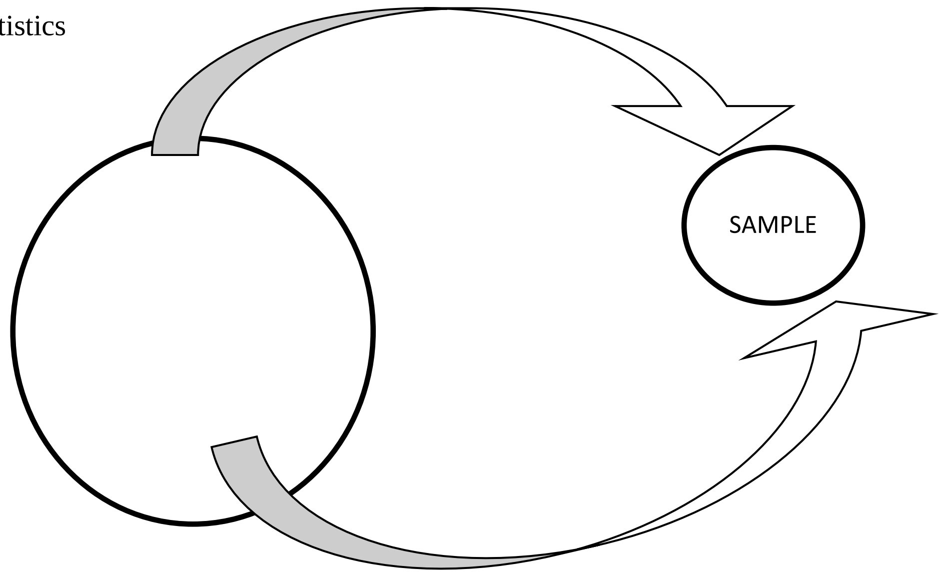
AP Statistics – Chapter 7 Notes §7.1 What is Sampling Distribution?

Parameter vs. Statistics



Identify the population, the parameter, the sample, and the statistic in each of the following settings.

A) A Gallup Poll asked a random sample of 515 US adults whether or not they believe in ghosts. Of the respondents, 160 said "YES".

Population: ____

Parameter: ____

Sample: ____

Statistic: ____

B) During the winter months, the temperatures outside the Murphy's cabin in Colorado can stay well below freezing (32° F, or 0° C) for weeks at a time. To prevent the pipes from freezing, Mr. Murphy sets the thermostat at 50° F. He wants to know how low the temperature actually gets in the cabin. A digital thermometer records the indoor temperature at 20 randomly chosen times during a given day. The minimum reading is 38° F.

Population: ____

Parameter: ____

Sample: ____

Statistic: ____

C) A pediatrician wants to know the 75th percentile for the distribution of heights of 10-year-old boys, so she takes a sample of 50 patients and calculates Q_3 = 56 inches.

Population: ____

Parameter: ____

Sample: ____

Statistic: ____

D) A Pew Research Center Poll asked 1102 12- to 17-year olds in the US if they have a cell phone. Of the respondents, 780 said “Yes”.

