Ch 5.5 Central Limit Theorem

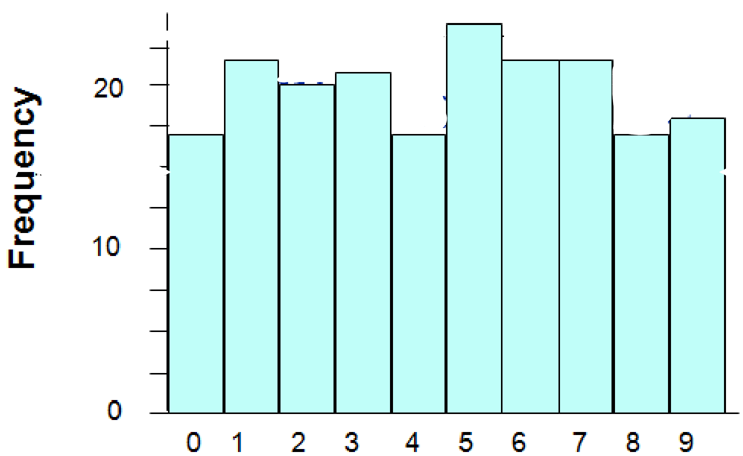
Definition

Sampling Distribution of the mean

the probability distribution of sample means, with all samples having the same sample size n.

Distribution of 200 digits from Social Security Numbers

(Last 4 digits from 50 students)

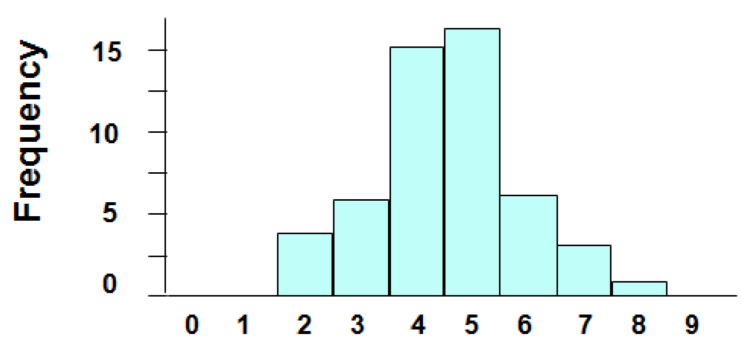


Distribution of 200 digits

Average of the #'s

SSN digits				$\bar{x}$
1	8	6	4	4.75
5	3	3	6	4.25
9	8	8	8	8.25
5	1	2	5	3.25
9	3	3	5	5.00
4	2	6	2	3.50
7	7	1	6	5.25
9	1	5	4	4.75
5	3	3	9	5.00
	3	4	[illegible]	
6	7	7	1	5.25
2	3	3	9	4.25
2	4	7	5	4.50
5	4	3	7	4.75
0	4	3	8	3.75
2	5	8	6	5.25
7	1	3	4	3.75
8	3	7	0	4.50
5	6	6	7	6.00

Distribution of 50 Sample Means for 50 Students



As the sample size increases, the sampling distribution of sample means approaches a normal distribution.

Central Limit Theorem

Given:

1. The random variable x has a distribution (which may or may not be normal) with mean μ and standard deviation σ.
2. Samples all of the same size n are randomly selected from the population of x values.

Central Limit Theorem

Conclusions:

1. The distribution of sample $\bar{x}$ will, as the sample size increases, approach a *normal* distribution.
2. The mean of the sample means will be the population mean μ.
3. The standard deviation of the sample means will approach $\sigma/\sqrt{n}$.

Practical Rules Commonly Used:

1. For samples of size n larger than 30, the distribution of the sample means can be approximated reasonably well by a normal distribution. The approximation gets better as the sample size n becomes larger.

2. If the original population is itself normally distributed, then the sample means will be normally distributed for <u>any sample size</u> n (not just the values of n larger than 30).

