We'll start today with a video of Pardis Sabeti 00: to 07:29

Dr. Sabeti is a professor at the Center for Systems Biology and Department of Organismic and Evolutionary Biology at Harvard University and a professor in the Department of Immunology and Infectious Disease at the Harvard School of Public Health.

Today: 7.3 Day 2

Tomorrow: Review

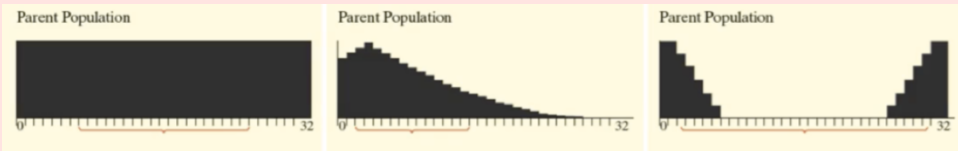
Wednesday: Test on Ch. 7

Get a laptop. Log in. Open up a browser.

Most Population Distributions are not Normal.

So, then the question becomes:

"What is the shape of the sampling distribution of the sample means, $\bar{x}$, from a non-Normal population."



You will need to open up an applet:

Google: **online statbook sampling**

Sampling Distributions - OnlineStatBook
onlinestatbook.com/stat_sim/sampling_dist/
This simulation lets you explore various aspects of **sampling** distributions. When the simulation begins, a histogram of a normal distribution is displayed at the ...

Notes on 7.3 Day 2 **December 10, 2018**

VS

The ACT test is scored with whole numbers from 0 to 36. We will use the website www.tinyurl.com/EKstats66 to take samples of ACT scores from SHS and Rockford HS.

Click "Begin" and you will see the population distribution of fake ACT scores from SHS.

1. Describe the shape, center, and variability of the *population distribution* of ACT scores for SHS.

   Approx. $\mu =$ $\sigma =$

2. Click "Animated" to take a sample of 5 ACT scores. Look at the grey 5 grey boxes. Estimate

   and List the 5 the scores here:________________ Estimate their sample Mean (blue box):____

VS

The ACT test is scored with whole numbers from 0 to 36. We will use the website www.tinyurl.com/EKstats66 to take samples of ACT scores from SHS and Rockford HS.

Click "Begin" and you will see the population distribution of fake ACT scores from SHS.

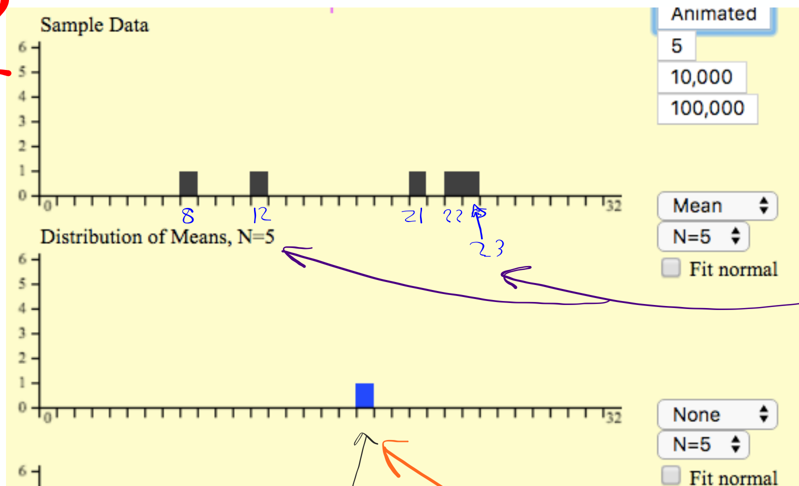
1. Describe the shape, center, and variability of the *population distribution* of ACT scores for SHS.

