Friday, February 1, 2019

- **Warm-Up**
  - Suppose that a particular candidate for a public office is in fact favored by 48% of all registered voters in the district. A polling organization will take a random sample of 500 voters and will use $\hat{p}$, the sample proportion, to estimate p. What is the approximate probability that $\hat{p}$ will be greater than 0.5, causing the polling organization to incorrectly predict the result of the upcoming election?

- **Distribution of Sample Means**

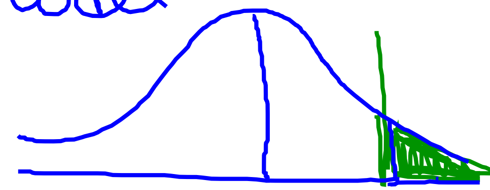
<u>Objectives</u>

Content: I will find the mean and standard deviation of a sampling distribution and apply the Normal model to determine probability.
Social: I will listen and focus on the lesson despite distractions.
Language: I will use correct vocabulary and clearly ask questions when I do not understand.

Warm-up

$N(0.48, 0.022)$ ← model

$\sqrt{\frac{(.48)(.52)}{500}}$



Suppose that a particular candidate for a public office is in fact favored by 48% of all registered voters in the district. A polling organization will take a random sample of 500 voters and will use $\hat{p}$, the sample proportion, to estimate p. What is the approximate probability that $\hat{p}$ will be greater than 0.5, causing the polling organization to incorrectly predict the result of the upcoming election?

10%

$500 \cdot 0.48 \geq 10$

$500 \cdot 0.52 \geq 10$

15-18% depending upon rounding

$z = \frac{0.5 - 0.48}{0.022} \approx 0.9$

<u>Objectives</u>

Content: I will find the mean and standard deviation of a sampling distribution and apply the Normal model to determine probability.

Social: I will listen and focus on the lesson despite distractions.

Language: I will use correct vocabulary and clearly ask questions when I do not understand.

Modeling the Distribution of Sample Means

- When working with sample means:
  - *The **mean** is the **population mean***
  - *The **standard deviation** is a variation on the population standard deviation based on the size of the sample*

$$N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

Assumptions and Conditions

The CLT requires essentially the same assumptions we saw for modeling proportions approximate 30

Sample Size Assumption

The sample size must be sufficiently large.

Independence Assumption

The sampled values must be independent of each other.

To check independence…

Randomization Condition

The data values must be sampled **randomly**.

or representative of population

To check independence…

10% Condition

When the sample is drawn without replacement, the sample size, *n*, should be no more than 10% of the population.

To check independence…

Large Enough Sample Condition

The CLT doesn't tell us how large a sample we need. For now, you need to think about your sample size in the context of what you know about the population.

Practice

$N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$

An elevator in large office building can safely carry up to 5000 pounds of people. A study shows that the population of elevator riders as a **μ=148** pounds and **σ=15.2** pounds. A sign allows up to 32 passengers and the elevator is filled with **32** randomly selected riders.

Does it meet our conditions for a Normal model?

- Randomization condition   Stated "randomly selected"
- 10% condition $32 \times 10 = 320 <$ total pop. of elevator riders
- Large enough sample condition $32 \geq 30$ good :)

What is the mean of the sampling distribution of μ?

$\mu = 148$

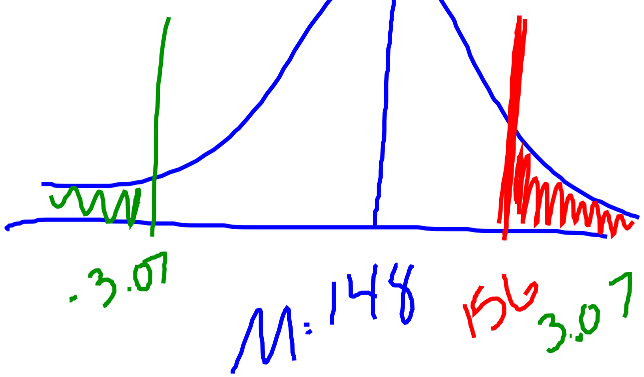
Find the standard deviation. $\sigma = \frac{15.2}{\sqrt{32}} = 0.687$

An elevator in large office building can safely carry up to 5000 pounds of people. A study shows that the population of elevator riders as a **μ=148** pounds and **σ=15.2** pounds. A sign allows up to 32 passengers and the elevator is filled with **32** randomly selected riders.

