Unit 2: Foundations for Inference

3. The Central Limit Theorem

(2.5)

2/24/2021

Recap from last time

1. Null hypothesis testing is a framework for quantifying evidence

2. Whenever we pick a standard of evidence that trades off Type I and Type II errors

3. We generally want to use two-sided tests, increasing our standard for evidence

Key ideas

1. Larger samples give us more precision

2. The Central Limit Theorem says that the Null distribution will generally approach the Normal distribution

3. Using theoretical distributions (instead of shuffled random distributions) makes statistical measures lossless compression

Why large samples matter

Suppose I want to know if I can guess the outcomes of coin flips better than chance.

I flip the coin four times and guess correctly three out of four times!

What can we conclude?
Nothing!

Intuition: How likely am I to guess all 4 correctly by chance?

Each correct guess has chance guessing probability of .5. So guessing 4 in a row is

$$.5 * .5 * .5 * .5 = .0625$$

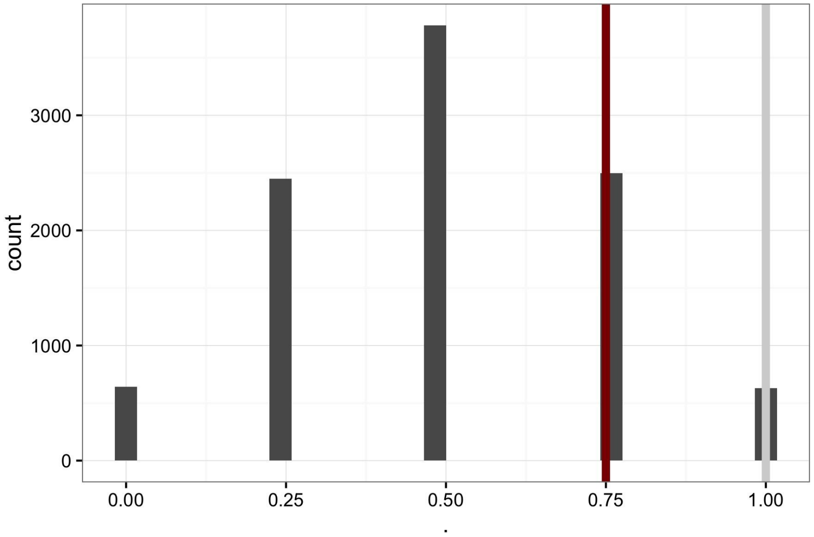
So even if guess ALL of them correctly, we still couldn't reject the null

If our sample is too small, we can never reject the null

Even if I have superhuman guessing ability, I can't tell if I flip 4 coins.

I do not have enough **statistical power** to detect the effect, even if the Alternative Hypothesis is true!

So what does power depend on?



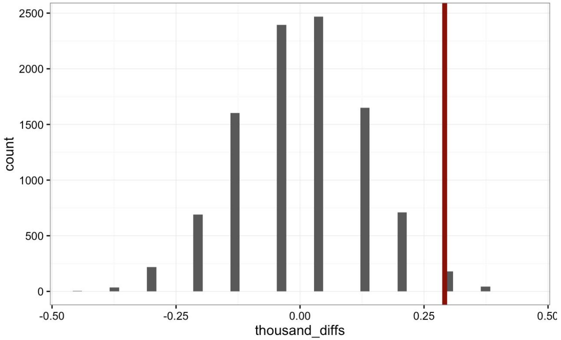
Statistical power depends on...

My ability to reject the Null Hypothesis depends on:

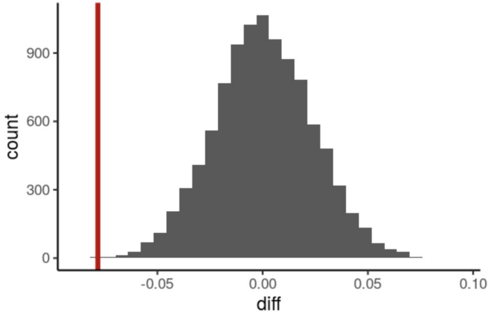
- The size of my sample
- The size of the difference between the True value of the population parameter and the value of the Null distribution population parameter
- My p-value criterion

It is shockingly easy to be in a regime where you can't infer anything no matter how the data turn out!

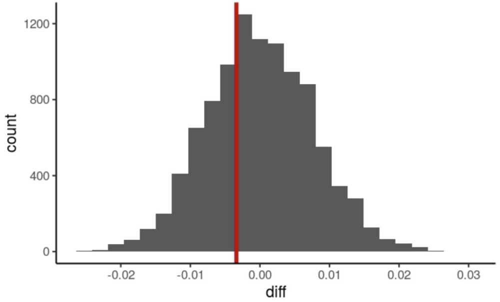
Our null distributions so far



Difference in proportion of women and men promoted



Difference in proportion of cardiac arrests during meetings and non-meetings at teaching hospitals

