## The Central Limit Theorem Section 6-5

- Objectives:
  - Use the central limit theorem to solve problems involving sample means for large samples

## Introduction

- Statisticians are interested in knowing:
  - How individual data values vary about the mean for a population (Section 6.4)
  - How the means of samples of the same size taken from the same population vary about the population mean (Section 6.5)

- Extremely important since it forms the foundation for estimating population parameters and hypothesis
- Tells us that if the sample size is large enough, the distribution of the sample means can be approximated by a normal distribution, even if the original population is not normally distributed.

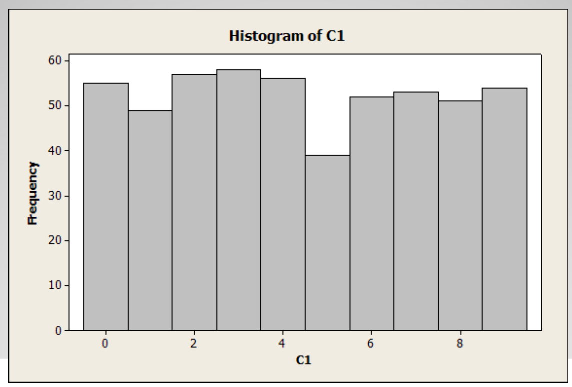
## Central Limit Theorem

- As the sample size n increases without limit, the shape of the distribution of the sample means taken with replacement from a population with mean, $\mu$, and standard deviation, $\sigma$ will approach a normal distribution. This "new" distribution will have a mean, $\mu_{\bar{x}}$ and an **adjusted** standard deviation,

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

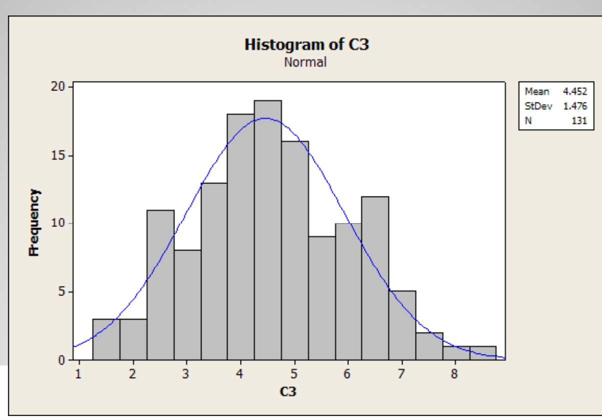
## Central Limit Theorem

- The last four digits of your Social Security number are supposedly assigned at random
- Collect last four digits from 131 individuals
- If we considered each digit individually (524 digits), the distribution appears to be….. and the mean is 4.452 and standard deviation is 2.897




- However, if we consider the last four digits as a group and calculate the mean of each group (131 groups of 4), then the distribution appears to be ….. and the mean is

  4.452 and standard deviation is 1.476



## Guidelines for applying CLT

- When the original variable is normally distributed, the distribution of the sample means will automatically be normally distributed for any sample size n.
- When the distribution of the original variable might not be normal, a sample size of 30 or more is needed to use a normal distribution to approximate the distribution of the sample means. (The larger the sample, the better the approximation will be)

## Example

- Engineers must consider the breadths of male heads when designing motorcycle helmets. Men have head breadths that are normally distributed with a mean of 6.0 inches and a standard deviation of 1.0 inch
- If ONE male is randomly selected, find the probability that his head breadth is less than 6.2 inches
- The Safeguard Helmet company plans an initial production run of 100 helmets Find the probability that 100 randomly selected mean have a head breadth less than 6.2 inches

- The serum cholesterol levels in men aged 18-24 are normally distributed with a mean of 178.1 and a standard deviation of 40.7 (units are in mg/100 mL and the data are based on National Health Survey)
  - If one man aged 18-24 is randomly selected, find the probability that his serum cholesterol is greater than 260, a value considered "moderately high"
  - The Providence Health Maintenance Organization wants to establish a criterion for recommending dietary changes if cholesterol levels are in the top 3%. What is the cutoff for men aged 18-24?
  - If 9 men aged 18—24 are randomly selected, find the probability that their mean serum cholesterol level is between 170 and 200.

## Example