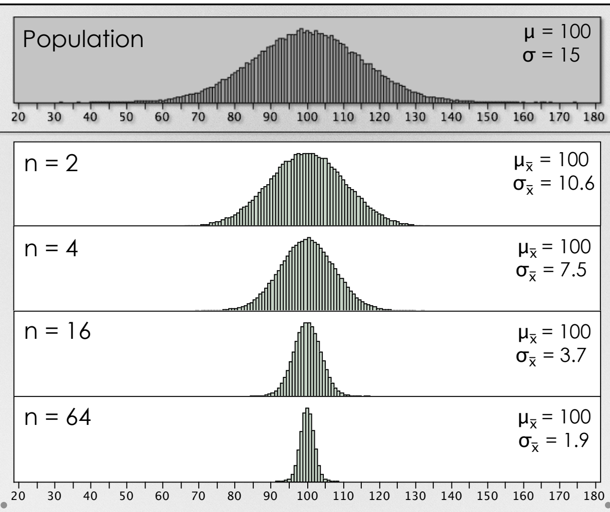
- Sampling distribution always has a mean = μ
- As sample size (n) increases, the standard deviation of the sampling distribution decreases by the square root of the sample size (n)

$$\frac{\sigma}{\sqrt{n}}$$

51



52

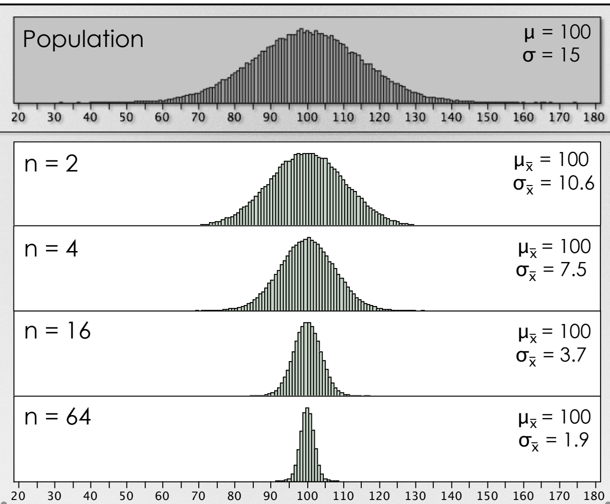


53

Characteristics of the Sampling Distribution of $\overline{X}$

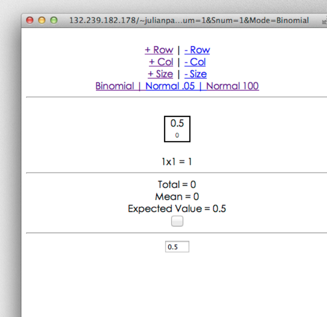
- Sampling distribution always has a mean = μ
- As sample size (n) increases, the standard deviation of the sampling distribution decreases by the square root of the sample size (n) $\dfrac{\sigma}{\sqrt{n}}$
- The sampling distribution is *normally* distributed
  …under what conditions?