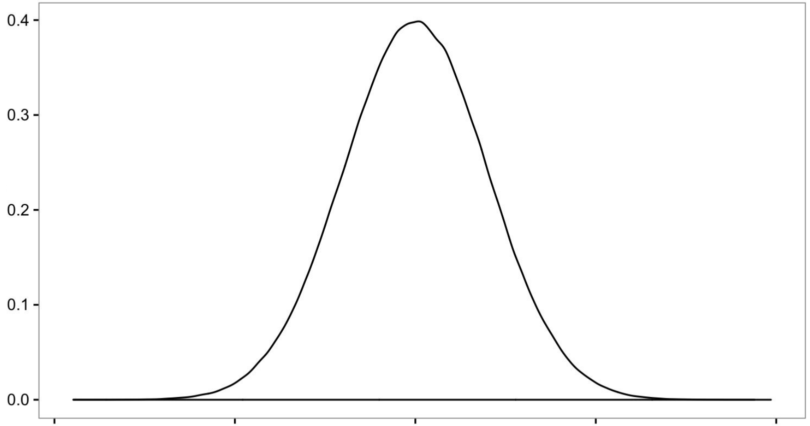


Difference in proportion of cardiac arrests during meetings and non-meetings at non-teaching hospitals

What do these distributions have in common?

The Central Limit Theorem

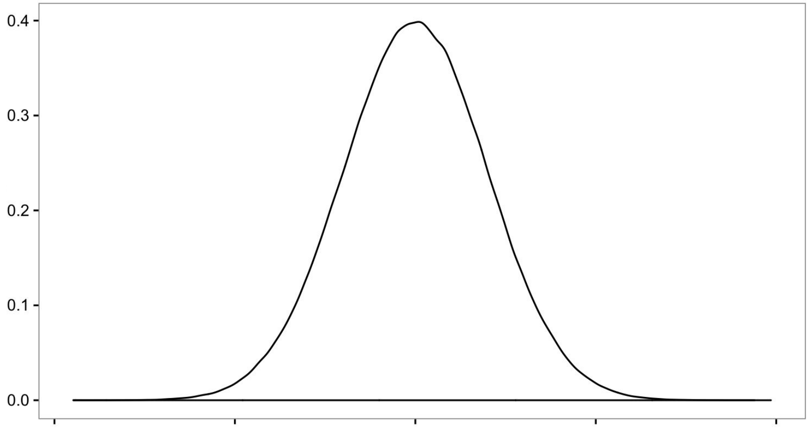
The null distribution for a proportion (or difference of proportions) will approximate the Normal Distribution as the sample size approaches infinity.



https://gallery.shinyapps.io/CLT_prop/

The Central Limit Theorem

The null distribution for a mean of a distribution of **any** shape will also approach the Normal as the sample size approaches infinity



https://gallery.shinyapps.io/CLT_mean/

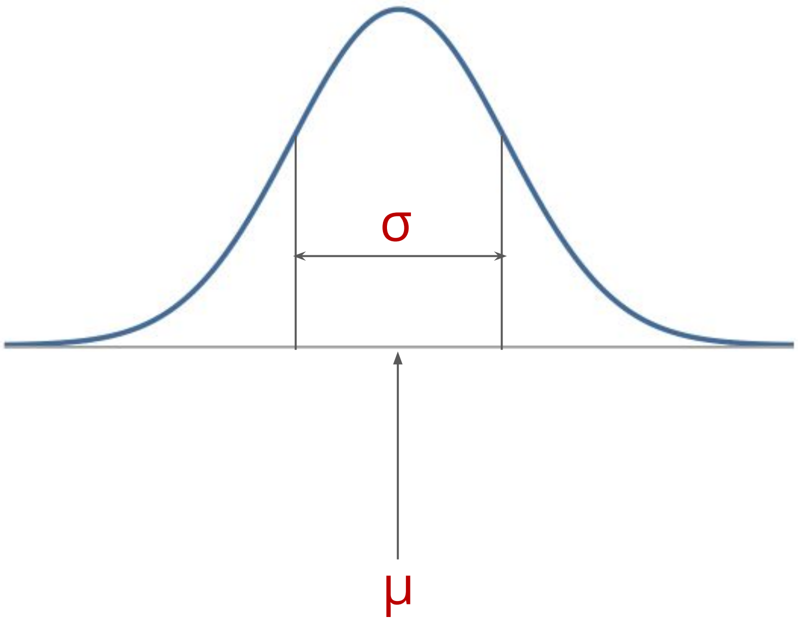
That's why the Normal Distribution is everywhere!

