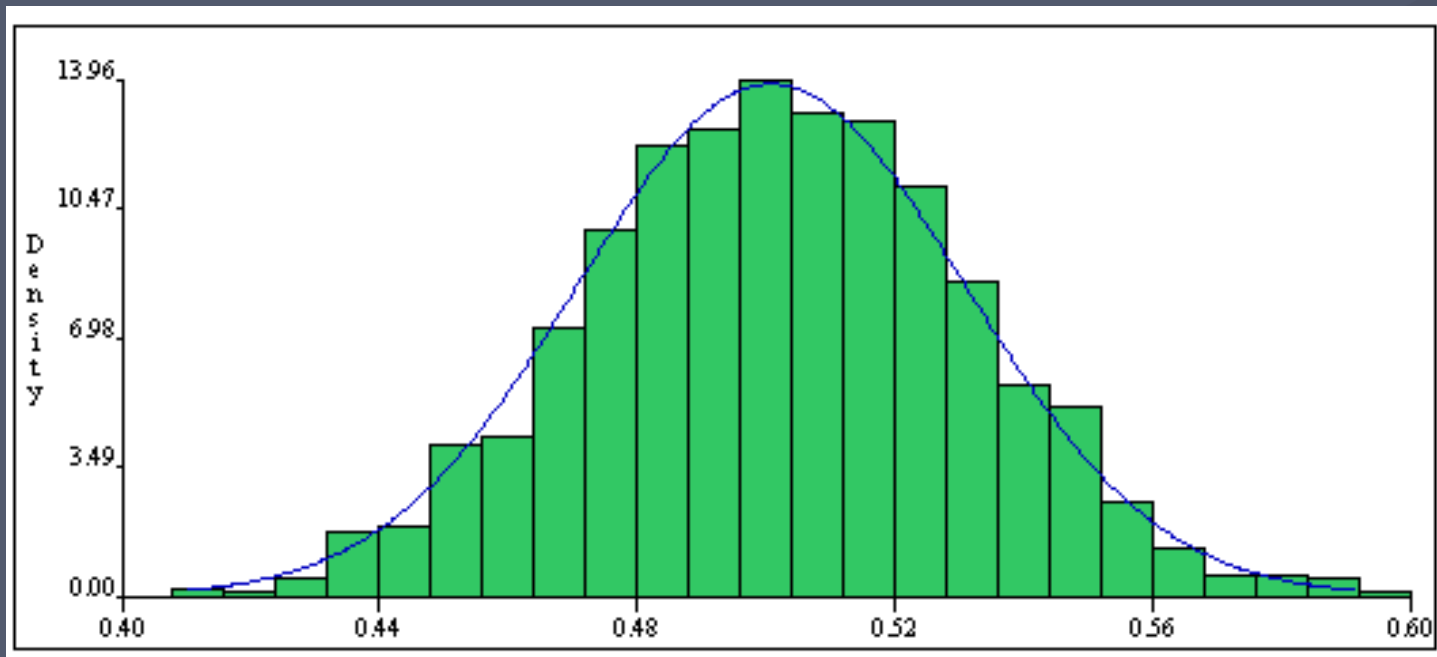


Central Limit Theorem





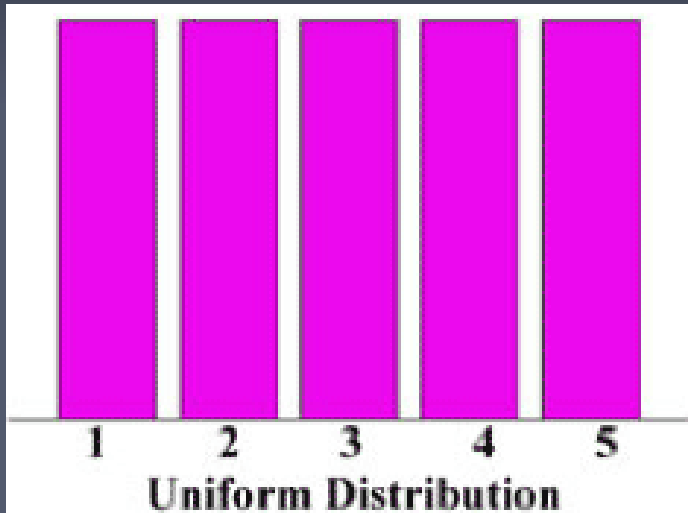
**REMEMBER
IT**

- Complete Tasks 1 and 2 on worksheet.

Your task today...

- Is to be a noticer
 - As you see each slide ask yourself:
 - What do I notice?

