# Math 361

*Day 9*

The Binomial Test

Inv. 1.3 – Tim or Bob?

# Review: Do I have ESP?

Let's say I got 10 cards correct in 25 tries.

If we assume that playing the game is a Binomial Random Process with X = # correct, then we can compute

$$P(X \geq 10) = P(X=10) + P(X=11) + \ldots + P(X=25) = 0.017$$

using the formula for P(X=k) with $\pi=0.2$ and n=25

*Assuming I am just guessing, you'd expect to see me get 10 or more matches in 25 cards only 1.7% of the time…*

# Do I have ESP?

*Assuming I am just guessing, you'd expect to see me get 10 or more matches in 25 cards only 1.7% of the time…*

…so there's evidence **against** the null model that says I'm just guessing!

# Learning Objectives

1. Distinguish between the **population** and **sample** associated with a study.

2. Distinguish between the **parameter** and **statisitic** associated with a study.

3. Write **null** and **alternative hypotheses** for a research question.

4. Interpret a **p-value** in context

5. List the steps of a **Binomial Test**

# Inv. 1.3: Binomial Test

n = 26

Who is on the left,

Bob

or

Tim?

tim = 21, not-tim = 5

No discussion of your responses!

# Results of Tim/Bob Survey for (c)

Of 26 students, 21 chose Tim as on the left

# Inv. 1.3 – group work

Try parts (a), (b), (c), (d), (e)

# Terminology Detour – Tim/Bob Survey

- Observational Units? students
- Variable? tim on left?
- Type of variable? Binary
- Parameter vs. Statistic

(part f)

# Terminology Detour – Tim/Bob Survey

## Parameter vs. Statistic

- A numerical summary from a random process or *population.*
- Usually notated with Greek letters, e.g. σ, μ, π,…
- Example: π is the probability a student randomly picks Tim

Population – the entire group of interest

- A numerical summary from a *sample*.
- Usually notated with accented Latin letters, e.g. $\bar{x}$, $\hat{p}$, $s$
- Example: $\hat{p}$ is the proportion of the class who pick Tim

Sample - a subset of the population whose data was recorded

1. Distinguish between the **population** and **sample** associated with a study.

2. Distinguish between the **parameter** and **statisitic** associated with a study.

# Practice

Example 1

Do voters make judgments about political candidates based on his/her facial appearance?

Researchers investigated this question in a study published in Science (Todorov, Mandisodka, Goren, & Hall, 2005).

Participants were shown pictures of two candidates and asked who has the more competent looking face. Researchers then predicted the winner to be the candidate whose face was judged to look more competent by most of the participants. For the 32 U.S. Senate races in 2004, this method predicted the winner correctly in 23 of them.

Population:

all elections

Sample:

32 elections

Parameter:

$\pi$ = proportion of elections won by "compentent looking" person

Statistic:

$\hat{p} = 23/32$

# More Practice

Example 2

A statistics class at Cal Poly collected data on a well-known campus legend.

Each student was asked to specify one of the four tires to answer in a situation where you have to make up which tire had recently been flat on your car. The prior conjecture is that a higher number than would be expected due to chance alone would pick the right front tire.

In this class, 24 of 54 students in class chose the right front tire (a tire identified in advance as being one that people tend to pick out of the four).

Population:

Sample:

Parameter:

Statistic:

# Inv. 1.3 – parts (g) and (h)

**Question:** Do most people identify "Tim" as the guy on the left?

**Parameter:** let π be the proportion of people who pick "Tim" to be on the left.

| Null Hypothesis: | Alternative Hypothesis: |
| --- | --- |
| $\pi = 0.5$ | $\pi > 0.5$ |
| i.e. people are just guessing. | i.e. the majority of people pick Tim to be on the left. |

3. Write **null** and **alternative hypotheses** for a research question.

# Practice

Example 1

Do voters make judgments about political candidates based on his/her facial appearance?

Researchers investigated this question in a study published in Science (Todorov, Mandisodka, Goren, & Hall, 2005).

Participants were shown pictures of two candidates and asked who has the more competent looking face. Researchers then predicted the winner to be the candidate whose face was judged to look more competent by most of the participants. For the 32 U.S. Senate races in 2004, this method predicted the winner correctly in 23 of them.

Parameter of interest:

Null Hypothesis:

Alternative Hypothesis:

# More Practice

Example 2

A statistics class at Cal Poly collected data on a well-known campus legend.

Each student was asked to specify one of the four tires to answer in a situation where you have to make up which tire had recently been flat on your car. The prior conjecture is that a higher number than would be expected due to chance alone would pick the right front tire.

In this class, 24 of 54 students in class chose the right front tire (a tire identified in advance as being one that people tend to pick out of the four).