Notation

the mean of the sample means

$$\mu_{\bar{x}} = \mu$$

the standard deviation of sample mean

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

(often called standard error of the mean)

$$\text{z score} = \frac{x - \mu_{\bar{x}}}{\frac{\sigma}{\sqrt{n}}}$$

Steps for finding probabilities:

1. Is it a normal distribution? If not is the sample greater than 30, then use the central limit theorem.

2. Find the **z score** $= \dfrac{x - \mu_{\bar{x}}}{\frac{\sigma}{\sqrt{n}}}$ ← gives you the new standard deviation

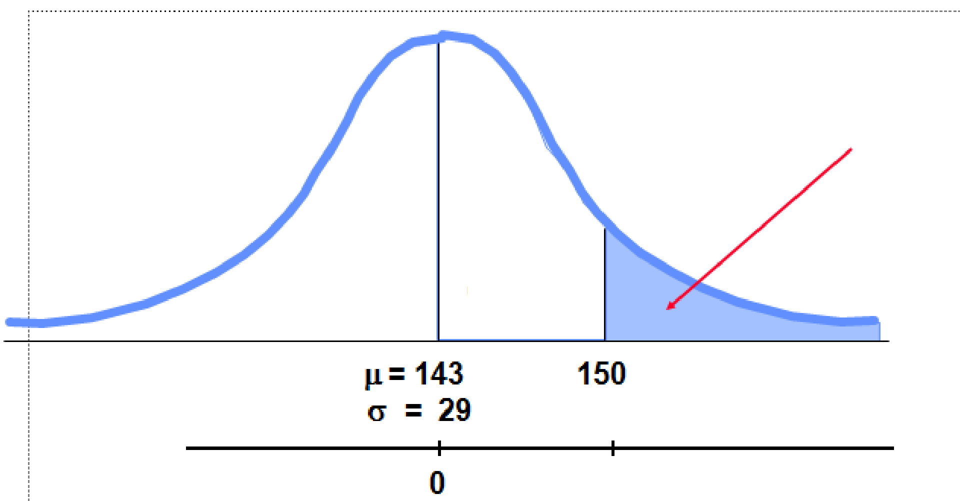
3. Then find the probability that corresponds with the z-score.

4. Determine if need to add or subtract,etc.

5. If asks about being unusual use $p < .05$

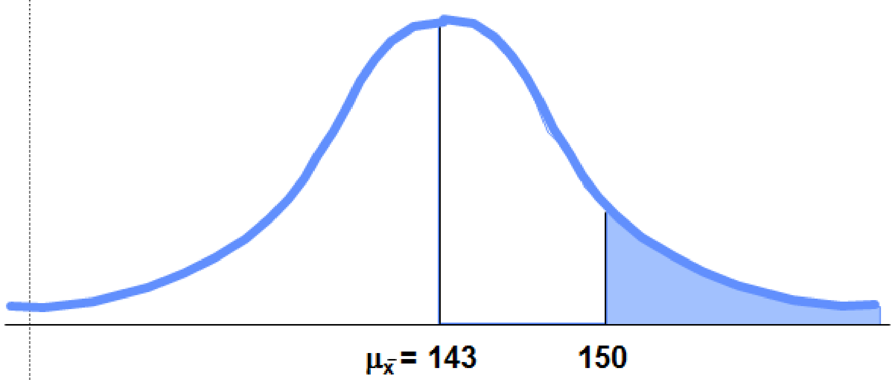
Ex1 **Example:** Given the population of women has normally distributed weights with a mean of 143 lb and a standard deviation of 29 lb,

a.) if one woman is randomly selected, the probability that her weight is greater than 150 lb. is ?



1b) **Example:** **Given the population of women has normally distributed weights with a mean of 143 lb and a standard deviation of 29 lb,**

b.) if 36 different women are randomly selected, find the probability that their mean weight is greater than 150 lb.