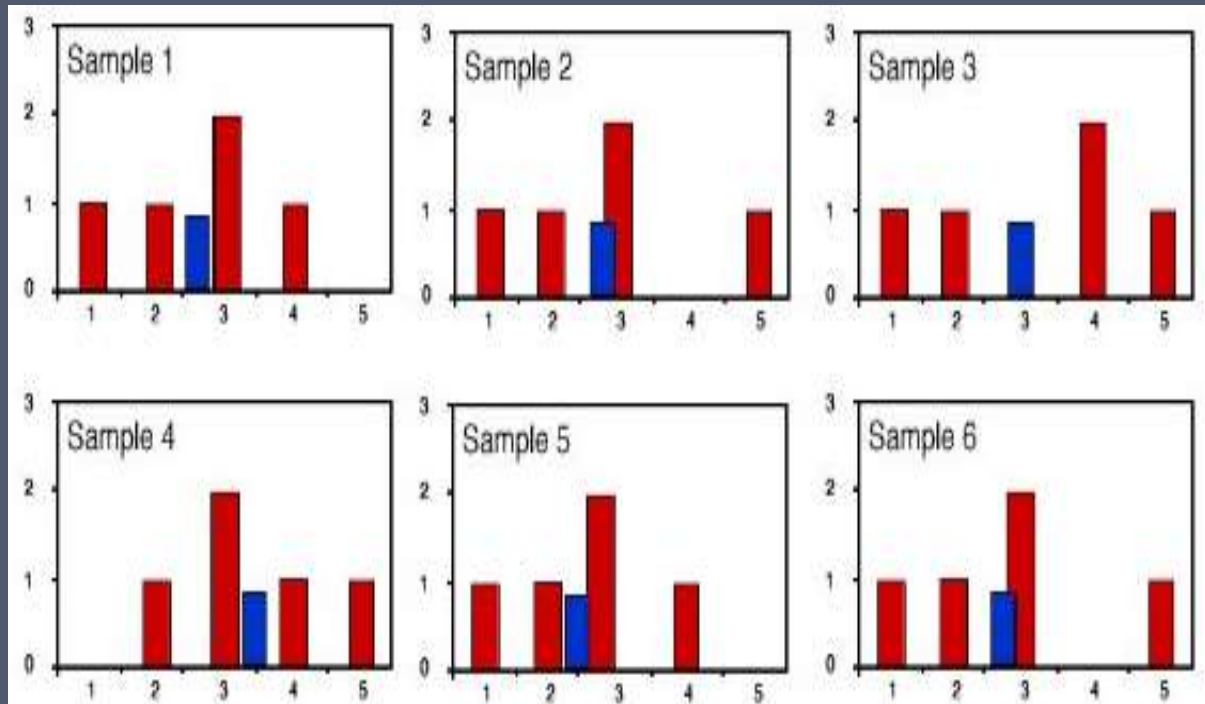
Intro to CLT



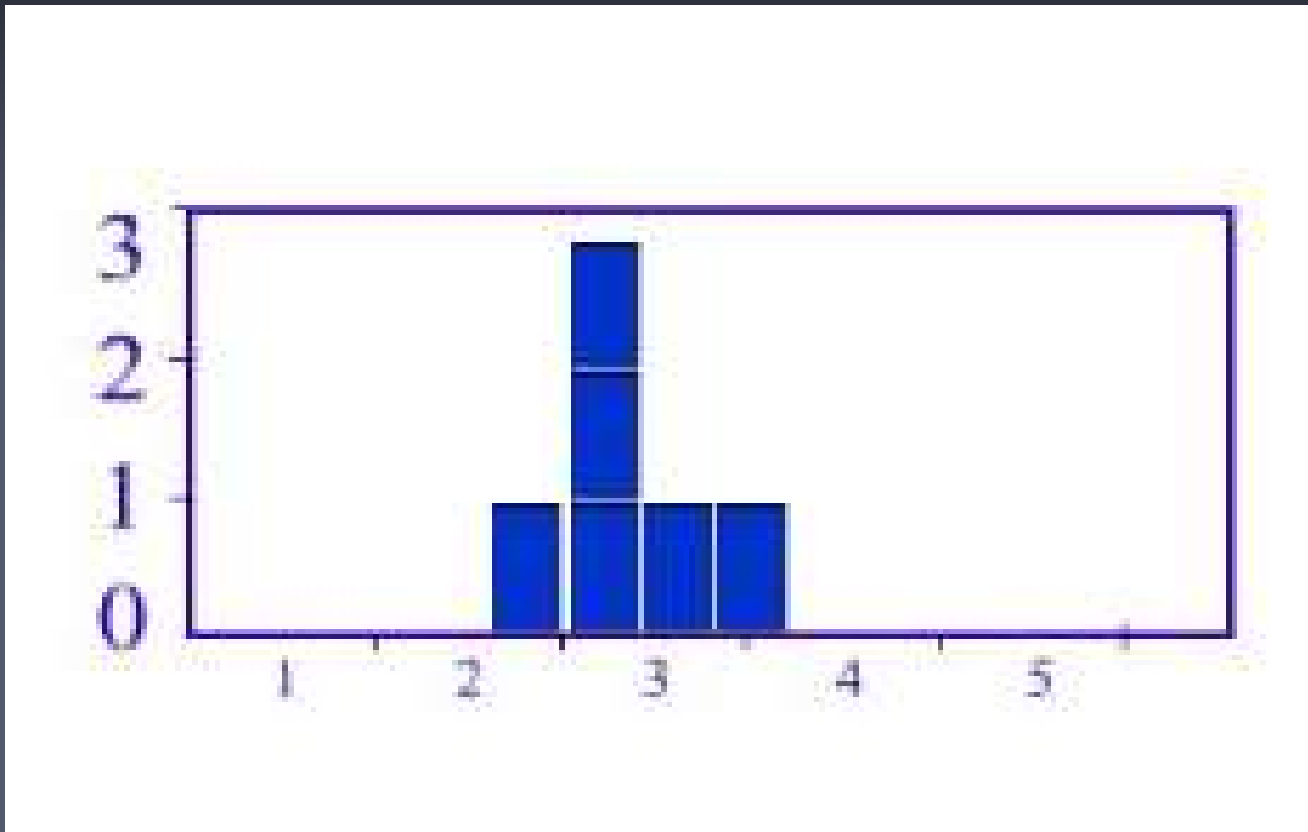
- We will begin by drawing samples from a population that has a uniform distribution.
- The scores in the **population** range from 1 to 5. The mean of the **population** is 3.0 and the standard deviation of the **population** is 1.41 [$\mu=3.0$; $\sigma=1.41$].
- Each level of the variable had the same probability of occurring. The population distribution is presented left.

Six successive samples...

