

9.3 Sample Means

sample mean has a distribution
with a mean and standard deviation

<http://www.rossmanchance.com/applets/SampleData/SampleData.html>



Draw 100 samples of size 10. Check each mean and standard deviation in the Year graph.

http://www.ruf.rice.edu/~lane/stat_sim/sampling_dist/



Draw 100 samples of size 5, then increase the size each time. Check what happens to the mean and standard deviation as n increases.

large university lecture hall courses
mean number of songs on their iPods

Same mean in every auditorium?

No. We usually expect different samples to yield different statistics.



mean number of songs on their iPods in all groups of 300 students
graph each mean found

probably a bell curve, maybe almost Normal, like these:



sampling distribution of the sample mean:

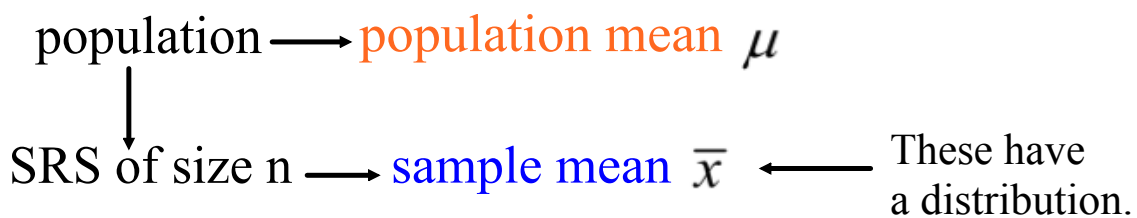
S approx. Normal

O few, if any

C \approx mean number of songs on iPods in population

S since $n=300$, spread's smaller than if n had 100 & larger than if n had been 400

Sampling Distribution of the Sample Mean is the distribution of sample means from all possible samples of this size.



mean of the sampling distribution of \bar{x} is exactly μ .

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

standard deviation of the sampling distribution of \hat{p} is

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

as long as the population is at least 10
times as large as the sample ($N \geq 10n$).

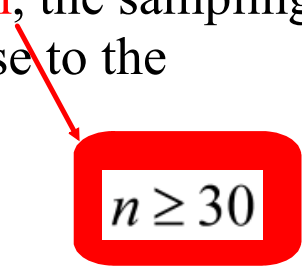
Why $N > 10n$? Because of the methods we use. Larger samples are so much better that our standard deviation formula is no longer a good estimate. If our sample is more than 1/10 the population, the assumptions, conditions, methods, and formulas we have won't work.

Normal population \Rightarrow Normal sampling distribution
of the sample mean

non-Normal population \Rightarrow *nearly* Normal sampling distribution
of the sample mean
if the sample size is at least 30
($n \geq 30$)

Central Limit Theorem

Draw an SRS of size n from **any population whatsoever** with mean μ . When n is **large enough**, the sampling distribution of sample mean \bar{x} is close to the normal distribution $N(\mu, \frac{\sigma}{\sqrt{n}})$.


$$n \geq 30$$

Ex. Consider adults of some snake species that is $N(40,6)$ in cm.

Take random samples of $n = 4$ adult snakes.

- The mean of the sample means =

$$\mu_{\bar{x}} = \mu = 40$$

- Verify n for the standard deviation

I would assume that 4 snakes is $< 10\%$ of the pop.

The standard deviation for the mean length =

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{6}{\sqrt{4}} = \frac{6}{2} = 3$$

- Verify condition for a Normal approximation:

$n = 4 < 30$, but the pop. has a normal dist.

The sampling distribution of the sample mean is approximately

Normal:

$$N(40, 3)$$

Then the 68-95-99.7 Rule says that

- 68% of samples of $n=4$ such adult snakes will have an average length between 37 and 43 cm.
- 95% of samples of $n=4$ such adult snakes will have an average length between 34 and 46 cm.
- 99% of samples of $n=4$ such adult snakes will have an average length between 31 and 49 cm.

It would be rare for a sample of 4 such adult snakes to have an average length over 49 cm or under 31 cm.

Ex. One year, the mean ACT score was 20.8 & the standard deviation was 4.8. Suppose we take a random samples of $n = 36$ scores.

- The mean sample mean ACT score = $\mu_x = \mu = 20.8$

- Verify n for the standard deviation:

36 students represent less than 10% of all students who've taken ACT
The standard deviation for the sample mean ACT score =

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{4.8}{\sqrt{36}} = \frac{4.8}{6} = .8$$

Verify the Condition for Normal Approximation:

$n = 36 \geq 30$ so we can use a normal approx.

The sampling distribution of the sample mean ACT score is approximately Normal:

$$N(20.8, .8)$$

So, by the 68-95-99.7 Rule,

There's a 99.7% chance that samples of $n=36$ ACT scores will have an average between 18.4 & 23.2.

Almost every sample of $n=36$ ACT scores will have a mean between 18.4 & 23.2.

If a sample of $n=36$ ACT scores has a mean more than 23.2, that is unusually high (lower than 18.4 would be unusually low).