5000 ÷ 32 = 156.25

Find the probability that the 5000 pound limit would be exceeded. Do you have any concerns about this elevator ride?

σ = 2.687



$$Z = \frac{156.25 - 148}{2.687}$$

= 3.07

$P(z \geq 3.07)$ = 0.11%

0.0011

$$N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

A traffic study shows that for a certain bridge, the mean number of occupants in a car is **1.85** people, and the standard deviation is **0.31** people. We take a random sample of **64 cars**.

Does it meet our conditions for a Normal model?

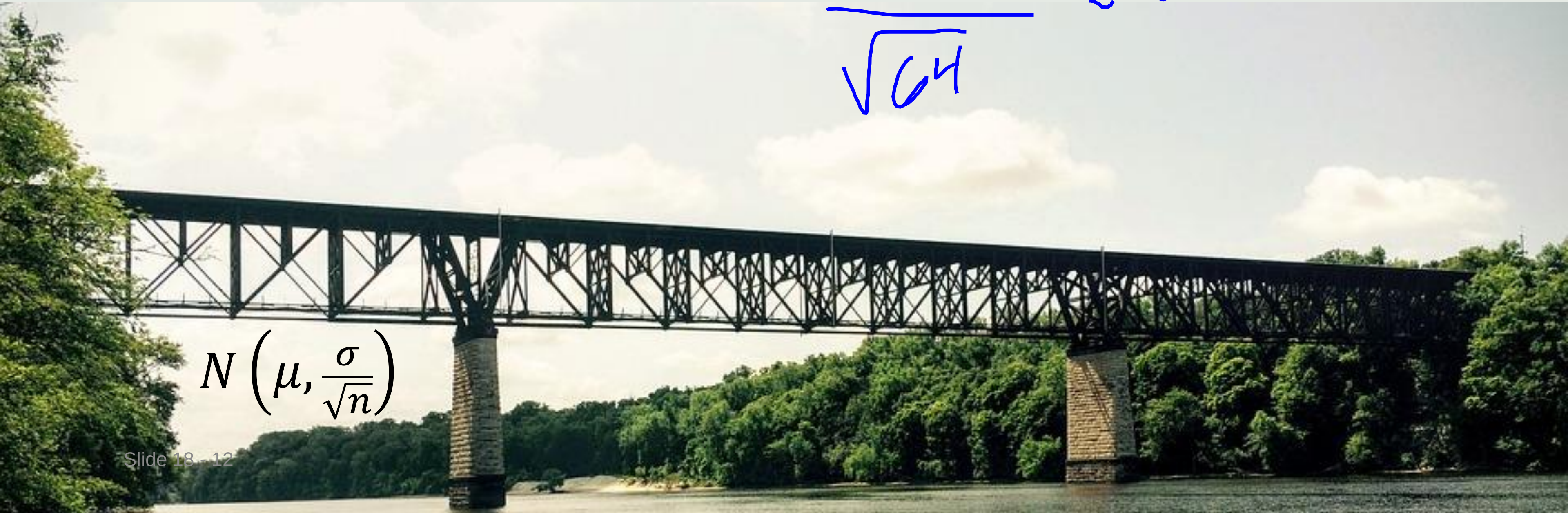
- Randomization condition
- 10% condition — $64 \times 10 = 640 \le$ to pop. of cars
- Large enough sample condition — $64 \ge 30$ - great ☺

What is the mean of the sampling distribution of μ?

$\mu = 1.85$

Find the standard deviation.

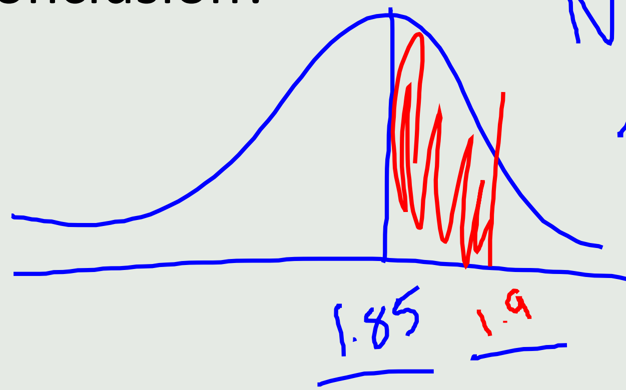
$$\frac{0.31}{\sqrt{64}} \approx 0.038$$



$$N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

A traffic study shows that for a certain bridge, the mean number of occupants in a car is 1.85 people, and the standard deviation is 0.31 people. We take a random sample of 64 cars.