- Let's start small and draw 6 samples randomly from the **uniform population** presented on the previous page.
- Each sample has only 5 subjects ($N=5$). The **central limit theorem** requires that **all samples** must have the **same** sample size. The distribution of scores in each of these samples is presented below.
- The sample mean is shown in blue.



Distribution of sample means



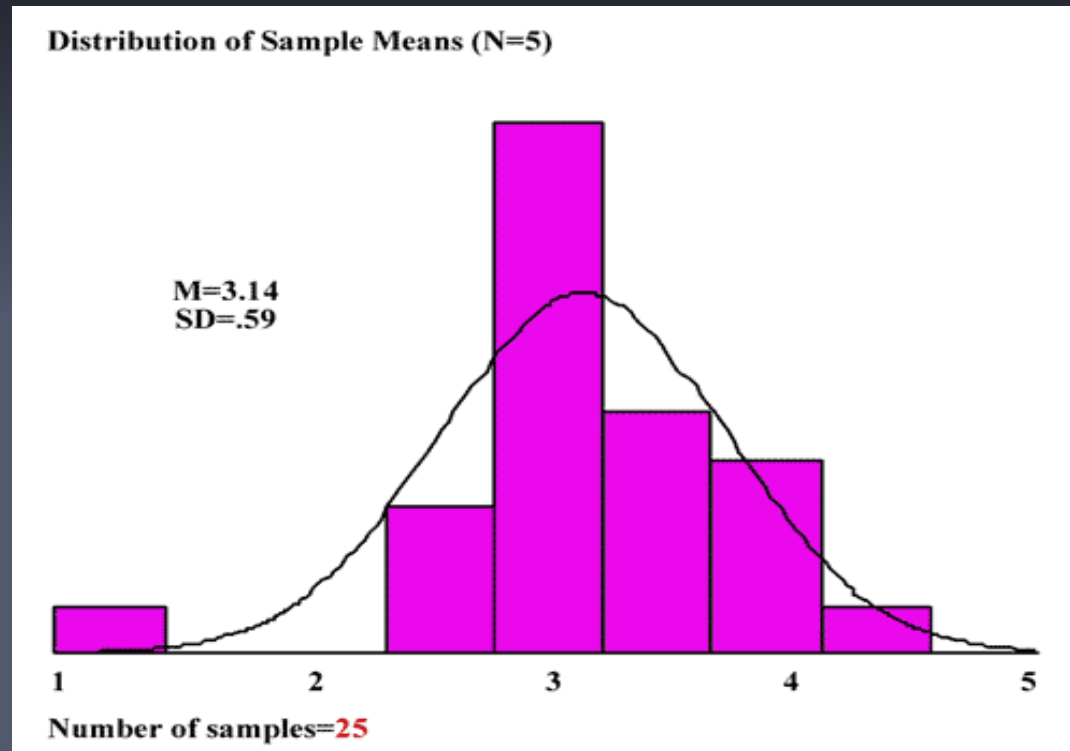
■ What do you notice?

25 additional sample means...

Sample Means

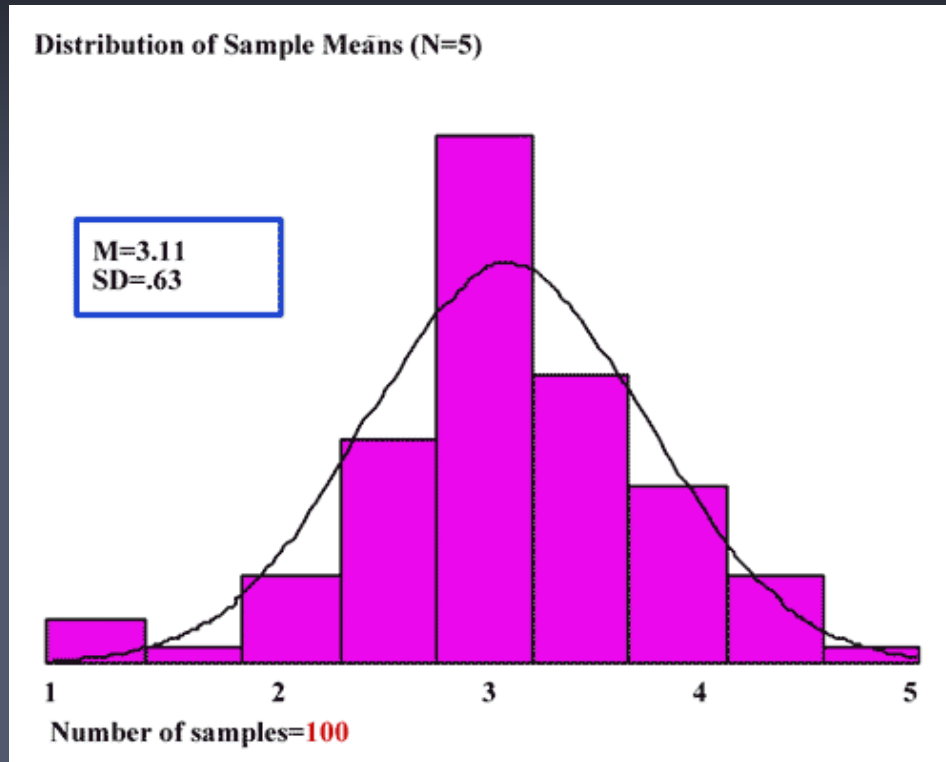
2.60	2.60	3.60	3.40	3.80
2.80	2.80	2.80	3.80	2.80
3.20	3.20	3.20	2.80	3.00
3.40	3.40	2.80	1.40	2.60
3.40	4.20	4.00	3.20	3.80

Frequency Distribution



- What do you notice?
- Let's see what happens when we randomly select 100 samples from the same population.

