

Statistical inference in one sentence



Cassie Kozyrkov [Follow](#)

Jan 11, 2019 · 9 min read

Every hypothesis test — from STAT101 to your scariest PhD qualifying exams — boils down to one sentence. It's the big insight of the 1920s that gave birth to most of the statistical pursuits you encounter in the wild today. You can derive our discipline from it, so if you want to understand statistics, graffiti this sentence at eye-level and meditate on it daily.

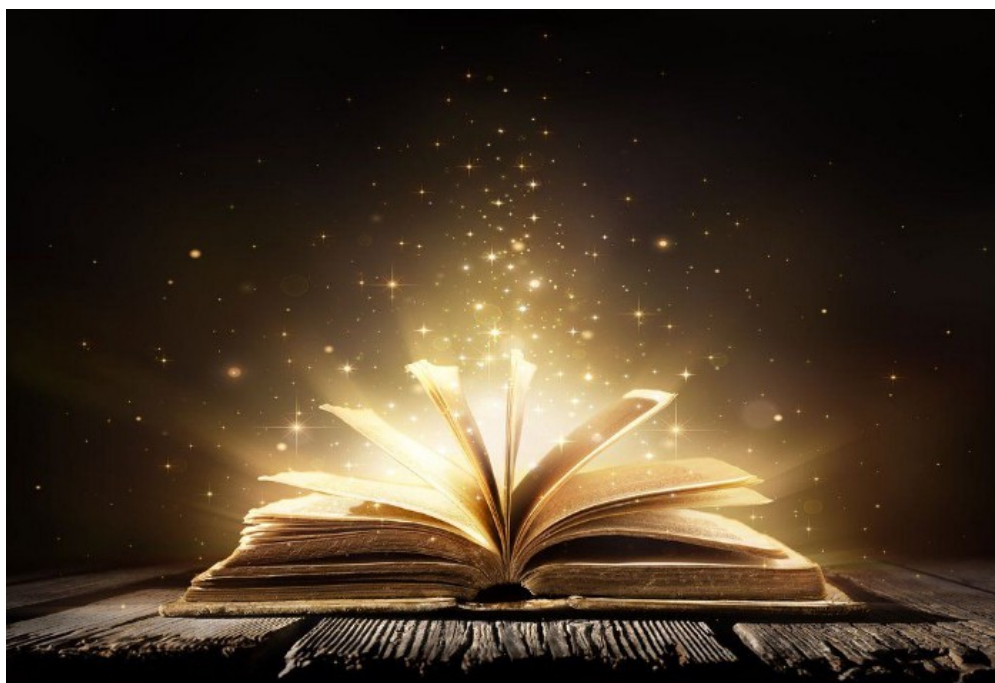


Image: [SOURCE](#)

Enough preamble! Here's the magical incantation itself:

“Does the evidence that we collected make our null

hypothesis look ridiculous?”

I’m not kidding; that’s all there is to it. Classical hypothesis testing is this. Every. Single. Time. Seeing it stripped of its teeth and claws might even feel like a letdown for those of you carrying around STAT101 scars. Others of you might be struggling to make heads or tails of it, so let’s look at a gentle example. *If this example isn’t gentle enough, read [this one](#) instead.*



Hypothesis testing with aliens

You’ve just been selected for the ultimate adventure: searching planets for alien life. Unfortunately, as with every dream job, there is... a manager. Your evil manager has given you a rather paltry user interface. It only has two buttons: YES and NO.



This is the entirety of your control panel. YES, there is alien life here. NO, there’s no alien life here. There is no way to input maybes, comments, or hedging.

In a further stroke of villainy, your manager has not given you the budget to search an entire planet. All you're able to do is land, pick a direction, start walking until your oxygen supply gets iffy, then head back and press one of those two buttons. Since you'll only be landing on big planets and you don't have enough oxygen in the tank to comb every inch of their surface, you'll face uncertainty: you might end up not knowing what the true answer is.

Step 1: What's the default action?

Every hypothesis test starts in the same place. A decision-maker selects a default action. This is the action you commit to taking if you don't examine any evidence. In other words, if you don't even land on this planet, will you press YES or NO?

This isn't a question with one right answer. It's an MBA question that really depends on the politics of your space exploration company, so we'll play through this example with both possible defaults. If you're like most readers, you'd prefer the NO button as default, so let's go with that one first.

Default action: Press the NO button.

Step 2: What's the alternative action?

...and here you were expecting statistics to be hard. The alternative action is simply what you will do if you don't go with your default.

Alternative action: Press the YES button.

If you read my breakdown of how this all works, you'll recall that the only way you'd end up pressing YES is if the evidence makes you feel stupid about pressing NO.