Introducing the Normal Distribution

Unimodal and symmetric

Has two parameters:

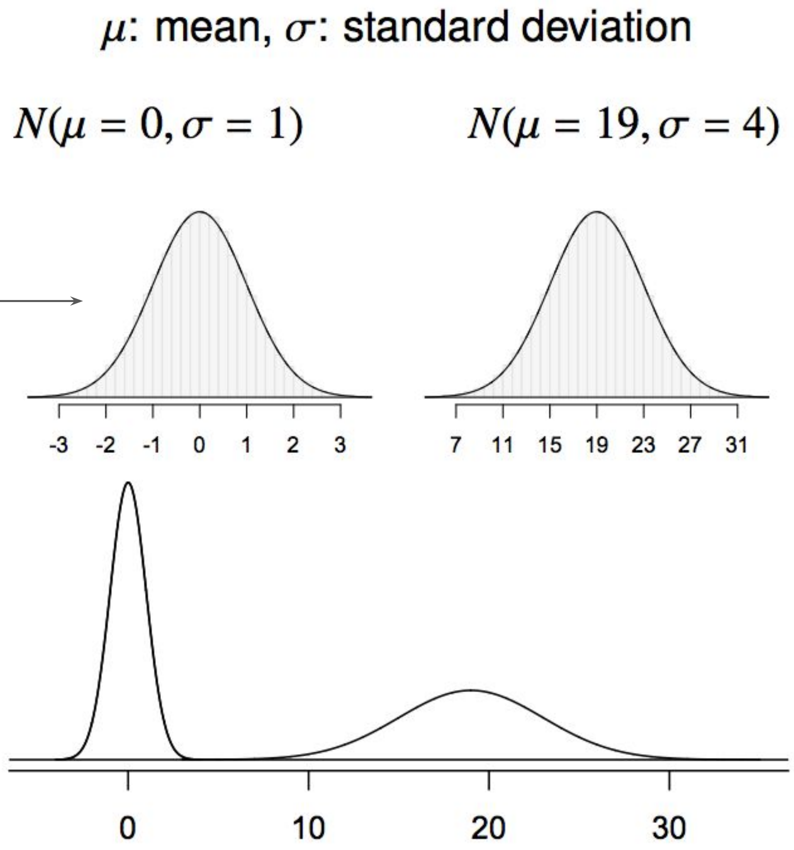
- Mean (μ)
- Standard deviation (σ)



The two parameters completely describe a Normal Distribution

Different Normal Distributions

Standard
Normal
Distribution



Descriptive statistics

What's the difference between .mp3 and .FLAC?
.jpeg and .png?

.mp3 and .jpeg are **lossy compression** -- they make data smaller by keeping only the most important parts of it.

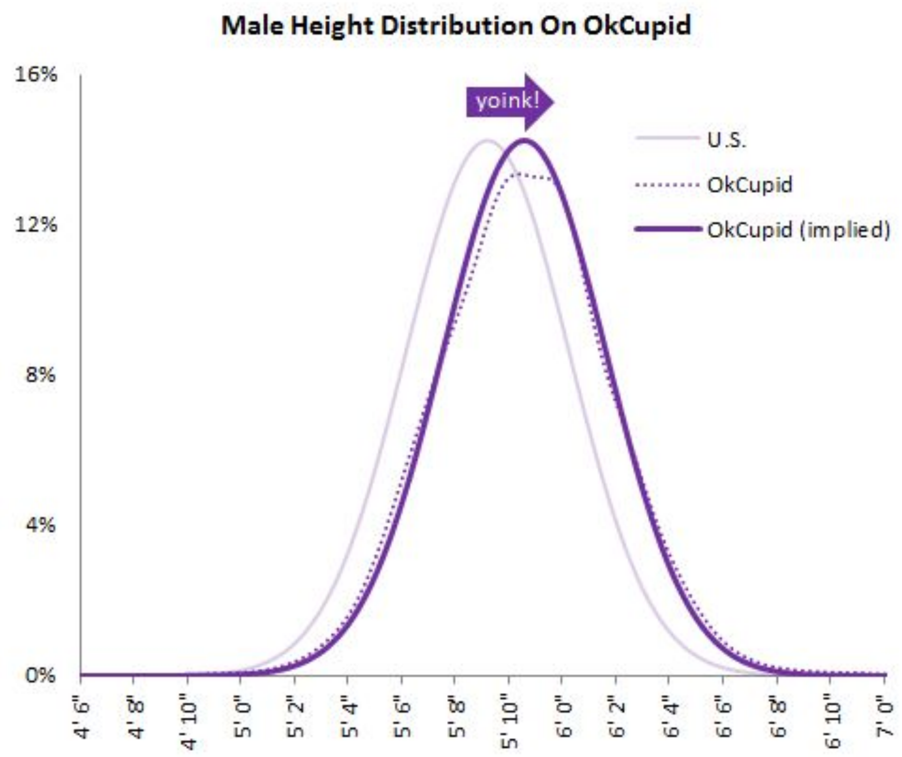
Descriptive statistics are kind of lossy compression:
What one/few number(s) that best represent my data.

But a distribution's parameters are **lossless compression**.
They tell you everything there is to know about it.

Detecting distortions by using a distribution's shape

OkCupid users are (likely) misreporting their heights in **two ways**.

What are they?



https://blog.okcupid.com/index.php/the-biggest-lies-in-online-dating/

Key ideas

1. Larger samples give us more precision

2. The Central Limit Theorem says that the Null distribution will generally approach the Normal distribution

3. Using theoretical distributions (instead of shuffled random distributions) makes statistical measures lossless compression