Compare and Contrast it



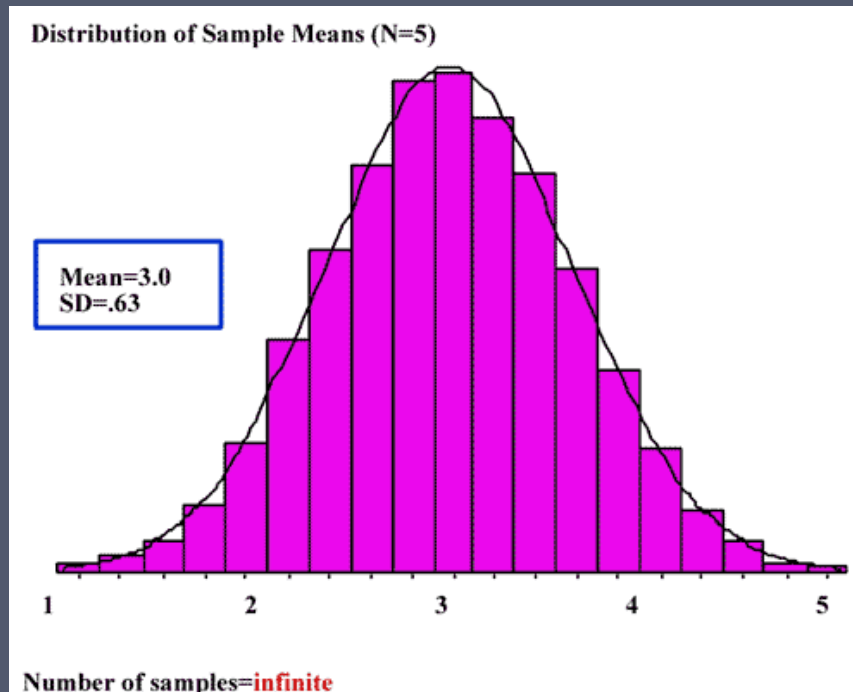
- Remember these sample means were taken from a uniform distribution!
- Compare and contrast this to the two previous frequency distributions...

Compare and Contrast it

- We can see that this distribution of sample means is even closer to a normal distribution than the distributions calculated with 6 and 25 samples.
- The mean of this distribution is quite close to the population mean of 3
- The standard deviation is smaller than the population standard deviation of 1.41

Infinite samples

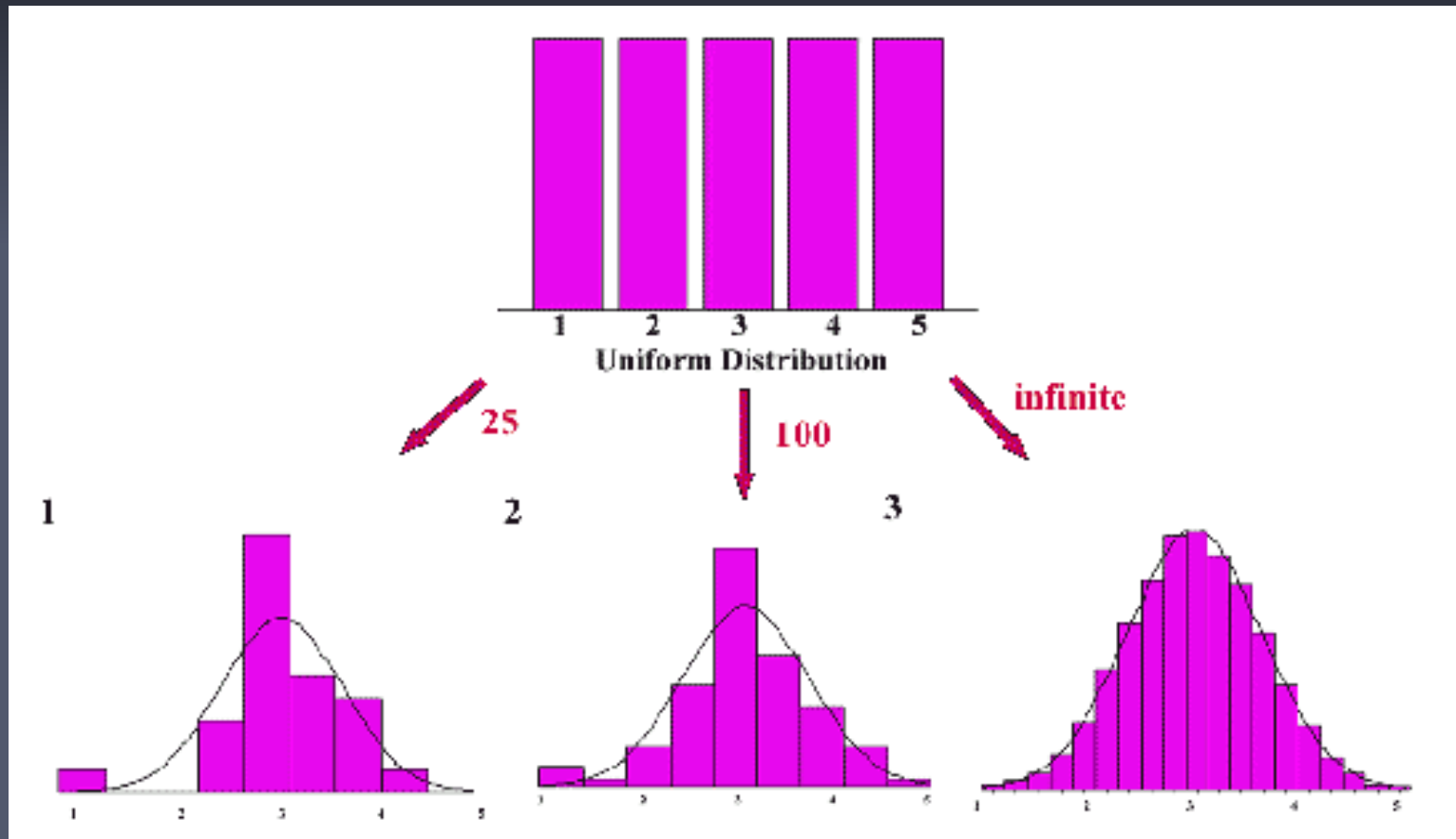
- Let's see what would happen if we were to randomly select an "infinite" number of samples from the same population.