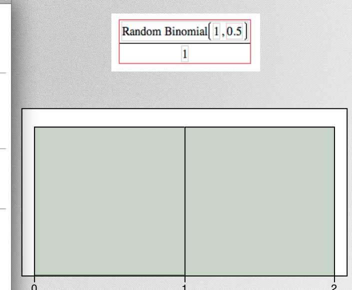
54



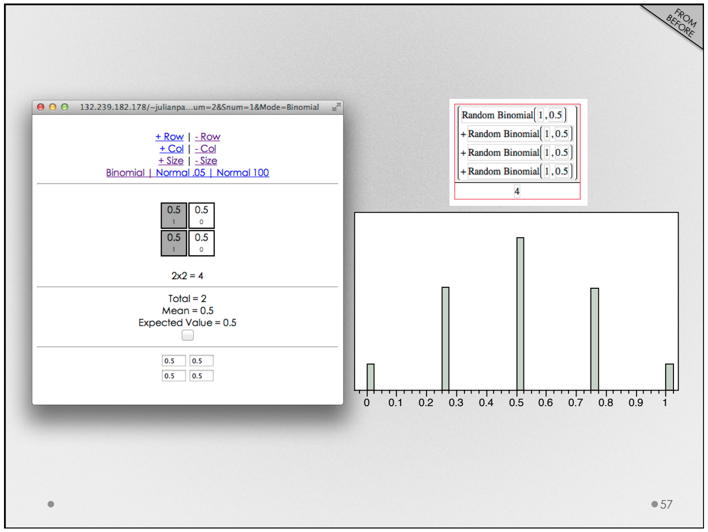
55

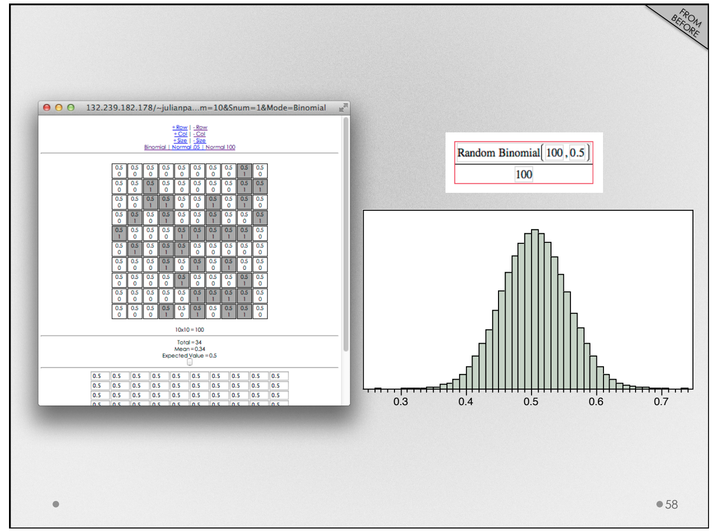
FROM BEFORE

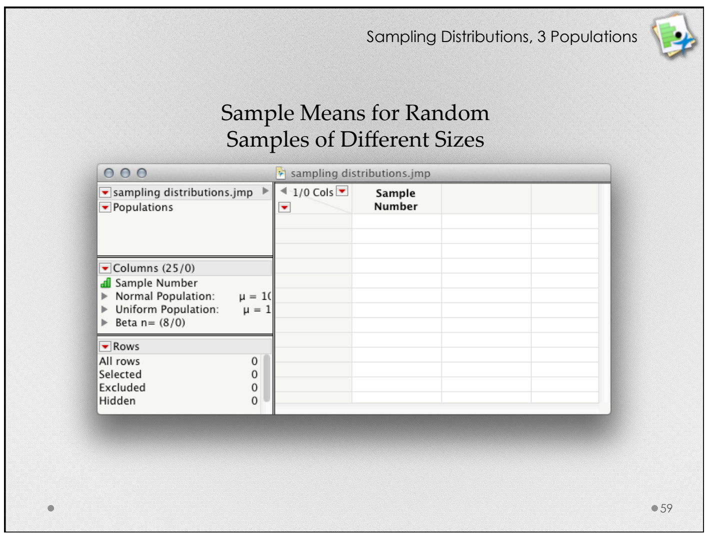


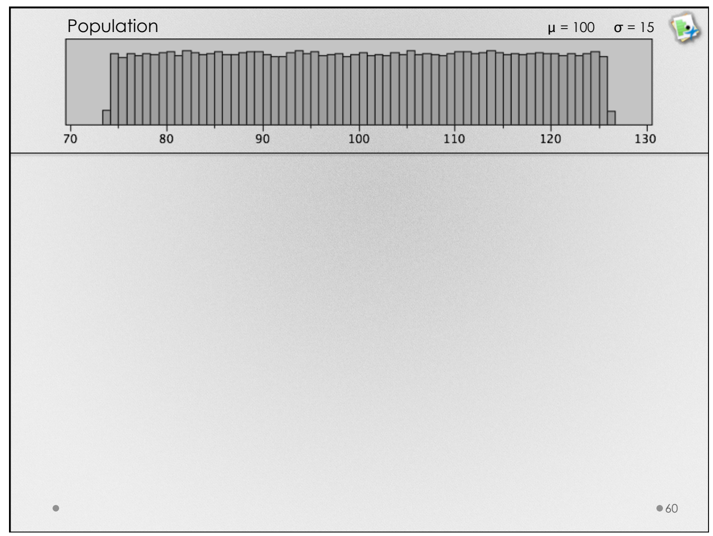


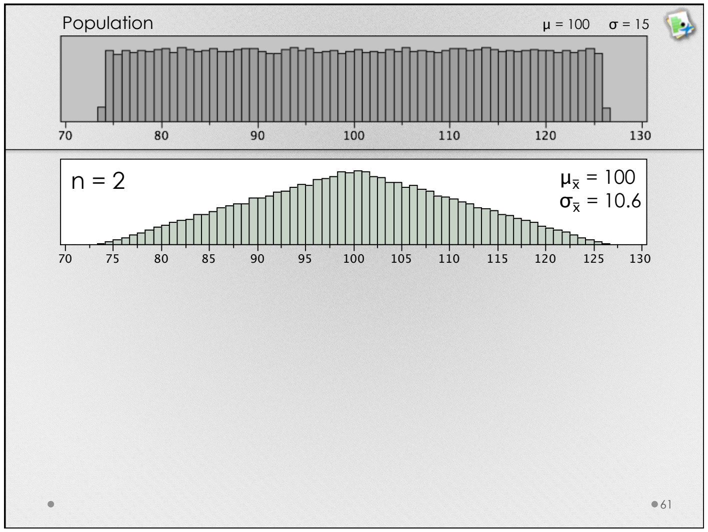
56

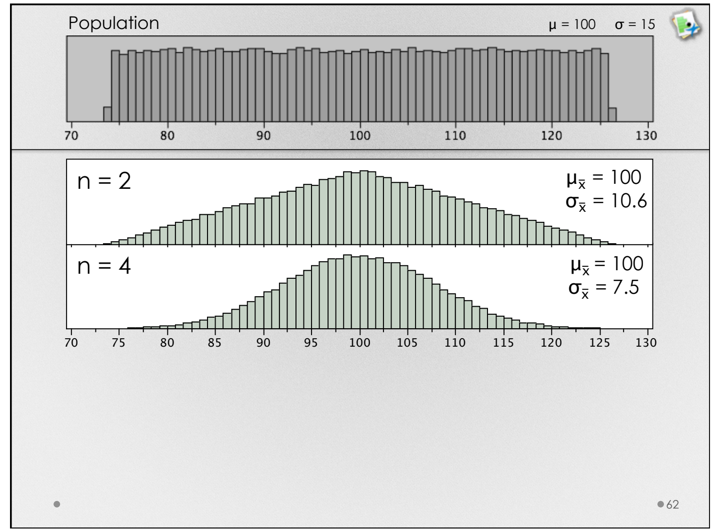

FROM BEFORE
+ Row | - Row
+ Col | - Col
+ Size | - Size
Binomial | Normal .05 | Normal 100
2x2 = 4
Total = 2
Mean = 0.5
Expected Value = 0.5
Random Binomial(1 , 0.5)
+ Random Binomial(1 , 0.5)
+ Random Binomial(1 , 0.5)
