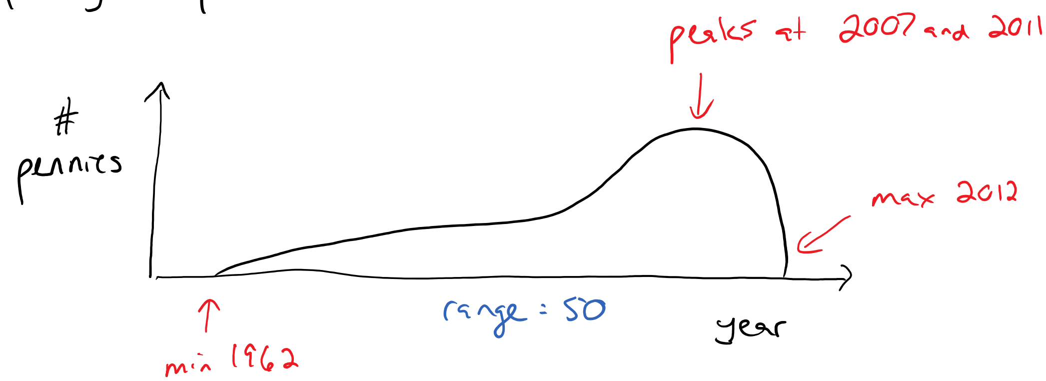
Section 5.5: The Central Limit Theorem

Thursday, November 21, 2019 9:53 AM

penny experiment



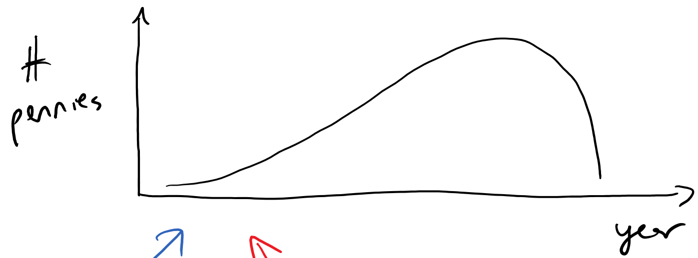
range: 19

what did we learn from this experiment?

- there is less variation (spread) in the graph of the average of 10 points than in the graph of the individual points

what do you think would happen if we were to graph also the average date for 25 pennies?

- it would be even more tightly clustered than the previous graph why is this?



the individual penny grap

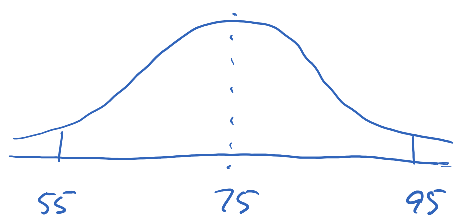
picking a penny with a date down here is somewhat unlikely

picking 10 pennies whose average is down here is <u>extremely</u> unlikely

whenever you take a sample you may in fact get a few outliers but most of your points will be from the "middle" of your graph (where most of the points are from) and so your average values will cluster towards the middle of the graph

consider how students do on a math test:

test scores will likely range from 55 to 95 with perhaps a few outliers



mean = 75
std dev = 10

how likely is it that an individual student will get > 95% ?

if the test scores are mound-shaped
Empirical Rule says that ~95% of measurements will lie within 2 std dev of mean, leaving ~5% outside

since the graph looks symmetrical, can say ~2.5% of the measurements will lie above the test score of 95

now consider the class average for a class of 40:

what is the probability that a class average is above 95%? Is it still ~2.5%?

No! this distribution is much narrower, so above 95 is way more than 2 std dev above the mean

but how much narrower?

turns out that it depends on the sample size

$$\sigma_{group} = \frac{\sigma_{individuals}}{\sqrt{n}}$$

Central Limit Theorem

given: a random variable x (which can have

any distribution – skewed, multimodal, symmetrical)
with mean μ and standard deviation σ

select samples of size n from this population – such that all possible samples of size n have the same chance of being selected

then: the distribution of sample means will as the sample size increases approach a normal distribution

with population mean $\mu_{group} = \mu_{individual}$

and standard deviation $\sigma_{group} = \dfrac{\sigma_{individual}}{\sqrt{n}}$

practical considerations:

for samples with $n \geq 30$, the distribution of the sample means can be approximated well by a normal distribution

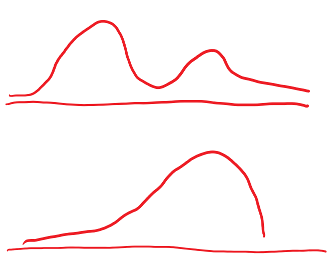
note also: if the original population is normal to begin with, then the sample means will be normally distributed for <u>any</u> sample size.