Parameter of interest:

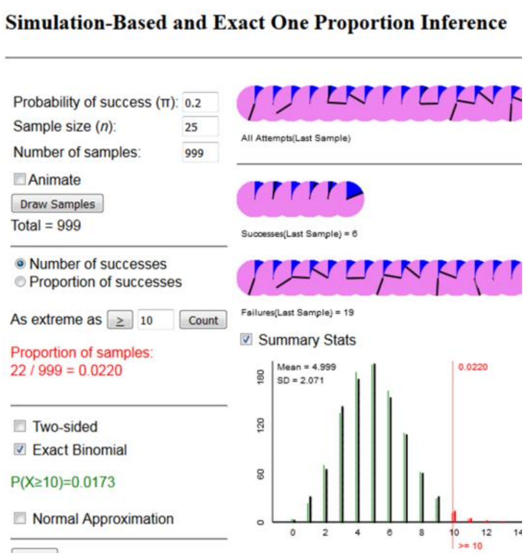
Null Hypothesis:

Alternative Hypothesis:

# Definition of P-value

*In order to decide whether we believe the null hypothesis, we'll compute a **p-value**.*

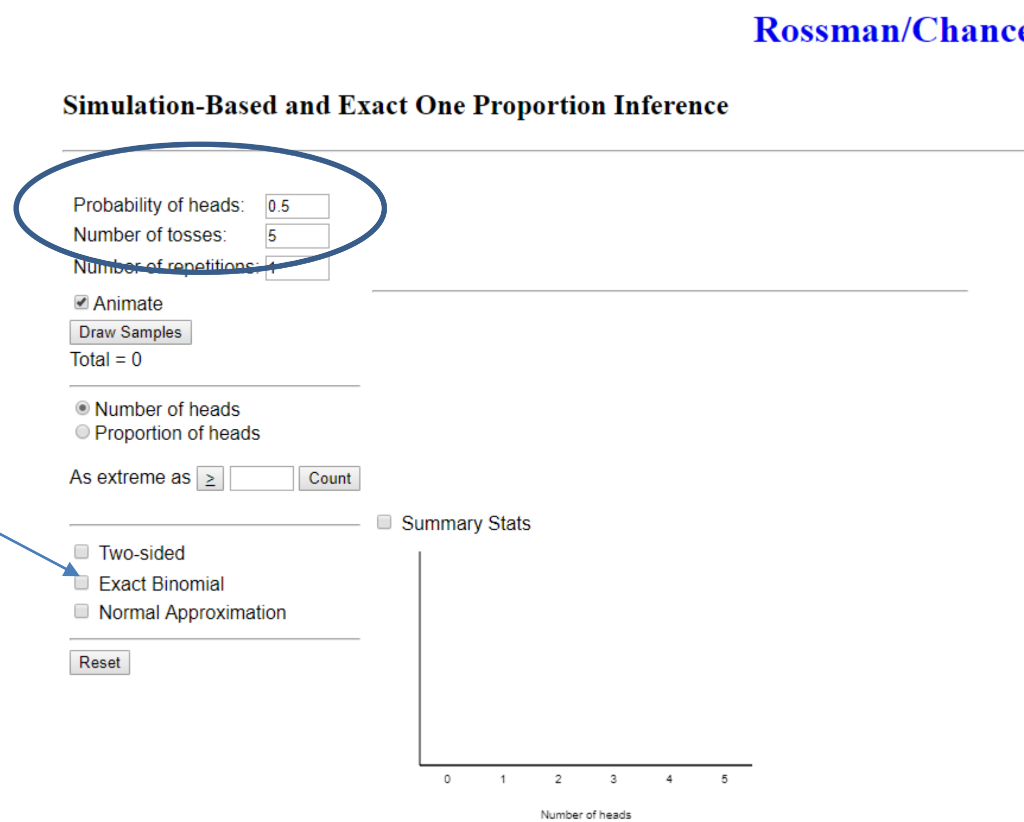
A **p-value** is defined as the probability of seeing a statistic as extreme as ours if the *null model* is true.




We can use the exact formula or coin toss simulation from the One-Proportion Applet IF our data collection is a Binomial Random Process

# Inv. 1.3 – part (j)

Need to identify n and π to compute the p-value with the **exact Binomial** probability formula