+ Random Binomial(1 , 0.5)
4
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
57

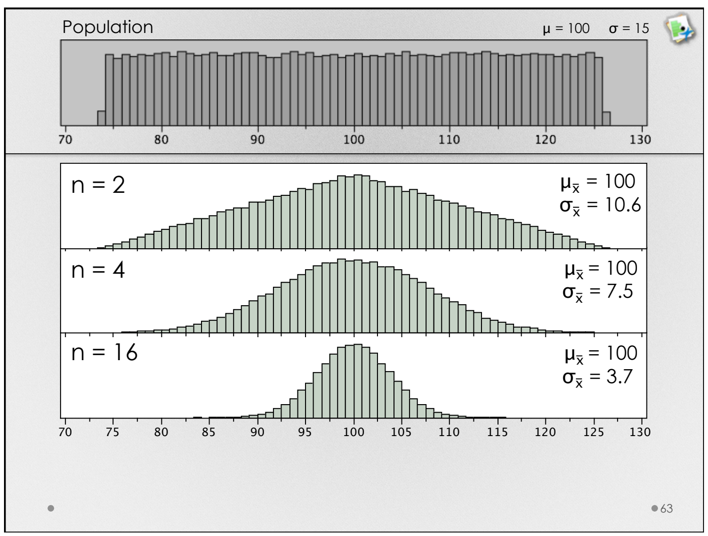

FROM BEFORE
Random Binomial(100 , 0.5)
100
0.3 0.4 0.5 0.6 0.7
58

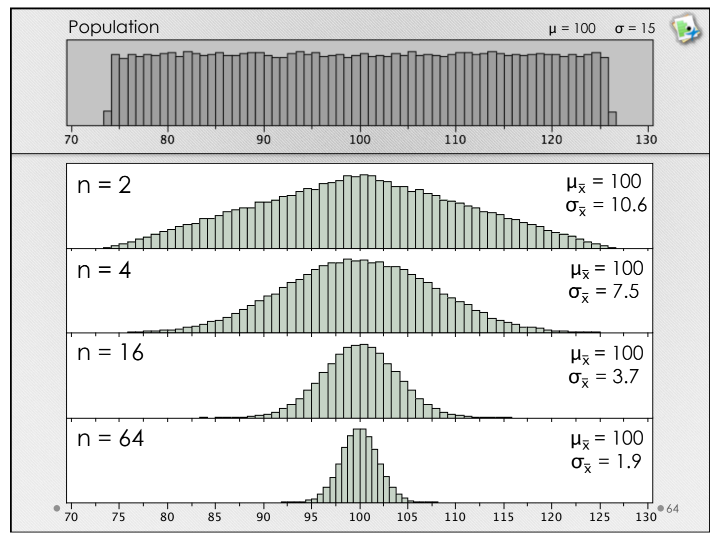

Sampling Distributions, 3 Populations
Sample Means for Random
Samples of Different Sizes
sampling distributions.jmp
Populations
Columns (25/0)
Sample Number
Normal Population:
Uniform Population:
Beta n= (8/0)
Rows
All rows 0
Selected 0
Excluded 0
Hidden 0
1/0 Cols
Sample Number
59