What do you notice?

- The sampling distribution of the mean is normally distributed.
- The mean of the normal distribution is the same as the population mean.
- The standard deviation of the normal distribution is equal to the population standard deviation divided by the square root of the sample size.

Summary so far



Let's increase the sample size.

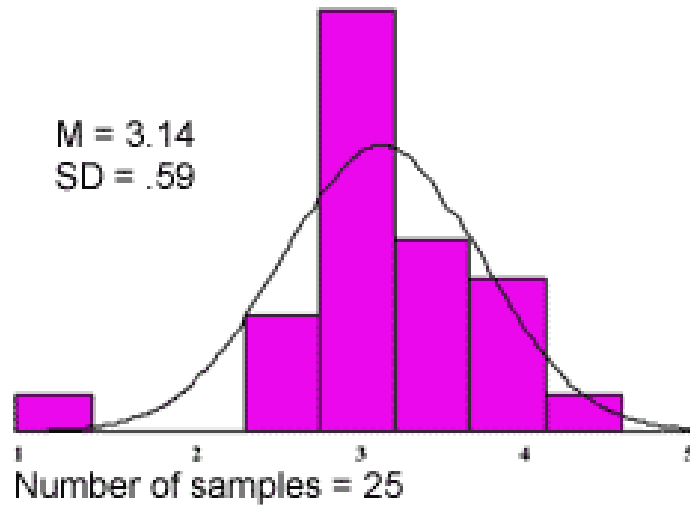
- Let's increase the sample size from 5 to 25...

Sample Means

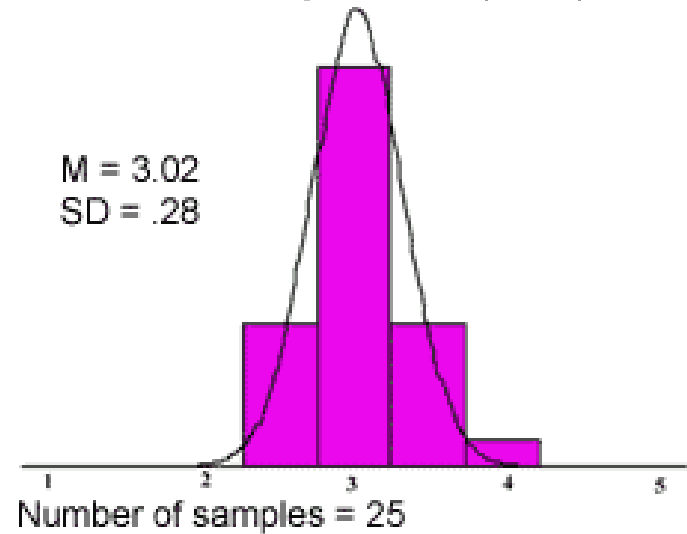
3.20	3.28	3.32	3.24	3.00
2.92	3.72	3.04	3.00	2.64
2.64	3.20	2.88	2.92	2.52
2.88	2.68	2.88	2.68	3.00
3.28	3.12	3.16	3.36	2.92

What do you notice

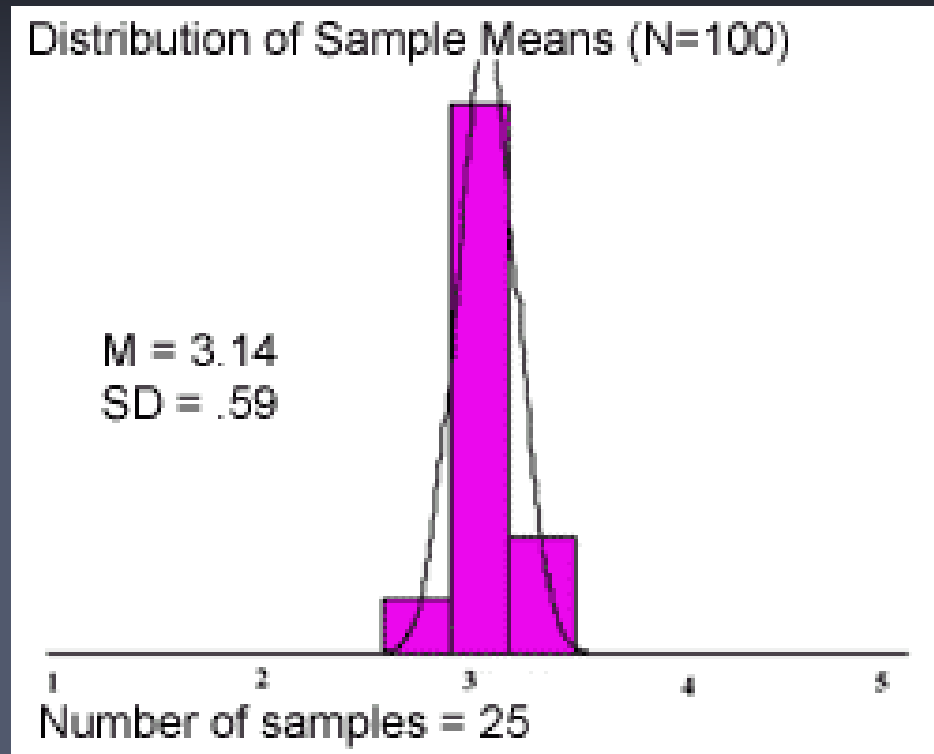
Distribution of Sample Means (N=5)



Distribution of Sample Means (N=25)



Let's increase n to 100...