   Approx. Normal $\mu = 16$ $\sigma = 5$

2. Click "Animated" to take a sample of 5 ACT scores. Look at the grey 5 grey boxes. Estimate

   and List the 5 the scores here:________________ Estimate their sample Mean (blue box):____

2

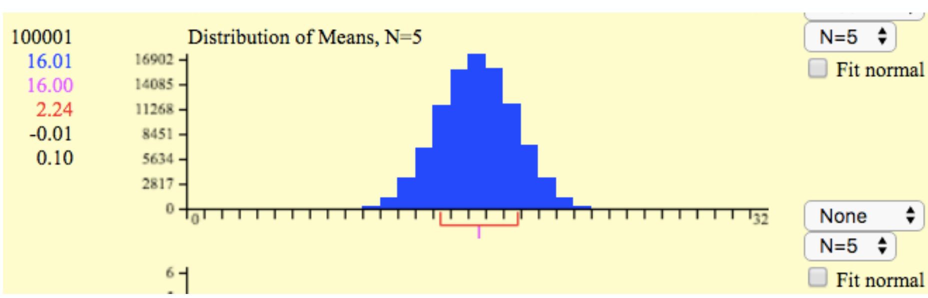


Estimated ACT Scores

17.2

Average of sample $\bar{x}$

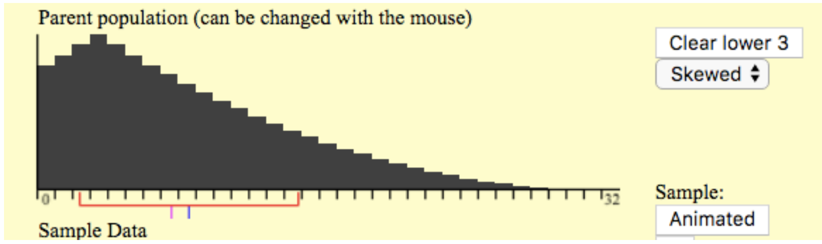
3.



from last class

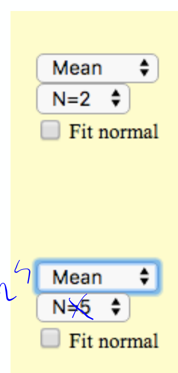
Since the population distribution is apporox. Normal, the Sampling Distribution (of the means) is also approx Normal.

4.



- Skewed Right
- More students had lower ACT scores and fewer did well

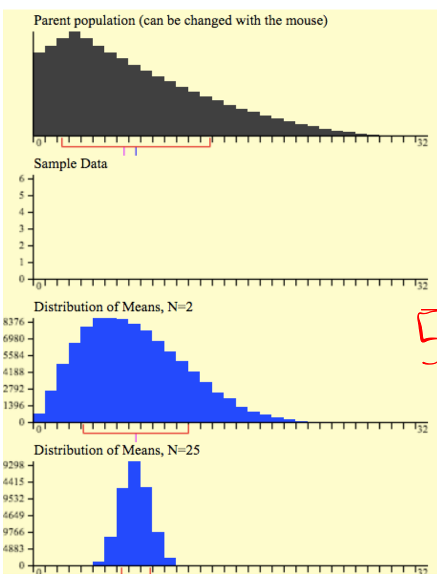
5.



N=25

Change both of the bottom two dropdown menus to "Mean". The first one should be "N=2" and the second one should be "N=25". The click "10,000" to take 10,000 samples.

5. Describe the shape of the sampling distribution of $\bar{x}$ when N = 2.



5. Skewed right, but not as strongly skewed as the population.

6. Approx. Normal

Large Sample Size

$n \geq ?$

The Central Limit Theorem

Draw an SRS of size n from any population with mean μ and finite standard deviation σ. The **central limit theorem (CLT)** says that when n is large, the sampling distribution of the sample mean $\bar{x}$ is approximately Normal.