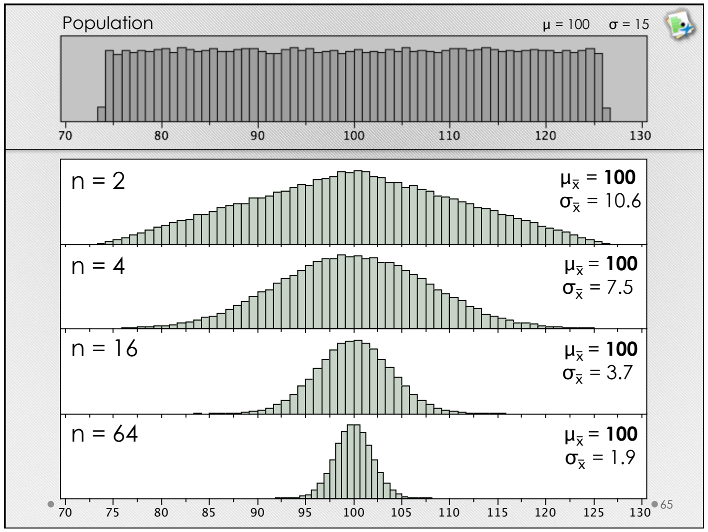

Population
μ = 100 σ = 15
70 80 90 100 110 120 130
60

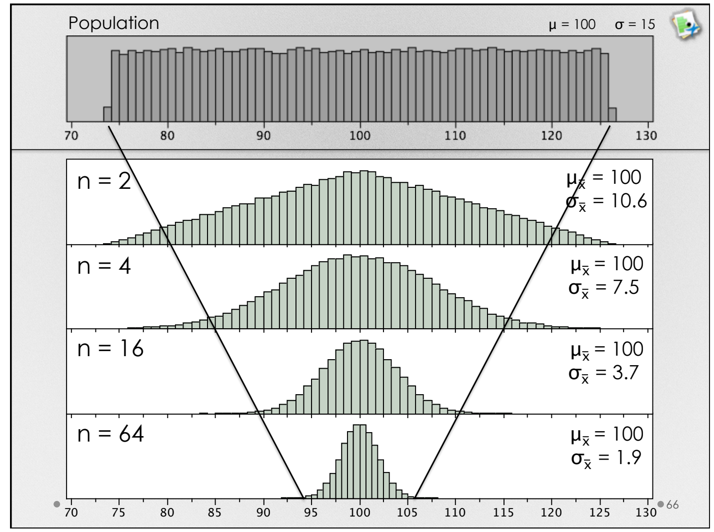

Population
μ = 100 σ = 15
70 80 90 100 110 120 130
n = 2
$\mu_{\bar{x}}$ = 100
$\sigma_{\bar{x}}$ = 10.6
70 75 80 85 90 95 100 105 110 115 120 125 130
61

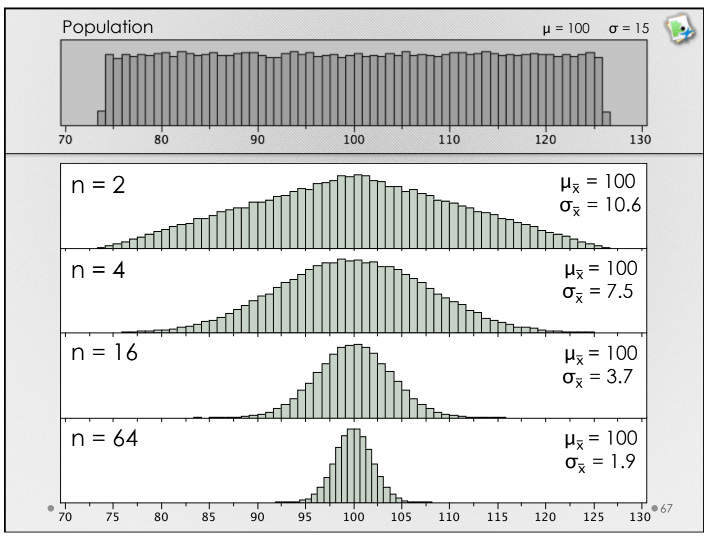

Population
μ = 100 σ = 15
70 80 90 100 110 120 130
n = 2
$\mu_{\bar{x}}$ = 100
$\sigma_{\bar{x}}$ = 10.6
n = 4
$\mu_{\bar{x}}$ = 100
$\sigma_{\bar{x}}$ = 7.5
70 75 80 85 90 95 100 105 110 115 120 125 130
62

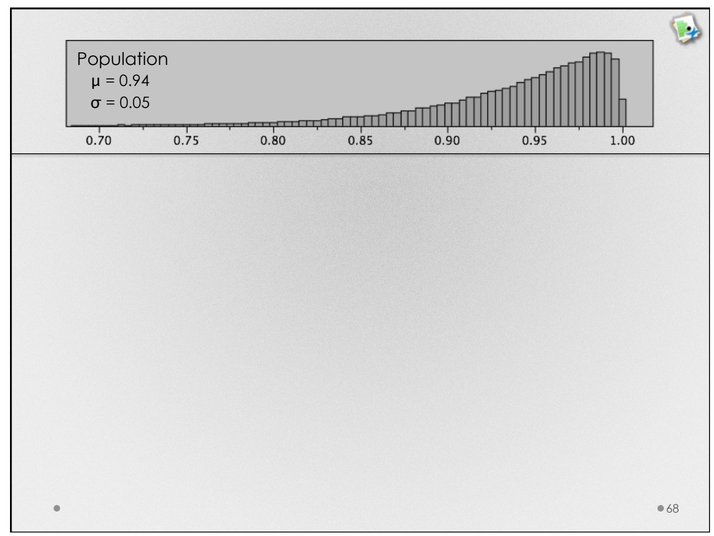

Population
μ = 100 σ = 15
70 80 90 100 110 120 130
n = 2
$\mu_{\bar{x}}$ = 100
$\sigma_{\bar{x}}$ = 10.6
n = 4
$\mu_{\bar{x}}$ = 100
$\sigma_{\bar{x}}$ = 7.5
n = 16
$\mu_{\bar{x}}$ = 100
$\sigma_{\bar{x}}$ = 3.7
70 75 80 85 90 95 100 105 110 115 120 125 130
63