---

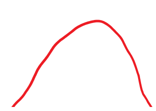
summary:

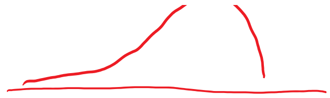
if individual population is non-normal

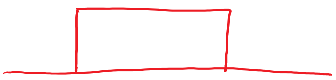
distribution for group



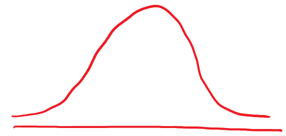
large sample $\xrightarrow[n \geq 30]{}$



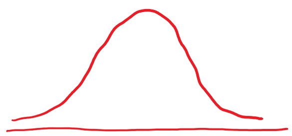




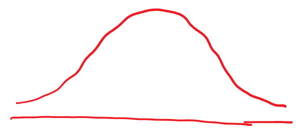
$\xrightarrow[n \geq 30]{}$



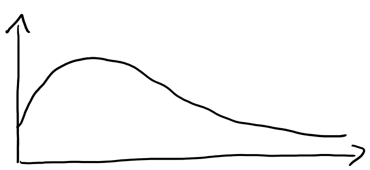
if individual population is normal to begin with



$\xrightarrow[\text{sample size}]{\underline{\text{any}}}$



example: The time students wait for the bus has the distribution shown below with mean 4.5 minutes and standard deviation 4.0 minutes



a) describe the shape and symmetry of this distribution

unimodal, skewed right

b) The wait times for 100 students are measured and the average / mean value is calculated. What would the distribution look like for these mean values?

mean of a group? ✓

is $n \geq 30$? ✓

the means will be normally distributed

c) what would be the mean and standard deviation for the distribution in part (b)?

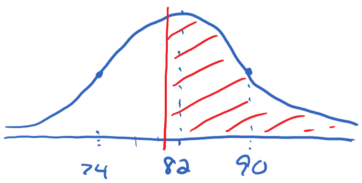
	individual	group of 100
mean	4.5	4.5
std dev	4.0	$\frac{4.0}{\sqrt{100}} = 0.4$

example: suppose that at a large university, the dean of admissions has determined that the scores of a first-year class on a math placement test are normally distributed with a mean of 82 and std dev of 8.

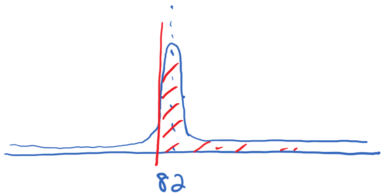
a) what is the probability that any one student drawn at random from the class has a score of at least 80?

b) what is the probability that the mean score of a random group of 64 students is at least 80?

	individual	group size = 64
mean	82	82
std dev	8	$\frac{8}{\sqrt{64}} = 1$
shape	normal because question said so	normal because the group has $n \geq 30$ (CLT)

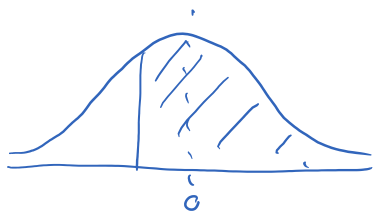


individuals



group of 64

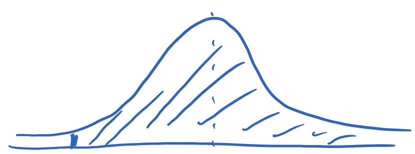
a) $z = \dfrac{x - \mu}{\sigma} = \dfrac{80 - 82}{8} = -0.25$



area = 0.0987

$p = 0.5 + 0.0987 = 0.5987$ or $\boxed{60\%}$

b) $z_{group} = \dfrac{x - \mu_{group}}{\sigma_{group}} = \dfrac{80 - 82}{1} = -2$



area = 0.4772

$p = 0.5 + 0.4772$
$= 0.9772 \approx \boxed{98\%}$

2019/11/27

example: The weight of luggage checked by airline passengers is a random variable with mean 50 lbs and a std dev of 30 lbs. For a plane seating 100 passengers the total luggage limit is an average of 57.5 lbs per passenger.

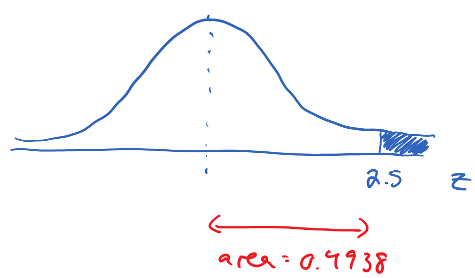
What is the probability that the luggage limit will be exceeded?

answer:

	individual	group of 100
mean	50	50
std dev	30	$\frac{30}{\sqrt{n}} = \frac{30}{\sqrt{100}} = 3$
shape	unknown	normal (because $n \geq 30$)

$$z_{group} = \frac{x - \mu_{group}}{\sigma_{group}} = \frac{57.5 - 50}{3}$$

$$= 2.5$$



$$P = 0.5 - 0.4938$$
$$= 0.0062 \quad \text{or} \quad \boxed{0.6\ \%}$$