Shape of the Sampling Distribution of the Sample Mean $\bar{x}$

- **If the population distribution is Normal, the sampling distribution of $\bar{x}$ will also be Normal, no matter what the sample size n is**

The Central Limit Theorem

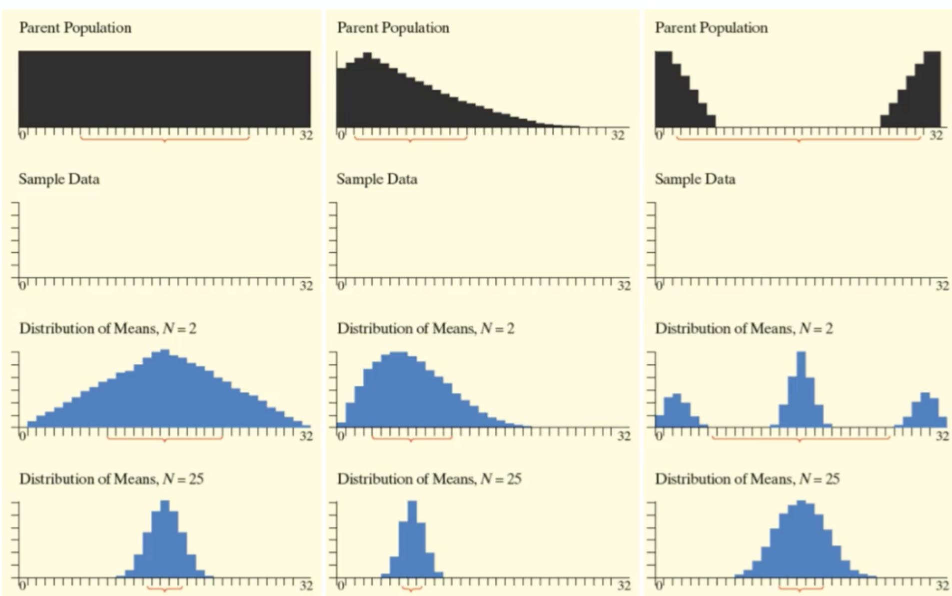
Draw an SRS of size n from any population with mean μ and finite standard deviation σ. The **central limit theorem (CLT)** says that when n is large, the sampling distribution of the sample mean $\bar{x}$ is approximately Normal.

Shape of the Sampling Distribution of the Sample Mean $\bar{x}$

- **If the population distribution is Normal, the sampling distribution of $\bar{x}$ will also be Normal, no matter what the sample size n is**
- **If the population distribution is not Normal, the sampling distribution of $\bar{x}$ will be approximately Normal when the sample size is large ($n \geq 30$ in most cases). If the sample size is small and the population distribution is not Normal, the sampling distribution of x will retain some characteristics of the population distribution (e.g., skewness).**

Now go back to the APP and try to create an **abnormal** Population Distribution

1. Select Custom (instead of skewed)
2. Click on a point on the population graph to insert a bar of that height OR click on a point on the horizontal axis, and drag up to define a bar. You can also shorten bars.
3. Make a distribution that looks as strange as you can.
4. Now go back and take 100,000 samples.
5. What do you notice about the sampling distributions?



The Central Limit Theorem

Important ideas:

The Central Limit Theorem

Important ideas:

CENTRAL LIMIT Theorem

The Sampling Distribution of the mean ($\bar{x}$) will be appox. Normal if the sample size is large ($n \geq 30$)

CLT is only about the shape of the distribution of the sample mean.

So ..... don't refer to it as an explanation of how the variability of a sampling distribution decreases as the sample size goes up.

The Central Limit Theorem

Important ideas:	Probability
CENTRAL LIMIT Theorem The Sampling Distribution of the mean ($\bar{x}$) will be appox. Normal if the sample size is large ($n \geq 30$)	If appox. Normal, then z-score of sampling distr. $Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$

not for Sampling Distrib. of proportions

Auto Care-Center

Auto Care Center: Keith is the manager of an auto-care center. Based on service records of 3500 customers from the past year, the time (in hours), that a technician requires to complete a standard oil change and inspection follows a right-skewed distribution with μ = 30 minutes and σ = 20 minutes. For a promotion, Keith randomly selects 40 current customers and offers them a free oil change and inspection if they redeem the offer during the next month. Keith budgets an average of 35 minutes per customer for a technician to complete the work. Will this be enough?

(a) Describe the shape of the sampling distribution of $\bar{x}$ for samples of 40 randomly selected customers. Justify your answer.