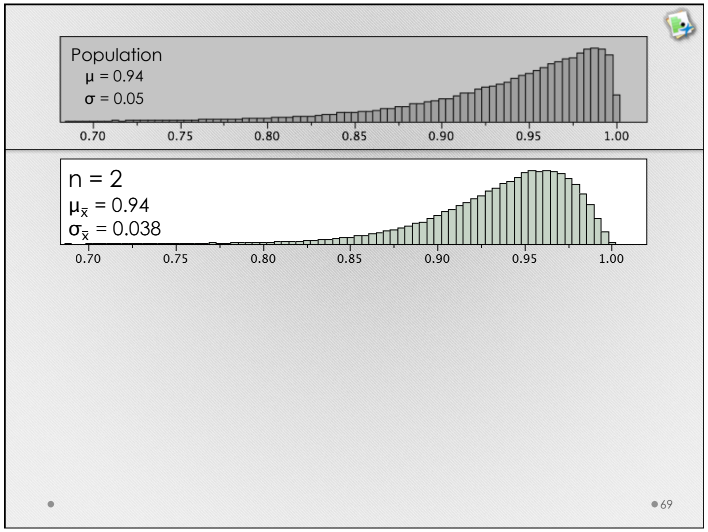

Population
μ = 100 σ = 15
70 80 90 100 110 120 130
n = 2
$\mu_{\bar{x}}$ = 100
$\sigma_{\bar{x}}$ = 10.6
n = 4
$\mu_{\bar{x}}$ = 100
$\sigma_{\bar{x}}$ = 7.5
n = 16
$\mu_{\bar{x}}$ = 100
$\sigma_{\bar{x}}$ = 3.7
n = 64
$\mu_{\bar{x}}$ = 100
$\sigma_{\bar{x}}$ = 1.9
70 75 80 85 90 95 100 105 110 115 120 125 130
64

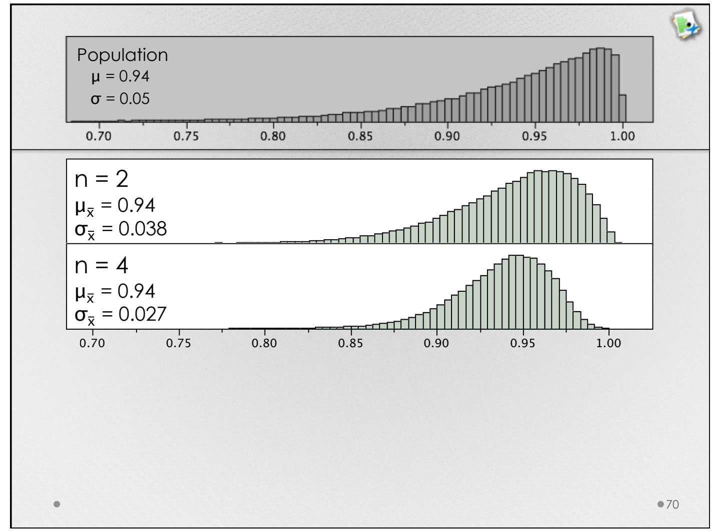

Population
μ = 100 σ = 15
70 80 90 100 110 120 130
n = 2
$\mu_{\bar{x}}$ = **100**
$\sigma_{\bar{x}}$ = 10.6
n = 4
$\mu_{\bar{x}}$ = **100**
$\sigma_{\bar{x}}$ = 7.5
n = 16
$\mu_{\bar{x}}$ = **100**
$\sigma_{\bar{x}}$ = 3.7
n = 64
$\mu_{\bar{x}}$ = **100**
$\sigma_{\bar{x}}$ = 1.9
70 75 80 85 90 95 100 105 110 115 120 125 130
65

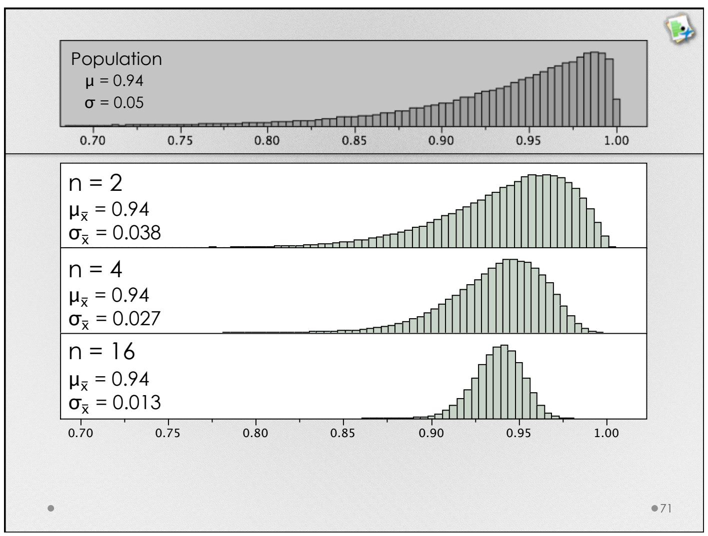

Population
μ = 100 σ = 15
70 80 90 100 110 120 130
n = 2
$\mu_{\bar{x}}$ = 100
$\sigma_{\bar{x}}$ = 10.6
n = 4
$\mu_{\bar{x}}$ = 100
$\sigma_{\bar{x}}$ = 7.5
n = 16
$\mu_{\bar{x}}$ = 100
$\sigma_{\bar{x}}$ = 3.7
n = 64
$\mu_{\bar{x}}$ = 100
$\sigma_{\bar{x}}$ = 1.9
70 75 80 85 90 95 100 105 110 115 120 125 130
66