Step 3: What's the null hypothesis?

You've just landed on a planet and you're asking yourself, *"If I knew everything about this planet, which circumstances would make the NO button a happy choice?"* If there is no alien life on this planet. Bingo! That's the null

hypothesis (H_0).

H_0 : There is no alien life on this planet.

Step 4: What's the alternative hypothesis?

The alternative hypothesis (H_1) is everything that's true when the null is false.

H_0 : There is no alien life on this planet.

H_1 : There is alien life on this planet.

Ta-da! You've got your hypotheses set up and you're ready to gather and analyze some data.



Collect data

You're a diligent soul, so you don't just sail past planets hitting NO. You'll land your spacecraft, get out, and start walking in some miserable direction for three miserable hours, then trudge back. Over the course of all this, you have observed... no aliens.

Statistic: 0 aliens.

What have we learned that's interesting?

When I teach this in a live class, the typical response is, “*There were no aliens visible on this three-hour hike.*” That’s a subtly incorrect answer because of how we framed our decision-making.

How you frame your decision-making is important.
Not all decisions lend themselves to the approach
taught in STAT101.

By engaging in classical statistics, we agree to a legal contract that says that only the population is interesting to us. That’s the whole planet’s surface, not this puny sample of a three-hour walk.

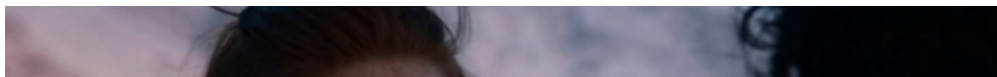
Sample statistic: 0 aliens on the 3h hike.

Population parameter: ?? aliens on the whole planet.

If we were doing analytics, we might be excited by this little factoid we just observed, but that’s not what we’re here for. We’re doing statistics, so everything that’s not informative about the whole planet is boring by definition. We can’t tell whether we saw no aliens because there aren’t any on the planet or because they’re under that other rock that we haven’t turned over yet. We have no way to distinguish between these two possibilities. So, let’s try again. It’s a one-word answer. What have we learned that’s interesting?

Nothing. We have learned nothing interesting.

This is amazing. Do you see what has happened here?





We have just analyzed data and we have (correctly!) learned nothing beyond it. How often do we let ourselves do that? Say it with me: I learned nothing and I'm proud of it!

You should get into the habit of learning nothing more often, because if you insist on learning something beyond the data every time you test hypotheses, you will learn something stupid.

Top highlight

When you are doing the kind of statistical inference that involves confidence intervals and p-values, learning nothing is a very good thing.

This is not analytics!

If this chafes you, take a deep breath. You might be thinking like an analyst while venturing into statistics territory. Here be dragons!

Analytics cares about what's here, while statistics cares more about what isn't.

Everyone is qualified to do analytics: simply look at a dataset and summarize what you see. "These are the facts in this spreadsheet. No aliens observed." You'll be learning something interesting every time in analytics, because the scope of your interest is the data that's in front of your nose. Analytics has only one golden rule: *stick to the data and don't go beyond it*. In that safe space, excellence is measured in how speedy your data frolicking is and you can do no wrong... except accidentally venturing out into statistics. Scary things lurk in the spaces outside your data.

Looking beyond the spreadsheet without hurting yourself takes a different mindset, which is why statistics is trickier. What do we call one of those cowboy types who run around slinging math without understanding the philosophy? A hazard to themselves and others.

Subtle things matter when you do battle with the unknown.

Some people seem to think that whenever they analyze data, the universe owes them insights beyond the facts. If we're making an Icarus-like leap from what we know to what we don't, why would we expect it to be easy?

If you insist on learning something every time you test hypotheses, you will learn something stupid.

Embrace the possibility of learning nothing when you do statistics.
(Starting with this article?)

The beating heart of everything

Statistics is the science of changing your mind under uncertainty. We'll change our minds if we feel ridiculous about persisting in what our evidence calls a foolish endeavor, which is why every hypothesis test boils down to the same core question:

“Does the evidence that we collected make our null hypothesis look ridiculous?”

For homework, you can now go and derive most of statistics. (Or you can keep reading, that's okay too.)

Analyzing the alien data

We've seen no aliens on our walk and our null hypothesis is that there are no aliens on the planet. What's our answer to the big testing question? Does

the evidence make our null look ridiculous? How could it? No aliens in sample is completely consistent with no aliens in population.

Now imagine if instead of seeing no aliens on our walk, we saw this little green guy.




Supposing that's an alien (and not a pickle), what have we learned? If I told you that I have observed this alien and I'm still considering the possibility that there is no alien life on this planet, you will tell me that you have observed an idiot.