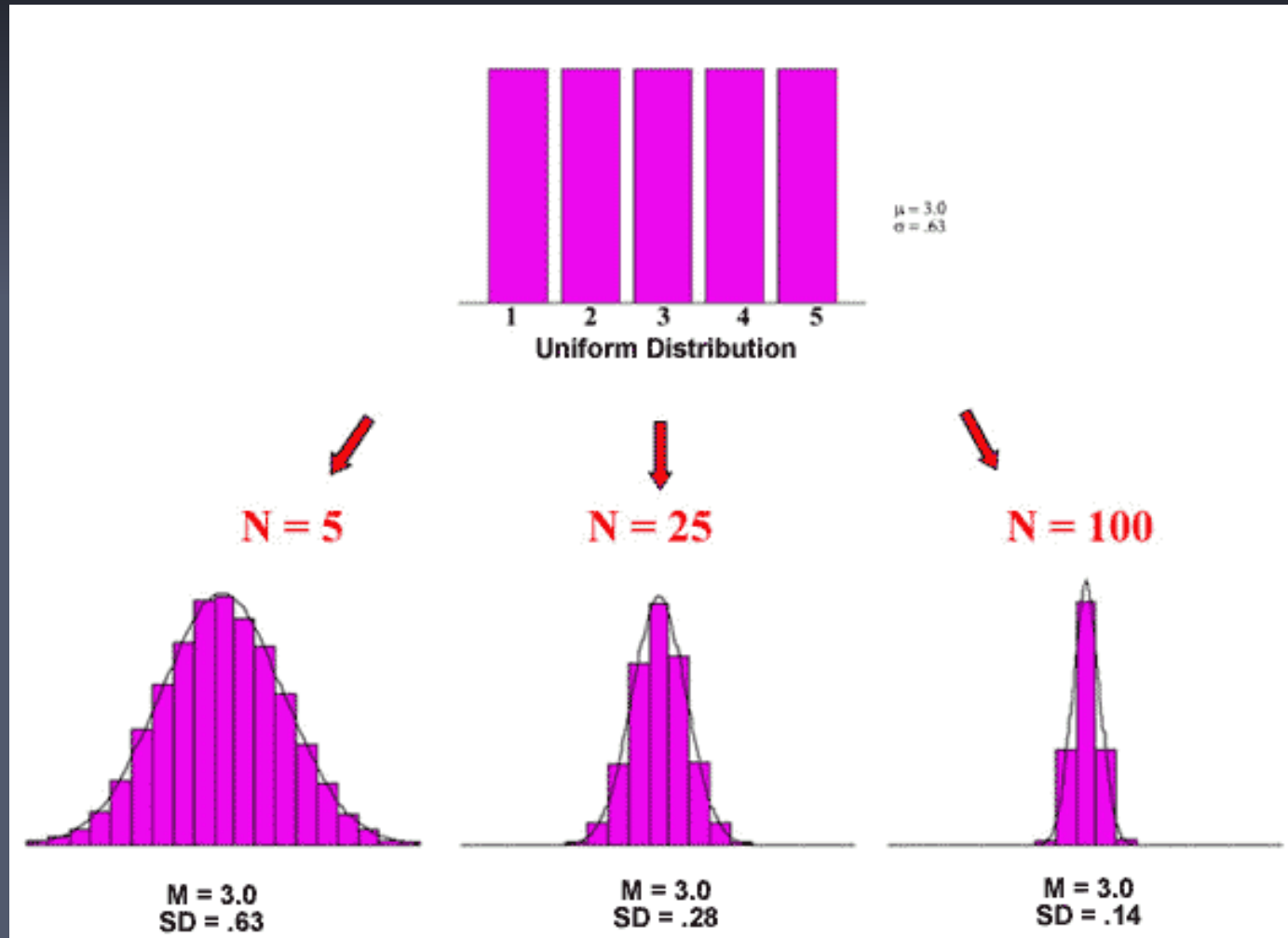


- What do you notice?
- What does this say about sample size?

Central Limit Theorem

- The **central limit theorem** states that when an **infinite** number of **successive random** samples are taken from a population, the **sampling distribution of the means** of those samples will become **approximately normally** distributed with mean μ and standard deviation σ/\sqrt{N} as the **sample size (N)** becomes **larger**, irrespective of the shape of the population distribution.