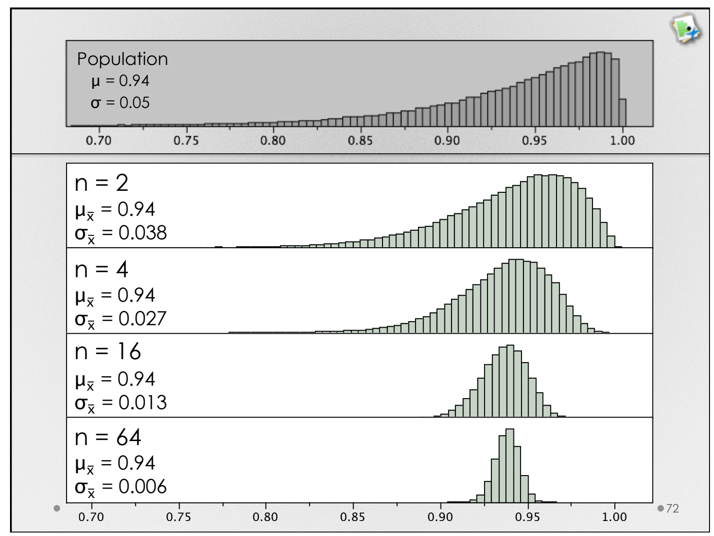

Population
μ = 100 σ = 15
70 80 90 100 110 120 130
n = 2
$\mu_{\bar{x}}$ = 100
$\sigma_{\bar{x}}$ = 10.6
n = 4
$\mu_{\bar{x}}$ = 100
$\sigma_{\bar{x}}$ = 7.5
n = 16
$\mu_{\bar{x}}$ = 100
$\sigma_{\bar{x}}$ = 3.7
n = 64
$\mu_{\bar{x}}$ = 100
$\sigma_{\bar{x}}$ = 1.9
70 75 80 85 90 95 100 105 110 115 120 125 130
67

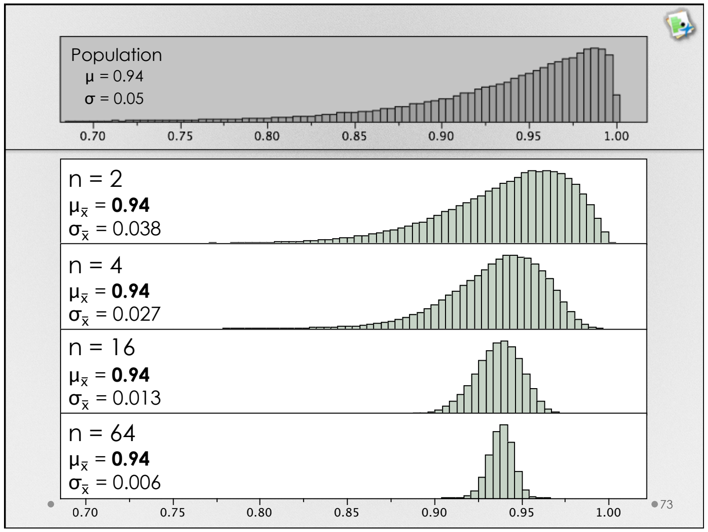

Population
μ = 0.94
σ = 0.05
0.70 0.75 0.80 0.85 0.90 0.95 1.00
68

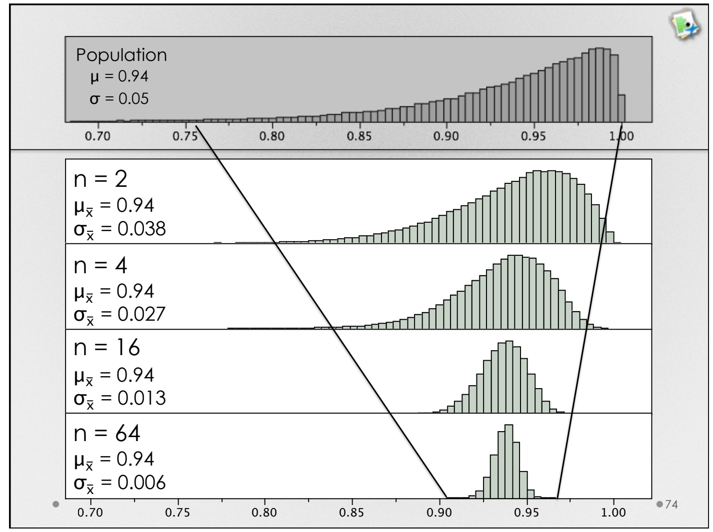

Population
μ = 0.94
σ = 0.05
0.70 0.75 0.80 0.85 0.90 0.95 1.00
n = 2
$\mu_{\bar{x}}$ = 0.94
$\sigma_{\bar{x}}$ = 0.038
0.70 0.75 0.80 0.85 0.90 0.95 1.00
69