This evidence makes my null hypothesis look ridiculous! What do we do when evidence makes a hypothesis look ridiculous? We don't cling to that nonsense. Get rid of it!

Since we always cunningly design our two hypotheses so that they span all possibilities, rejecting one corners us into accepting the other. As good Frequentists, we started with no opinions about the planet. We had a favorite action, sure, but you don't need opinions for that. Beginners seem to get tripped up on the difference between understanding buttons (actions) and understanding planets (hypotheses), but you won't, right?

H_0 : There is no alien life on this planet.



H_1 : There is alien life on this planet.

Testing: *Does the evidence we collected make H_0 look ridiculous?*

If yes:	If no:
<ul style="list-style-type: none"> • Result: Reject H_0. • Conclusion: H_1 is true. • Action: Switch. 	<ul style="list-style-type: none"> • Result: Fail to reject H_0. • Conclusion: None. • Action: Default.

Feel ridiculous? Reject!

If our evidence makes us answer “yes” to our big testing question, we reject this ridiculous hypothesis and make a conclusion in favor of the alternative. We now feel ridiculous about performing the default action, so we switch to the alternative action and press YES. So we have gained knowledge about the planet as a whole: there’s life on it!

Don’t feel ridiculous? Learn nothing.

What about the scenario where we answered “no” to our testing question? In STAT101 class, they teach you to write a convoluted paragraph when that happens. (“*We fail to reject the null hypothesis and conclude that there is insufficient statistical evidence to support the existence of alien life on this planet.*”) I’m convinced that the only purpose of this expression is to strain students’ wrists. I’ve always allowed my undergraduate students to write it like it is: we learned nothing interesting.

Congratulations, you’ve learned nothing!

Learning nothing may seem like a tragedy. We’ve put all this effort into collecting and analyzing our data... and what did we get out of it? Nothing?! Before we wail and beat our chests, remember that we weren’t here to know things. We’re here for decision-making and our endgame is a sensible choice of action, not knowledge. We’re here to press a button, dammit.

Well, when it comes to decision-making, this framework is actually pretty sturdy. Our default action is our insurance policy which makes it okay to learn nothing. It gives us a contract that says, “*If I know nothing, here is what*

I'm going to do about it."

By entering this inference game, we declared that we were happy to take our default action under ignorance... if that's not the case, we shouldn't be in statistics. None of this makes sense without a default action.

Our default action was to press the NO button, so that's what we do when we fail to reject the null hypothesis. We take the action we were happy with because there's no reason to change our minds. Is it the right action? Beats me! But we made an honest effort to talk ourselves out of it and now we're doing what we planned to do with a clean conscience.

Failing to reject the null hypothesis sure doesn't mean that we believe there are no aliens here. For all we know, they're hanging out just beyond that next rock formation. We'd be fools to conclude they're not here just because we didn't find them. If I spend 5 minutes looking for my keys to no avail, it doesn't mean that they're not in my apartment. It means I don't know where they are. There's a difference. (Is your spidey sense tingling uncomfortably? Then read [this](#).)

No reason to change your mind? Proceed with the default action as planned. Is it the right action?

- _(ツ)_/ - Welcome to uncertainty.

To summarize: the game of hypothesis testing is all about determining whether the evidence that we have collected makes our null hypothesis look ridiculous. Everything hinges on how we feel about changing our minds in light of the evidence.

To see what would be different in a parallel universe where the default action is YES instead of NO, read on [here](#). (Hint: Everything changes!)

Cassie Kozyrkov - Hacker Noon

Read writing from Cassie Kozyrkov in Hacker Noon. Chief Decision Intelligence Engineer, Google. ❤️ Stats, ML/AI, data...

hackernoon.com

Sign up for Get Better Tech Emails via HackerNoon.com

By HackerNoon.com

how hackers start their afternoons. the real shit is on hackernoon.com. [Take a look](#)

Your email

Get this newsletter

By signing up, you will create a Medium account if you don't already have one. Review our [Privacy Policy](#) for more information about our privacy practices.

Statistics

Data Science

Technology

Artificial Intelligence

Statistical Inference

Learn more.

Medium is an open platform where 170 million readers come to find insightful and dynamic thinking. Here, expert and undiscovered voices alike dive into the heart of any topic and bring new ideas to the surface. [Learn more](#)

Make Medium yours.

Follow the writers, publications, and topics that matter to you, and you'll see them on your homepage and in your inbox. [Explore](#)

Share your thinking.

If you have a story to tell, knowledge to share, or a perspective to offer — welcome home. It's easy and free to post your thinking on any topic. [Write on Medium](#)

[About](#)

[Help](#)

[Legal](#)