One day observers estimated the mean number of occupants to be between 1.85 and 1.9 people. How confident are you in this conclusion?



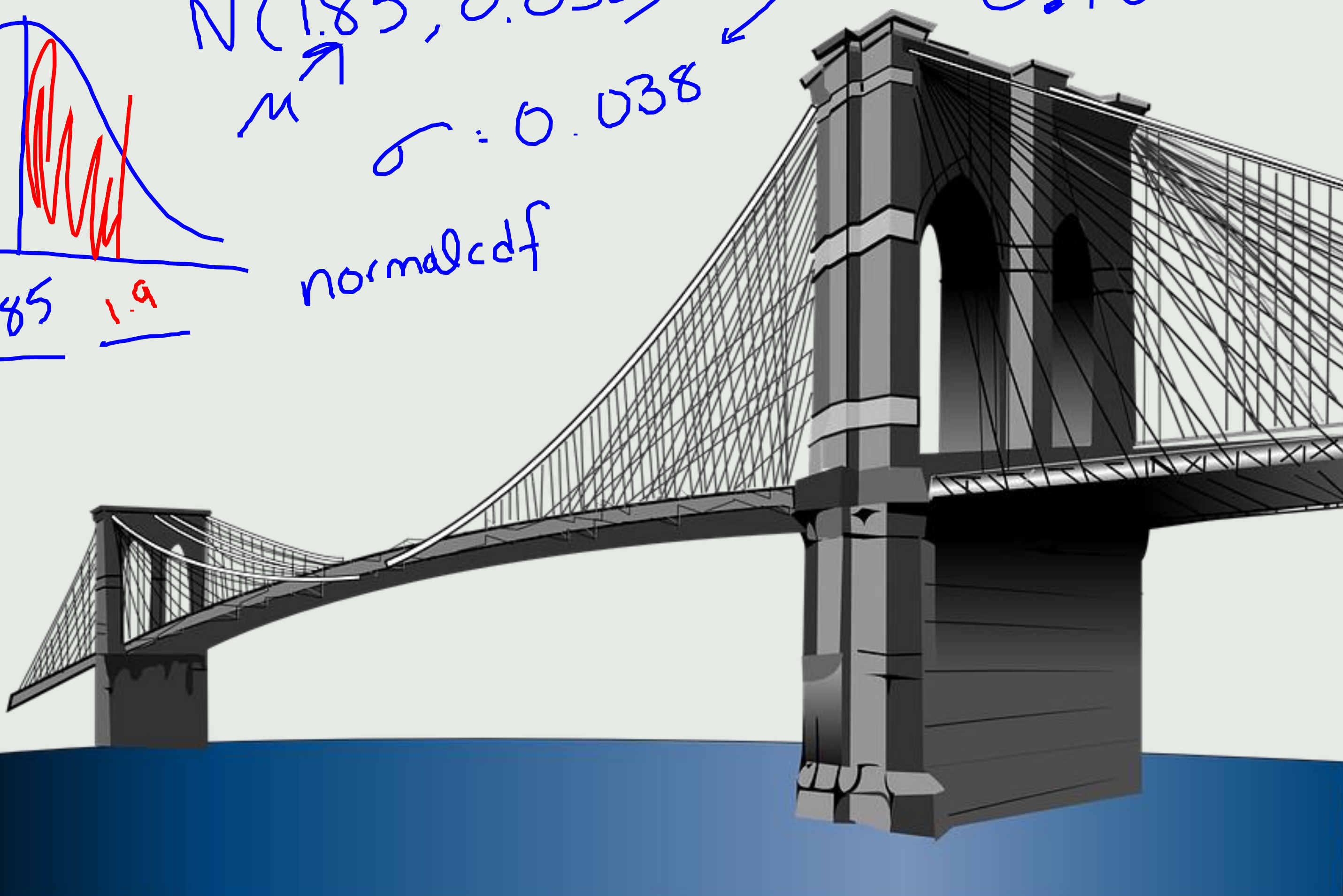
N(1.85, 0.038)

