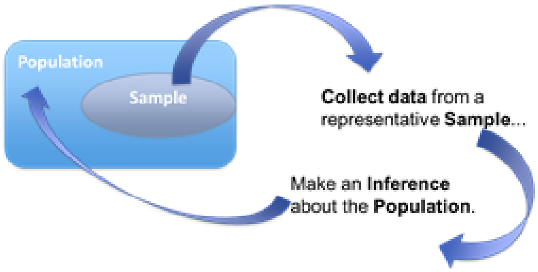
7.1 What Is A Sampling Distribution?

- DISTINGUISH between a parameter and a statistic.
- USE the sampling distribution of a statistic to EVALUATE a claim about a parameter.
- DISTINGUISH among the distribution of a population, the distribution of a sample, and the sampling distribution of a statistic.
- DETERMINE whether or not a statistic is an unbiased estimator of a population parameter.
- DESCRIBE the relationship between sample size and the variability of a statistic.

The process of statistical inference involves using information from a sample to draw conclusions about a wider population.



Parameters and Statistics

A **parameter** is a number that describes some characteristic of the population.
A **statistic** is a number that describes some characteristic of a sample.

	Population Parameter	Sample Statistic
Mean	μ	$\bar{x}$
Standard dev.	σ	s
Proportion	p	$\hat{p}$

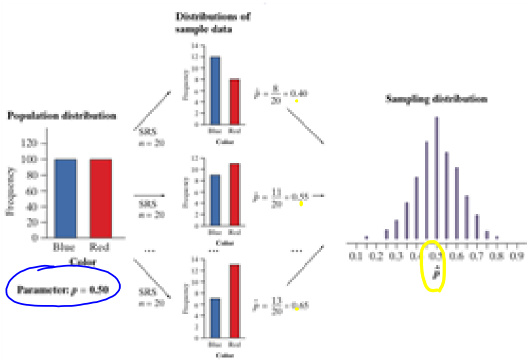
③ Parameter Statistic
$\mu = 2.5003$ $\bar{x} = 2.5009$

Sampling Variability

How can $\bar{x}$ be an accurate estimate of μ? After all, different random samples would produce different values of $\bar{x}$.

This basic fact is called sampling variability : the value of a statistic varies in repeated random sampling.

Sampling Distribution vs. Population Distribution

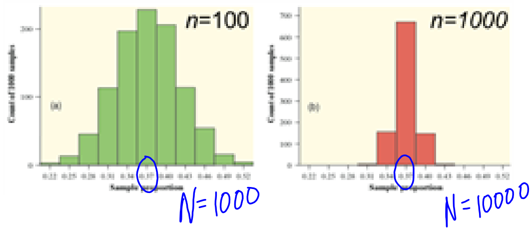


A statistic used to estimate a parameter is an **unbiased estimator** if the mean of its sampling distribution is equal to the true value of the parameter being estimated.

Describing Sampling Distributions

Spread: Low variability is better!

To get a trustworthy estimate of an unknown population parameter, start by using a statistic that's an unbiased estimator. This ensures that you won't tend to overestimate or underestimate.
Unfortunately, using an unbiased estimator doesn't guarantee that the value of your statistic will be close to the actual parameter value.



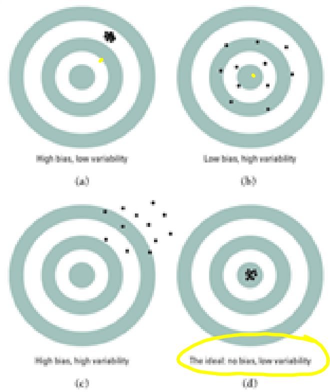
The **variability of a statistic** is described by the spread of its sampling distribution. This spread is determined mainly by the size of the random sample. Larger samples give smaller spreads. The spread of the sampling distribution does not depend much on the size of the population, as long as the population is at least 10 times larger than the sample.

Bias, Variability, and Shape

Both bias and variability describe what happens when we take many shots at the target.

Bias means that our aim is off and we consistently miss the bull's-eye in the same direction. Our sample values do not center on the population value.

High **variability** means that repeated shots are widely scattered on the target. Repeated samples do not give very similar results.

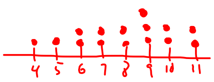


7. {2, 6, 8, 10, 10, 12}

$\mu = 8$ Range = 10

a)

2,6 4	6,8 7	8,10 9	10,10 10	10,12 11
2,8 5	6,10 8	8,10 9	10,12 11	
2,10 6	6,10 8	8,12 10		
2,10 6	6,12 9			
2,12 7				

b)  Slightly skewed left

10. $\mu = 64$ $\sigma = 2.5$

a) $\bar{x} = 62.4$ one of the 250 samples and has a mean of 62.4

b) Normal, No obvious outliers

c) 64.7 or more - Suprising?
Close to 25 dots $25/250 = .10$
So it is not suprising

d) 64.7 Top 10% - Tall

23. 200,000 Sample 250000

$\boxed{n = 10}$

n = 50
n = 100
n = 1000

Close to 200,000