Population: ____

Parameter: ____

Sample: ____

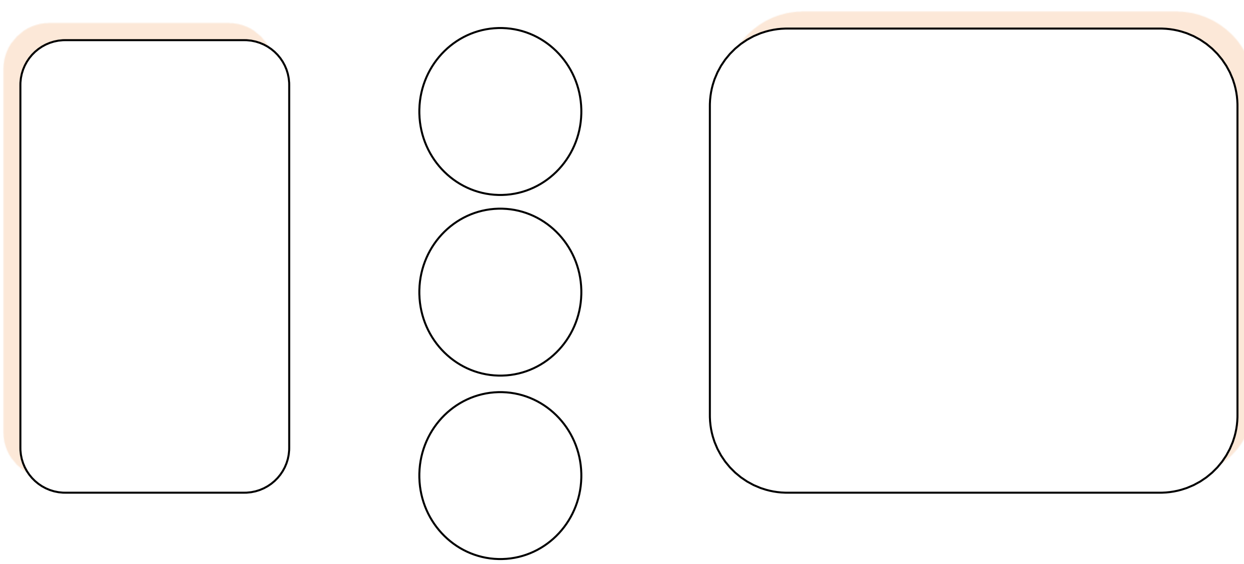
Statistic: ____

CYU: Pg.417

The ____________________________ of a statistic is the distribution of values taken by the statistic in _____ possible samples of the _________________ from the population. (It is the ideal pattern that would emerge if we looked at all possible samples)

**This is usually too difficult so instead we use simulation to imitate the process of taking many, many samples. (SEE Example on pg. 419)

Sampling repeatedly yields 3 distributions: RED or BLUE chips?



_____________________ __________________ ______________________________

The population distribution and the distribution of sample data describe individuals. A sampling distribution describes how a statistic varies in many samples from the population.

CYU: Pg.420

Describing Sampling Distributions:

Center:
A statistic used to estimate a parameter is an ________________________________ if the mean of its sampling distribution is ___________ to the true value of the parameter being estimated.

The value of an unbiased estimator will sometimes exceed the true value of the parameter and sometimes be less if we take many samples. Because its sampling distribution is centered at the true value, however, there is no systematic tendency to overestimate or underestimate the parameter. Hence, it is considered UNBIASED. (No favoritism). In statistics 'unbiased' does not mean perfect!

Unbiased Estimators	Biased Estimators
Sample variance Sample mean Sample proportion	Sample standard deviation Sample range

The ____________________________________ is described by the spread of its sampling distribution. This spread is determined primarily by the size of the random sample. Larger samples give smaller spread. The spread of the sampling distribution does not depend on the size of the population, as long as the population is at least 10 times larger than the sample.

CYU: Pg.426

To obtain a trustworthy estimate of an unknown parameter:

1. Use a statistic that is an unbiased estimator of the parameter.
   (This ensures that you won't tend to overestimate or underestimate)

2. Keep variability low!

3. Use larger samples – they are more likely to produce an estimate close to the true value of the parameter.

4. Consider the shape of the sampling distribution before doing inference.

See Fig. 7.8 Pg. 426

AP Statistics – Chapter 7 Notes §7.2 Sampling Proportions

Sampling Activity: Conclusions:

SHAPE: The sampling distribution of $\hat{p}$ can be approximated by a Normal curve. This depends on both the sample size n and the population proportion p.

CENTER: The mean of the distribution is ____________. This makes sense because the sample proportion $\hat{p}$ is an unbiased estimator of p.

SPREAD: For a specific value of p, the standard deviation ________ gets smaller as n gets larger. The value of ________ depends on both n and p.

**The standard deviation of the $\hat{p}$ distribution measures how far the sample proportions will be from the true proportion, on average, in repeated random samples of a particular size.