μ

σ = 0.038

normalcdf

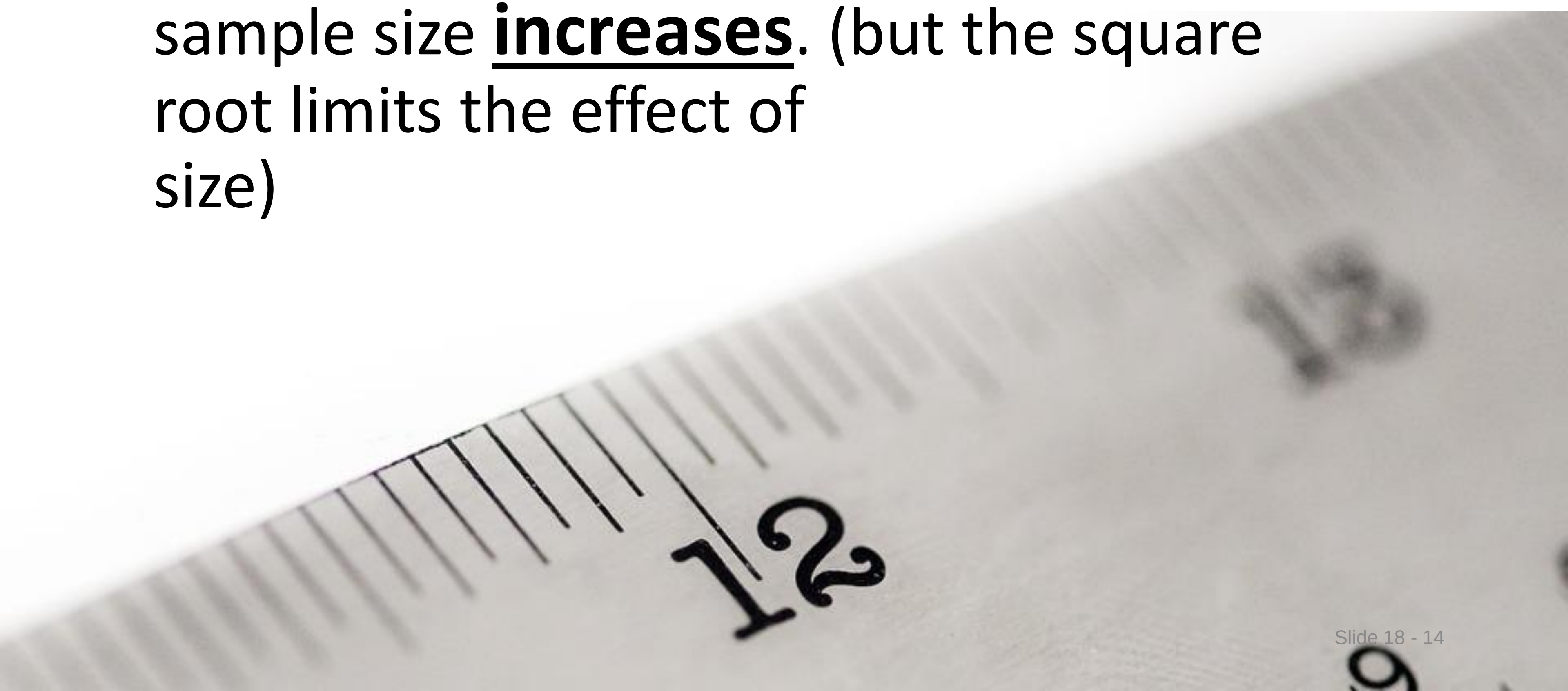
0.4058



$$N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

Size Matters

- The standard deviation of the sampling distribution declines *only* with the square root of the sample size
- Therefore, the variability **<u>decreases</u>** as the sample size **<u>increases</u>**. (but the square root limits the effect of size)



The Real World and the Model World

Be careful! Now we have *two* distributions to deal with.