# Inv. 1.3: Interpretation of **p-value**

What does our study result,

**the statistic, $\hat{p} = 0.9$,**

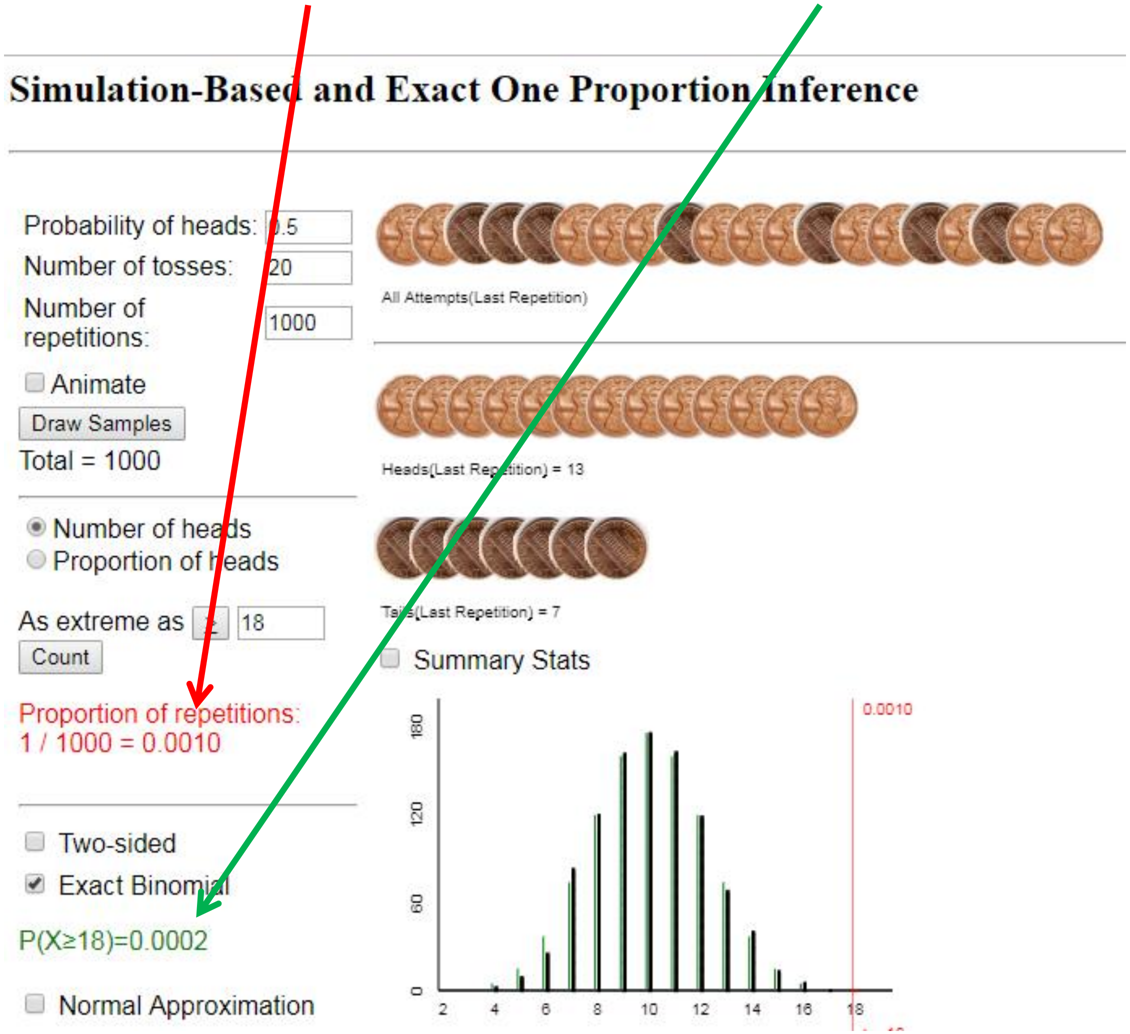
tells us about the

**parameter under the null model, π=0.5** ?

- **P-value** = *probability of seeing 18 out of 20 students ($\hat{p} = 0.9$) chose Tim as the guy on the left if they were randomly guessing (**π=0.5**)*

4. Interpret a **p-value** in context

# Inv. 1.3: Using technology to compute the **p-value** either by simulation or math



# Inv. 1.3: Drawing a conclusion from a p-value

What does our study result,

**the statistic, $\hat{p} = 0.9$,**

tells us about the

**parameter under the null model, π=0.5** ?

***There is strong evidence against the null model: students weren't just guessing.***

Let X = # of students who put Tim on left out of 20.

Assuming X is a **binomial random variable**,

p-value = P(X≥18) = 0.0002

so it would be **very unlikely** to see 18 of 20 students ($\hat{p} = 0.9$) chose Tim on left **if** they were randomly guessing (**π=0.5**).

# Let’s generalize the steps in Inv. 1.3

# A Test of Significance

Research question

**Null Hypothesis:** Nope: there's nothing going on

**Alternative Hypothesis :** Yes, something is going on

Collect data from a sample

Choose a random process that models the data collection well

Compute a p-value, *the probability of seeing results as extreme as the statistic if chance alone is at work*

If p-value is large,
there's no evidence against the **null hypothesis**.
If p-value is small,
there's evidence against
**the null hypothesis**.

# Binomial Test

Research question that involves parameter π from a Binomial Random Process

**$H_0$:** π = some number

**$H_a$:** π ≠ some number

Collect a binary variable from a sample of size n

Verify that the data collection is modelled well by a binomial process

Compute a binomial p-value, either through simulating a coin toss or the exact formula for a Binomial probability, assuming

π = some number

If p-value is large,
there's no evidence against **$H_0$**.
If p-value is small,
there's evidence against **$H_0$**.

5. List the steps of a **Binomial Test**