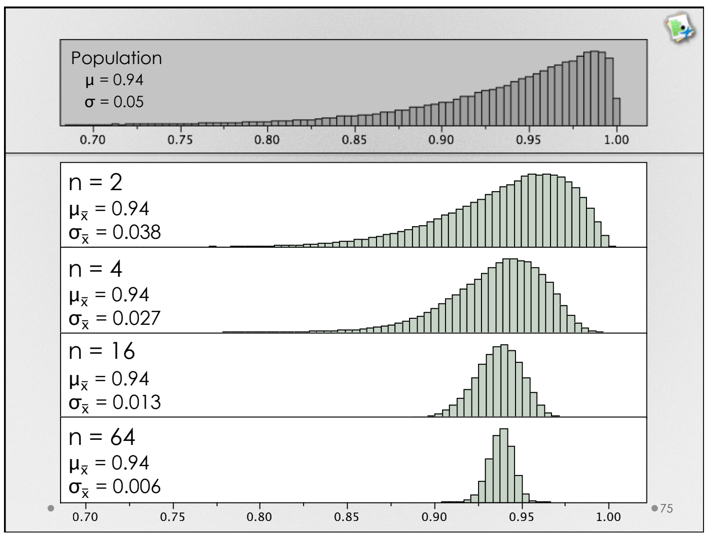

Population
μ = 0.94
σ = 0.05
0.70 0.75 0.80 0.85 0.90 0.95 1.00
n = 2
$\mu_{\bar{x}}$ = 0.94
$\sigma_{\bar{x}}$ = 0.038
n = 4
$\mu_{\bar{x}}$ = 0.94
$\sigma_{\bar{x}}$ = 0.027
0.70 0.75 0.80 0.85 0.90 0.95 1.00
70


Population
μ = 0.94
σ = 0.05
0.70 0.75 0.80 0.85 0.90 0.95 1.00
n = 2
$\mu_{\bar{x}}$ = 0.94
$\sigma_{\bar{x}}$ = 0.038
n = 4
$\mu_{\bar{x}}$ = 0.94
$\sigma_{\bar{x}}$ = 0.027
n = 16
$\mu_{\bar{x}}$ = 0.94
$\sigma_{\bar{x}}$ = 0.013
0.70 0.75 0.80 0.85 0.90 0.95 1.00
71


Population
μ = 0.94
σ = 0.05
0.70 0.75 0.80 0.85 0.90 0.95 1.00
n = 2
$\mu_{\bar{x}}$ = 0.94
$\sigma_{\bar{x}}$ = 0.038
n = 4
$\mu_{\bar{x}}$ = 0.94
$\sigma_{\bar{x}}$ = 0.027
n = 16
$\mu_{\bar{x}}$ = 0.94
$\sigma_{\bar{x}}$ = 0.013
n = 64
$\mu_{\bar{x}}$ = 0.94
$\sigma_{\bar{x}}$ = 0.006
0.70 0.75 0.80 0.85 0.90 0.95 1.00
72







Characteristics of the Sampling Distribution of $\bar{X}$

- Sampling distribution always has a mean = μ
- As sample size (n) increases, the standard deviation of the sampling distribution decreases by the square root of the sample size (n) $\boxed{\frac{\sigma}{\sqrt{n}}}$
- The sampling distribution is *normally* distributed
  …*under what conditions?*

76

Characteristics of the Sampling Distribution of $\overline{X}$

- Sampling distribution <u>always</u> has a mean = μ
- As sample size (n) increases, the standard deviation of the sampling distribution decreases by the square root of the sample size (n) $\frac{\sigma}{\sqrt{n}}$
- The sampling distribution is *normally* distributed *when population is normal <u>or</u> samples are large, n > ~30*

77

Characteristics of the Sampling Distribution of $\overline{X}$

- Sampling distribution <u>always</u> has a mean = μ
- As sample size (n) increases, the standard deviation of the sampling distribution decreases by the square root of the sample size (n) $\frac{\sigma}{\sqrt{n}}$
- The sampling distribution is *normally* distributed *when population is normal <u>or</u> samples are large, n > ~30*