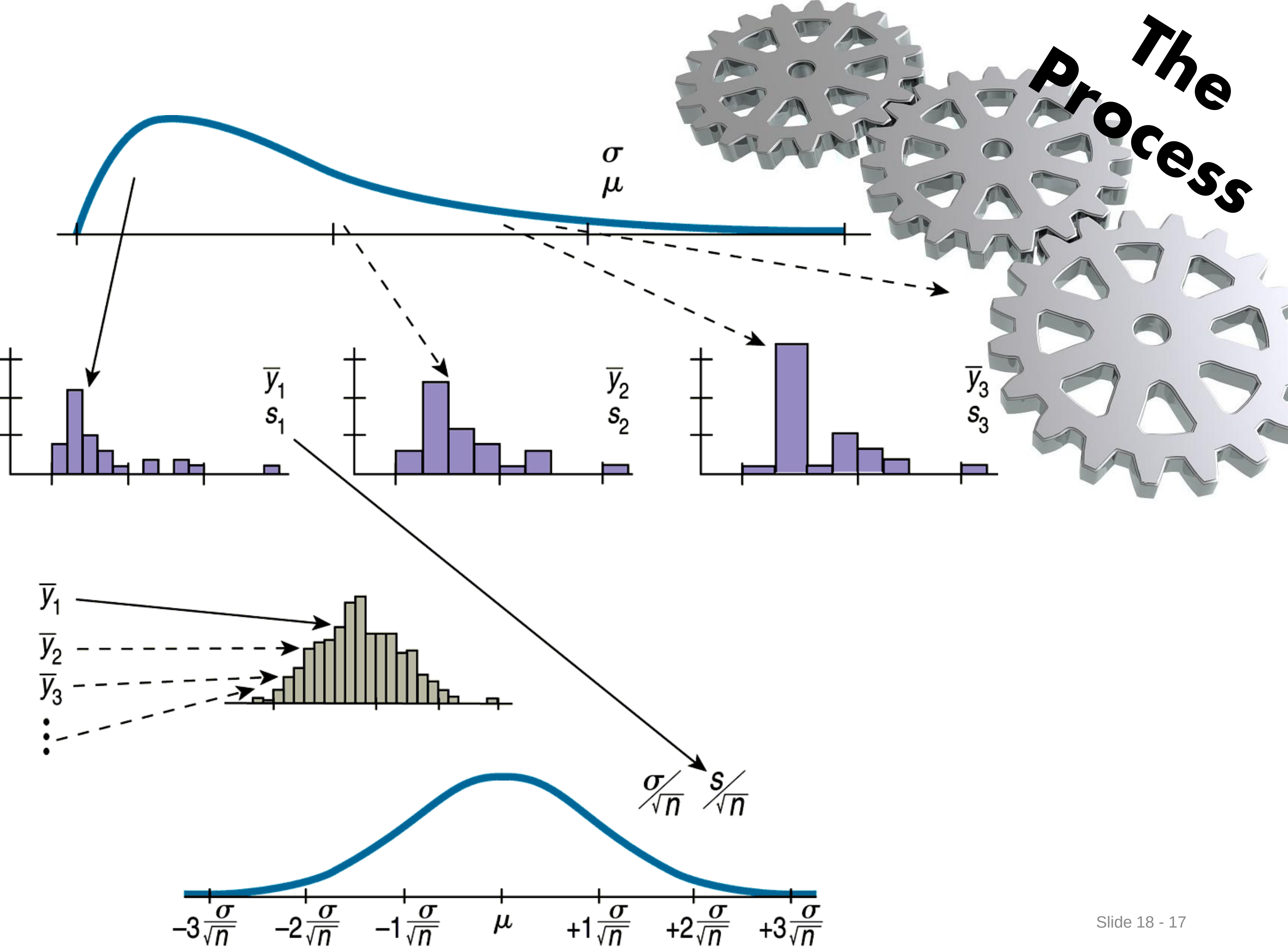
- The first is the real world distribution of the sample, which we might display with a histogram.
- The second is the math world *sampling distribution* of the statistic, which we model with a Normal model based on the Central Limit Theorem.

Don't confuse the two!

BASIC TRUTHS

1. Sampling distributions arise because samples vary. Each random sample will have different cases and, so, a different value of the statistic.
2. Although we can always simulate a sampling distribution, the Central Limit Theorem saves us the trouble for means and proportions.

The Process



σ
μ
$\bar{y}_1$
s_1
$\bar{y}_2$
s_2
$\bar{y}_3$
s_3
$\bar{y}_1$
$\bar{y}_2$
$\bar{y}_3$
$\sigma/\sqrt{n}$
$s/\sqrt{n}$
$-3\frac{\sigma}{\sqrt{n}}$
$-2\frac{\sigma}{\sqrt{n}}$
$-1\frac{\sigma}{\sqrt{n}}$
μ
$+1\frac{\sigma}{\sqrt{n}}$
$+2\frac{\sigma}{\sqrt{n}}$
$+3\frac{\sigma}{\sqrt{n}}$

Homework

Page 436 (37-40)