Sampling Distribution of a Sample Proportion
Choose an SRS or size n from a population of size N with proportion p of successes. Then: • The mean of the sampling distribution of $\hat{p}$ is $\mu_{\hat{p}} = p$ • The standard deviation of the sampling distribution of $\hat{p}$ is: As long as the 10% condition is satisfied: $n \leq \frac{1}{10}N$ • As n increases, the sampling distribution of $\hat{p}$ becomes APPROXIMATELY NORMAL. Before you perform Normal calculations, check that the Normal Condition is satisfied: $np \geq 10 \ \ and \ \ n(1-p) \geq 10$

CYU: Pg. 437

Example: Going to college: (Normal Calculations involving $\hat{p}$)

A polling organization asks an SRS of 1500 first-year college students how far away their home is. Suppose that 35% of all first-year students actually attend college within 50 miles of home. What is the probability that the random sample of 1500 students will give a result within 2 percentage points of this true value? (Follow the 4-step process)

STATE:

PLAN:

DO:

CONCLUDE:

(Because 90% of all SRSs will give a sample proportion within 2 percentage points of the truth, then in 90% of the SRSs, the true proportion will be within 2 percentage points of the sample proportion)

Your Turn:

The superintendent of a large school district wants to know what proportion of middle school students in her district are planning to attend a four-year college or university. Suppose that 80% of all middle school students in her district are planning to attend a four-year college or university. What is the probability that an SRS of size 125 will give a result within 7 percentage points of the true value?

Mean and Standard Deviation of the Sampling Distribution of $\bar{x}$
Suppose that ____ is the mean of an SRS of size n drawn from a large population with mean _____ and standard deviation ______. Then: - -

Example: Moviegoing Students

Suppose that the number of movies viewed in the last year by high school students has an average of 19.3 with a standard deviation of 15.8. Suppose we take an SRS of 100 high school students and calculate the mean number of movies viewed by the members of the sample.

a) What is the mean of the sampling distribution of $\bar{x}$?

b) What is the standard deviation of the sampling distribution of $\bar{x}$? Check whether the 10% condition is satisfied.

Facts about the mean and standard deviation of $\bar{x}$: (These are true no matter what shape the population distribution has)

- The sample mean $\bar{x}$ is an unbiased estimator of the population mean ___.
- The values of $\bar{x}$ are less spread out for larger samples.
- The standard deviation decreases at the rate $\sqrt{n}$, so you must take a sample four times as large to cut the standard deviation of $\bar{x}$ in half.
- Use the formula ____________ ONLY when the population is at least 10 times as large as the sample (the _______ condition).

Exploring the Sampling Distribution Applet

Example: Young Women's Heights

The height of young women follows a Normal distribution with mean $\mu = 64.5$ inches and standard deviation $\sigma = 2.5$ inches.

a) Find the probability that a randomly selected young woman is taller than 66.5 inches. Show your work.

b) Find the probability that the mean height of an SRS of 10 women exceeds 66.5 inches. Show your work.

Suppose that the population is Normally distributed with mean μ and standard deviation σ. Then the sampling distribution of _____ has the Normal distribution with mean ____ and standard deviation _______, provided that the 10% condition is met.

CYU: Pg.448

What if the population distribution is not Normal?

CENTRAL LIMIT THEOREM (CLT)
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.

See Ex: A Strange Population Distribution (Pg. 450)

Conclusions:

NORMAL CONDITION FOR SAMPLE MEANS:

- If the population distribution is Normal, then so is the sampling distribution of $\bar{x}$. This is true no matter what the sample size n is.
- If the population distribution is not Normal, the central limit theorem tells us that the sampling distribution of $\bar{x}$ will be approximately Normal in most cases if $n \geq 30$.

See Example: Servicing Air Conditioners (Calculations using the CLT) Pg.451

Example: Mean Texts

Suppose that the number of texts sent during a typical day by a randomly selected high school student follows a right-skewed distribution with a mean of 15 and a standard deviation of 35. Assuming that students at your school are typical texters, how likely is it that a random sample of 50 students will have sent more than a total of 1000 texts in the last 24 hours? (SPDC)