Sample Size

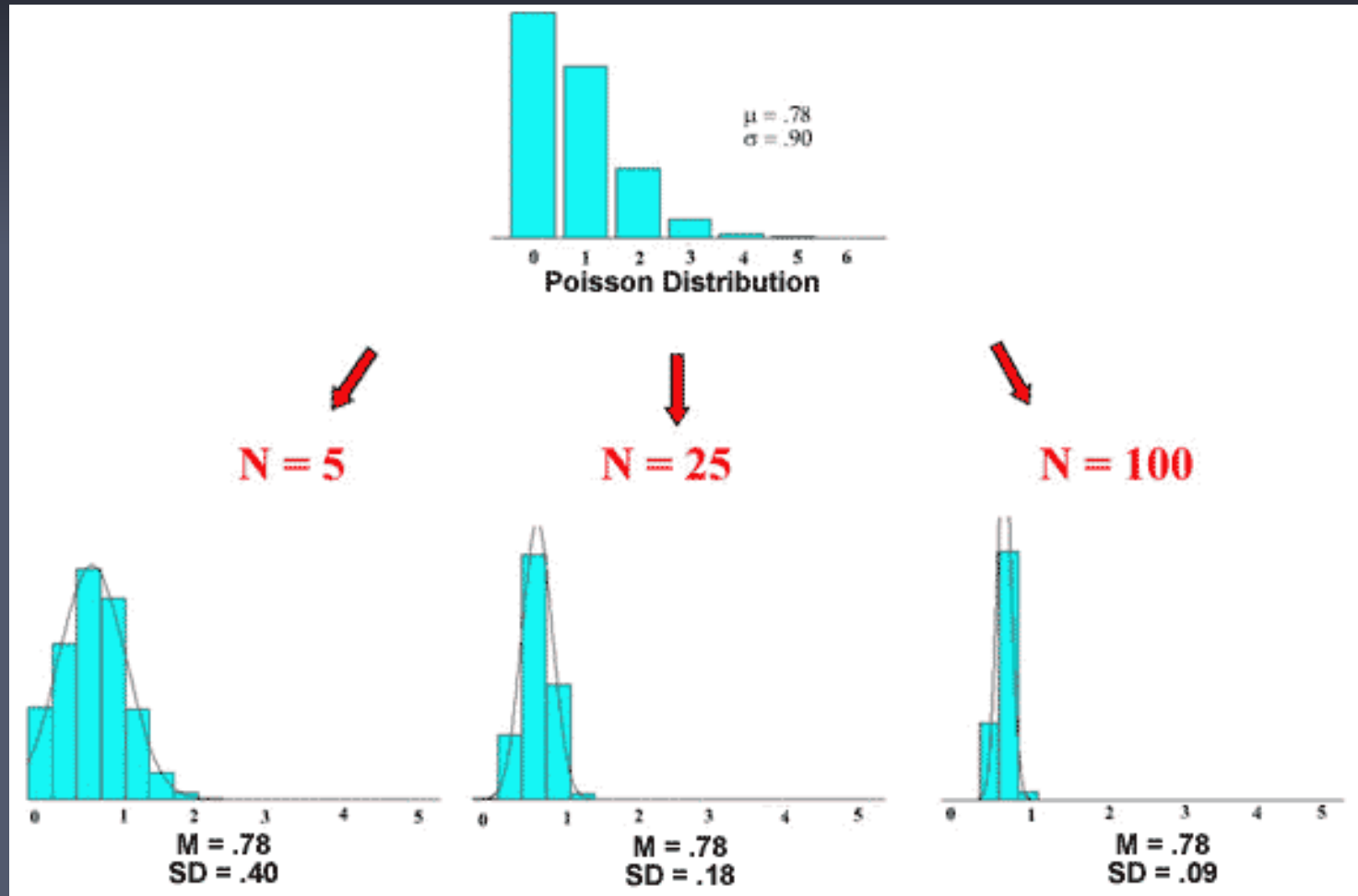


■ What do you notice? – fill in on handout

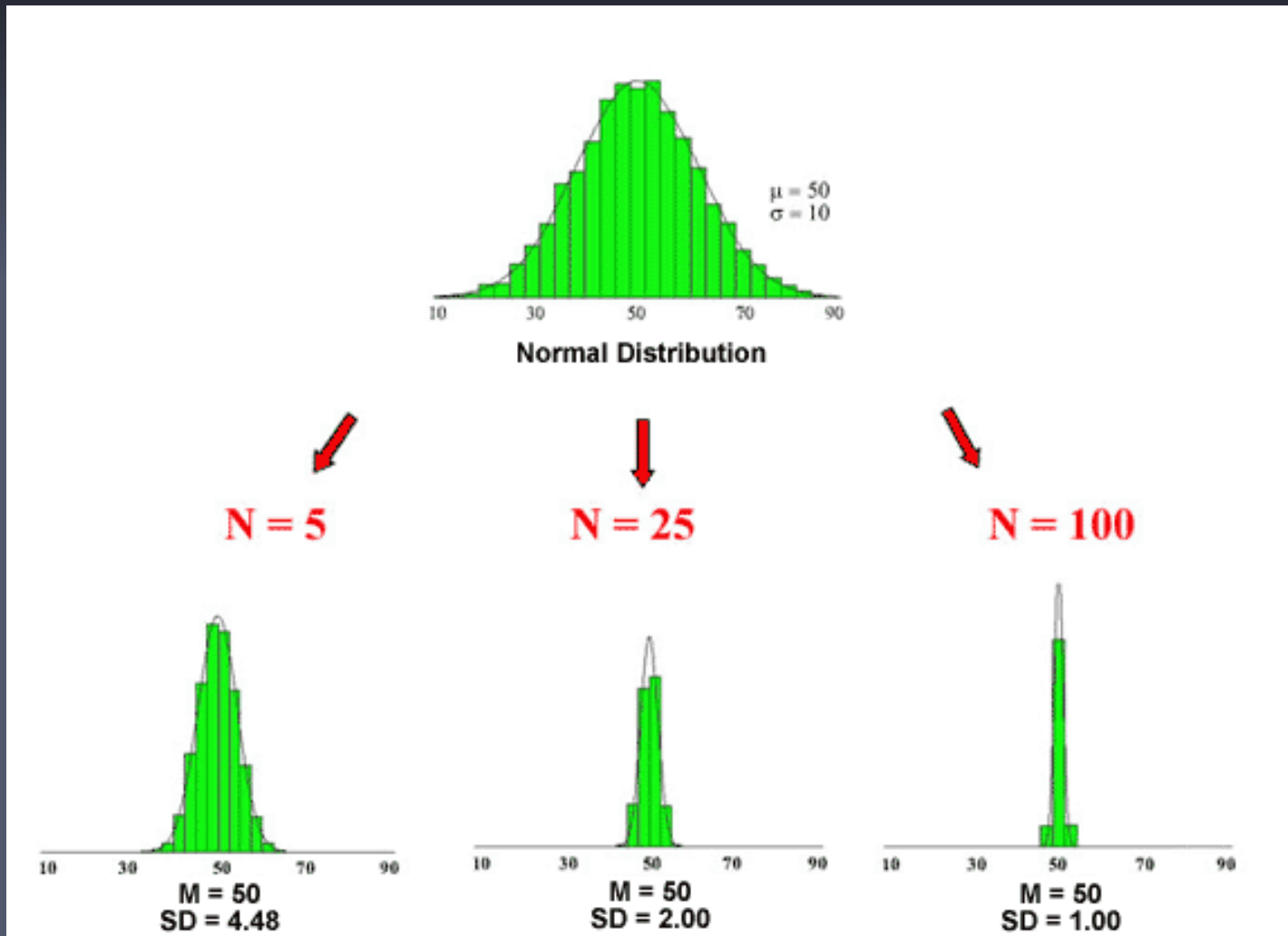
What can we see from these comparisons?