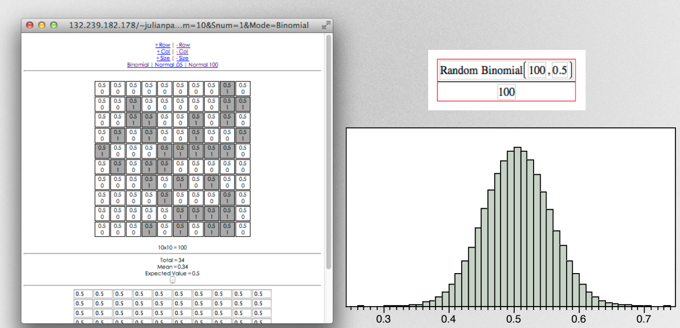
78

The Central Limit Theorem

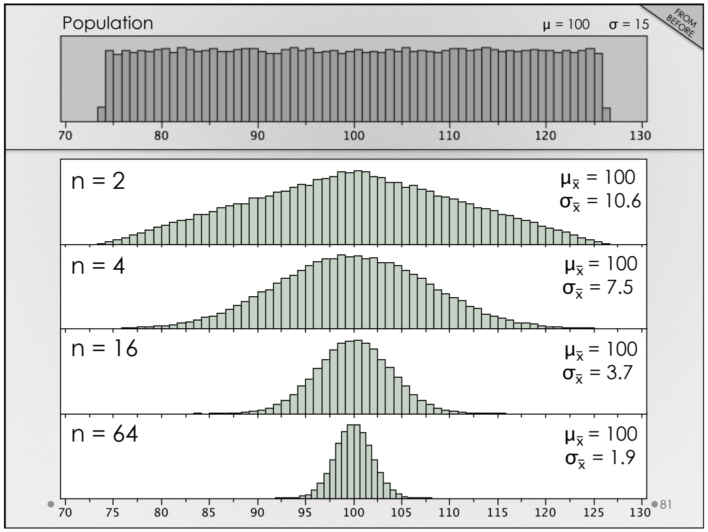
the distribution of the sum of a large number of independent random variables approaches a normal distribution

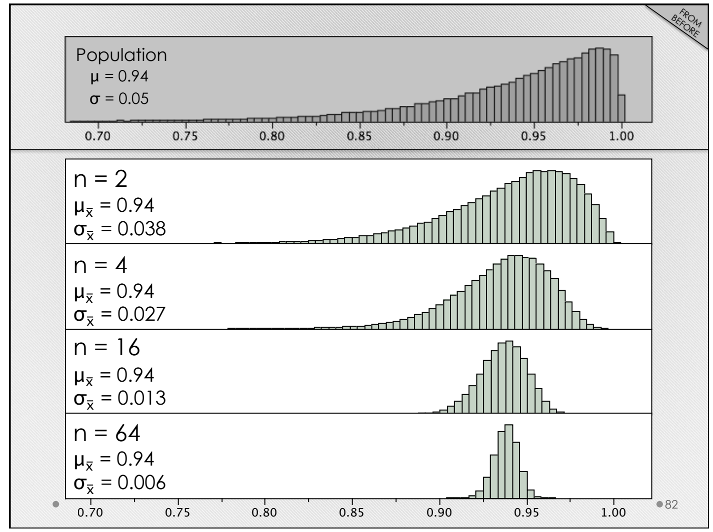
79

FROM BEFORE



80



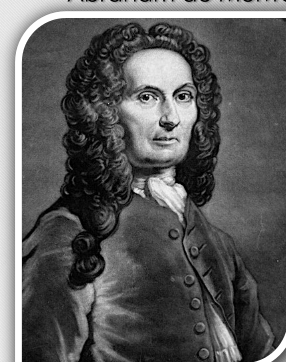


The Central Limit Theorem

the distribution of the sum of a large number of independent random variables approaches a normal distribution

83

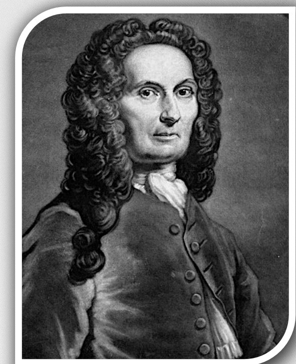
Abraham de Moivre



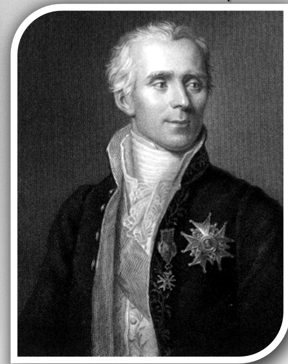
The Doctrine of Chances
1718

84

Abraham de Moivre



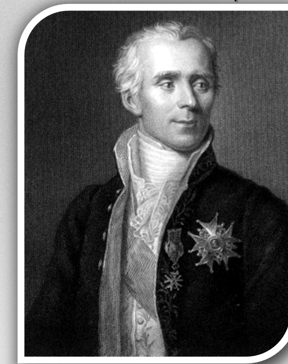
Pierre-Simon Laplace



85

Théorie Analytique des Probabilités
1820

Pierre-Simon Laplace

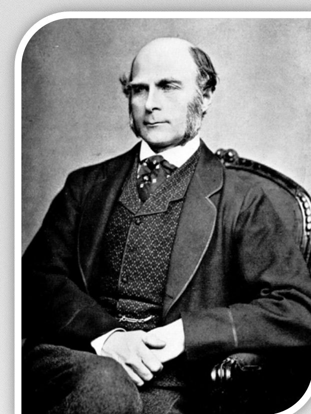


86

The Central Limit Theorem

the distribution of the sum of a large number of independent random variables approaches a normal distribution

87



Sir Francis Galton
1822 - 1911

88