(b) Find the mean and standard deviation of the sampling distribution of $\bar{x}$. Be sure to check the 10% condition. (this is a good time to point out that we are talking about sampling distribution of the mean, not a population --- formulas are not the same!)

Auto Care Center: Keith is the manager of an auto-care center. Based on service records of 3500 customers from the past year, the time (in hours), that a technician requires to complete a standard oil change and inspection follows a right-skewed distribution with μ = 30 minutes and σ = 20 minutes. For a promotion, Keith randomly selects 40 current customers and offers them a free oil change and inspection if they redeem the offer during the next month. Keith budgets an average of 35 minutes per customer for a technician to complete the work. Will this be enough?

sample means

(a) Describe the shape of the sampling distribution of $\bar{x}$ for samples of 40 randomly selected customers. Justify your answer.

Approx. Normal by Central Limit Theorem since $40 \geq 30$

(b) Find the mean and standard deviation of the sampling distribution of $\bar{x}$. Be sure to check the 10% condition. (this is a good time to point out that we are talking about sampling distribution of the mean, not a population --- formulas are not the same!)

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

$\sigma_{\bar{x}} =$

Auto Care Center: Keith is the manager of an auto-care center. Based on service records of 3500 customers from the past year, the time (in hours), that a technician requires to complete a standard oil change and inspection follows a right-skewed distribution with μ = 30 minutes and σ = 20 minutes. For a promotion, Keith randomly selects 40 current customers and offers them a free oil change and inspection if they redeem the offer during the next month. Keith budgets an average of 35 minutes per customer for a technician to complete the work. Will this be enough?

(a) Describe the shape of the sampling distribution of $\bar{x}$ for samples of 40 randomly selected customers. Justify your answer.

Approx. Normal by Central Limit Theorem since $40 \geq 30$

(b) Find the mean and standard deviation of the sampling distribution of $\bar{x}$. Be sure to check the 10% condition. (this is a good time to point out that we are talking about sampling distribution of the mean, not a population --- formulas are not the same!)

$\mu_{\bar{x}} = \mu = 30$ min.

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{20}{\sqrt{40}} = 3.16$

10% condition

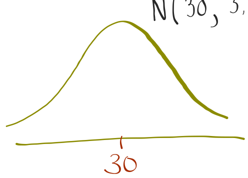
$40 < \frac{1}{10}(3500)$

(c) Calculate the probability that the average time it takes to complete the work exceeds 35 minutes.

(d) How much average time per customer should Keith budget if he wants to be 99% certain that he doesn't go "over budget"?

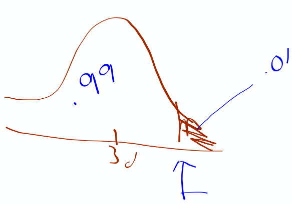
(c) Calculate the probability that the average time it takes to complete the work exceeds 35 minutes.

N(30, 3.16)

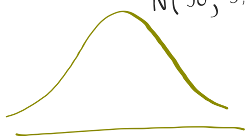


$$z = \frac{35-30}{3.16} = 1.58 \xrightarrow{\text{TABLE A}} \boxed{.0568}$$

(d) How much average time per customer should Keith budget if he wants to be 99% certain that he doesn't go "over budget"?



(c) Calculate the probability that the average time it takes to complete the work exceeds 35 minutes.

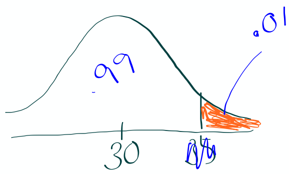


$N(30, 3.16)$

$z = \frac{35-30}{3.16}$

$= 1.58$ $\xrightarrow{\text{TABLE A}}$ $\boxed{.0568}$

(d) How much average time per customer should Keith budget if he wants to be 99% certain that he doesn't go "over budget"?



prob to the left

.99

invNorm(.99, 30, 3.16)

area to left

μ σ

~~= .0568~~

37.4 min

Pardis Sabeti

07:29 to....

7.3.......63-73 (odds) , and 74-76

study pp. 474-478