- With "**infinite**" numbers of successive random samples, the sampling distributions all have a **normal** distribution with a **mean of 3.0**, which is equal to the population mean ($\mu = 3.0$).
- As the **sample sizes increase**, the **variability** of each sampling distribution **decreases**. The **range** of the sampling distribution is **smaller** than the range of the original population.
- The **standard deviation** of each sampling distribution is equal to σ/\sqrt{N} (*where N is the size of the sample drawn from the population*).
- Taken together, these distributions suggest that the **sample mean** provides a **good estimate** of μ and that **errors in our estimates** (indicated by the **variability of scores** in the distribution) **decrease** as the **size of the samples** we draw from the population **increases**.

Starting with a Poisson Distribution



Starting with a Normal Distribution



Quiz

- Complete the quiz on your worksheet with a partner.

Quiz – Question One

- You randomly draw an infinite number of samples of 10 subjects from a population, calculate the mean for each sample, and plot the sample means. The distribution that you obtain from plotting sample means is called the...
 - A. Standard error of the mean
 - B. Sampling distribution of the mean
 - C. Sample means normal distribution
 - D. Sample error of the estimate

Answer:

B. Sampling distribution of the mean

Question 2

- A colleague plots a set of sample means drawn from 100 successive samples of 50 subjects. He is not sure what the distribution should look like and asks for your opinion. You tell him that the distribution should be approximately normally distributed and fairly smooth because...
 - A. The sampling distributions of all statistics eventually become perfectly normally distributed.
 - B. There are never outliers that will affect the distribution when you take over 25 samples.
 - C. He took a large number of large samples from the population.
 - D. The population from which it was drawn was normally distributed.

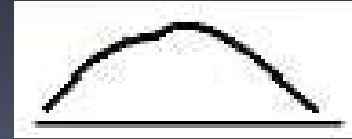
Answer

C. He took a large number of large samples from the population.

Question 3

- Which of the following distributions is most likely if you draw 15 samples of 6 subjects from a population whose distribution is unknown, calculate the mean of each sample, and graph the values of the sample means...

A.



B.



C.



D. Can't be determined

Answer

A.



Question 4

■ Which of these statements is true?

- A. The central limit theorem gives the exact probability of estimating the true mean.
- B. The central limit theorem only applies when the population distribution is normal.
- C. The range of values for the sampling distribution of means is larger than the range of scores in the population.
- D. The central limit theorem requires that all samples are randomly selected from a single population.

Answer:

- D. The central limit theorem requires that all samples are randomly selected from a single population.

Question 5

■ The central limit theorem states that the standard deviation of the mean must equal the:

- A. Sample standard deviation
- B. Population standard deviation
- C. Standard deviation of the distribution of sample means
- D. Average of scores across all random samples selected.

Answer:

C. Standard deviation of the distribution of sample means

Question 6

- How does the central limit theorem show us that the mean is a good estimator of the population mean (μ)?
 - A. It is the easiest measure of central tendency to calculate on all the separate samples.
 - B. The mean of the sampling distribution of the mean is always equal to μ .
 - C. The sampling distribution of μ is always close to the sample means.
 - D. Infinite sampling from a defined population is always more accurate than random sampling.

Answer

B. The mean of the sampling distribution of the mean is always equal to μ .

Question 7

- When comparing the size of the standard error of the mean with the size of the population standard deviation:
 - A. The standard error of the mean is always smaller.
 - B. The standard error of the mean is always larger.
 - C. The standard error is sometimes smaller and sometimes larger depending on sample size.
 - D. The standard error is only smaller than the population standard deviation when the population follows a Poisson distribution.

Answer

- The standard error of the mean is always smaller.

Question 8

■ When graphing the sampling distribution of the mean, the x-axis represents the:

- A. Sample size
- B. Population parameter
- C. Value of the sample mean
- D. Frequency occurrence

Answer

C. Value of the sample mean

Question 9

■ A statistics text presents the following sampling distribution:

■ Which of the following graphs is most likely to represent the population from which it was drawn?

A.



B.



C.



D. Can't be determined because the central limit theorem applies to any population distribution.

Answer

D. Can't be determined because the central limit theorem applies to any population distribution.

Key Ideas of it



- What are the key ideas of the Central Limit Theorem?