$\mu_{\bar{x}} = 143$

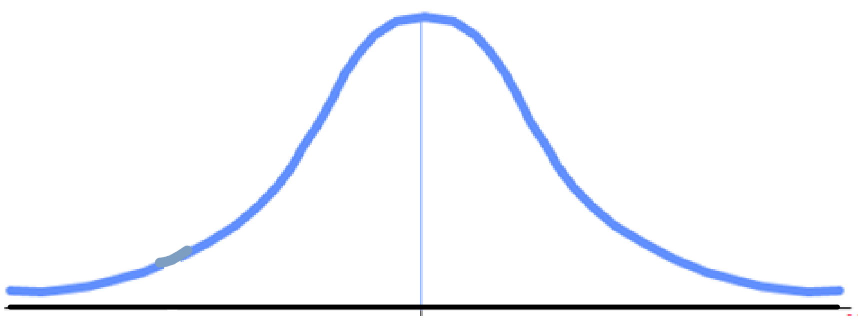
$\sigma_{\bar{x}} = \frac{29}{\sqrt{36}} = 4.83333$

$$z = \frac{150-143}{\frac{29}{\sqrt{36}}} = 1.45$$

$0.5 - 0.4265 = 0.0735$

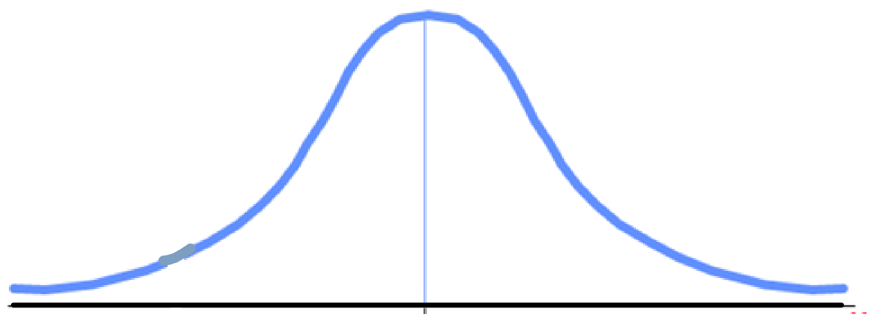
1c) **Example:** **Given the population of women has normally distributed weights with a mean of 143 lb and a standard deviation of 29 lb,**

if 15 **different women are randomly selected, find the probability that their mean weight is** less than 130?



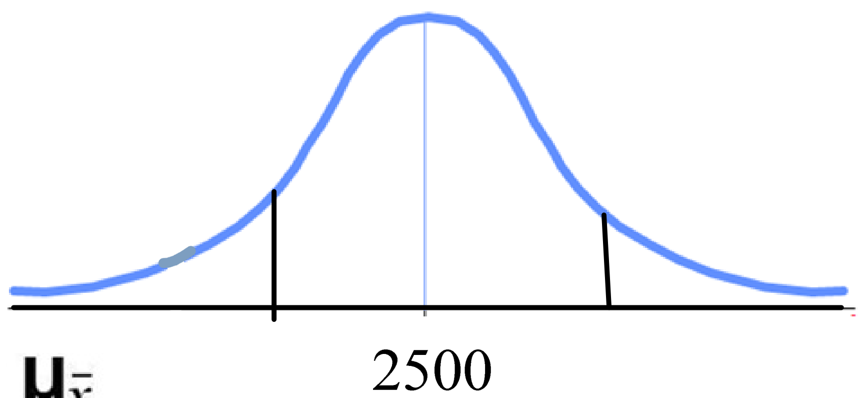
Is it unusual to weigh less than 130 pounds for a women?

Ex2a) In December in Mesa Verde National Park the average weight for a female deer is 63.0 kg and σ of 7.1 kg. If a doe weighs less than 54 kg, then she is considered to be undernourished. What is the probability that a single doe captured in December is undernourished?



2b)If the park has about 2200 does, what number do you expect to be undernourished?

Ex3a) A certain strain of bacteria occurs in all raw milk. The health dept. has found that it has a normal distribution with a μ of 2500 and σ = 300 count per ml of milk. In a sample of 42 samples of milk, what is the probabiity that average bacteria count for one day is between 2350 and 2650 per ml?



z score = $\dfrac{x - \mu_{\bar{x}}}{\frac{\sigma}{\sqrt{n}}}$

Ex3b) If you were the inspector and noticed the milk was not between 2350 and 2650, would you suspect something?

1 The lengths of pregnancies are normally distributed with a mean of 268 days and a standard deviation of 15 days. If 36 women are randomly selected, find the probability of having mean pregnancy between 268 days and 270 days?

A 0.2881

B 0.7881

C 0.2119

D 0.5517