How to evaluate evidence

a statistical proof of the resurrection

About me

Currently, I’m a lead data scientist at Yelp. Looking at data and evaluating evidence is a regular part of my day job.

In the past, I did a lot of teaching. I taught a variety of subjects to a variety of students at all levels of education.

I have a Master's degree in physics from U.C. Berkeley, and a blog where I discuss things like this.

The basic framework

How do we decide if something is true?

Is it going to rain tomorrow? Do you have the coronavirus? Did you win the lottery?

Did Jesus rise from the dead?

The discussion on the resurrection often ends up like this: “here’s the evidence”. “Well I don’t buy it”.

Fortunately, there’s a whole mathematical system for deciding such things. It’s called statistics.

And in statistics, **we use Bayes’ Rule** to decide if something is true.

Intrinsic chance + evidence = conclusion

Or, using the language of Bayes rule:

Prior odds X Bayes factor = Posterior odds

Consider the previous examples:

- Is it going to rain tomorrow? (given that it rained today?)
- Do you have the coronavirus? (given that you tested positive?)
- Did you win the lottery? (given that you claim that you did?)

Think through these - but really, **we want to assign numbers to these factors**.

Where do we get the numbers?

We use empirical data to get the odds. We do not use theoretical calculations without empirical backing. Consider again the following:

- Is it going to rain tomorrow? (what would be the empirical odds?)
- Do you have the coronavirus? (what would be a theoretical odds?)
- Did you win the lottery? (how often have people actually won in the past?)

Given that

Prior odds X Bayes factor = Posterior odds

we want empirical, numerical values for the prior odds and the Bayes factor

Getting the numbers

The empirical prior odds for the resurrection

The skeptic's case: "nobody else has ever risen from the dead"

This doesn't mean that the odds of Jesus's resurrection is zero. After 1000 heads in a row, the next flip can still be tails. Observing a million white swans doesn't mean that there aren't black swans.

The standard practice here says that the odds of a tail are about 1e-3, and odds of a non-white swan are about 1e-6.

For Jesus's resurrection, **assuming all other humans that ever existed stayed dead, this gives us a prior odds of 1e-11**.

The evidence in 1 Corinthians 15

> “For I delivered to you as of first importance what I also received: that Christ died for our sins in accordance with the Scriptures, that he was buried, that he was raised on the third day in accordance with the Scriptures, and that he appeared (1) to Cephas, then (2) to the twelve. Then he appeared (3) to more than five hundred brothers at one time, most of whom are still alive, though some have fallen asleep. Then he appeared (4) to James, then (5) to all the apostles. Last of all, as to one untimely born, he appeared also (6) to me.”

Summarized in this passage are **6 sets of people testifying to Jesus’s resurrection** - many of them are prominent historical figures.

Now we have to **quantitatively weigh this human testimony against the prior odds of 1e-11**.

How to quantify human testimony

You can quantify human testimony, **by seeing how much it changes the odds**. Remember:

prior odds X bayes factor = posterior odds

A strong testimony, with a big Bayes factor, changes the odds a lot.

Example: Is it raining where you live?

Prior odds: before hearing anything, I'd give it an odds of 1e-2 (1%).

Posterior odds: but if you testify that it is raining outside right now, then I'd update the odds to 1e2 (99%).

So your testimony would then have a strength/Bayes factor of 1e4.

Human testimony can be very powerful, but...

Say that I hand you a record of a chess game that I recorded. What is the strength of that testimony?

There's more than 1e120 possible chess games. So before you got my record, the prior odds for a specific game is about 1e-120. If you believe my record with a 90% certainty (1e1 posterior odds), this means the odds went from 1e-120 to 1e1 - so the Bayes factor is 1e121.

So **human testimony can be incredibly powerful - but this example is not directly applicable to the resurrection**.

Odds of the recorded game:

Before: 1e-120

After: 1e1

Bayes factor of the record: 1e121

Human testimony can be near-worthless, but...

How are you doing? Are you guys having fun? Are you understanding everything?

That was a trick question!

I don't really expect an honest answer here. Most speakers/teachers don't, unless there's a dramatic response of some kind. Other questions of this kind include things like "did you ever cheat?", "is the coin heads?", or "are you busy that day?"

In all these questions I don't expect to change my odds much from the answers - **so human testimony can be quite weak. But these examples are also not applicable to the resurrection.**

The right testimony: an unlikely, special claim

For the chess example, the claim wasn't "special" in any way. The recorded game was just one game among many possible games. This is NOT like the resurrection, so it's not applicable.

For the "are you having fun?" question, the claim wasn't unlikely enough. The priors for these questions are already quite large, so there isn't much more certain that they can get. Again this is NOT like the resurrection, so it's not applicable.

What we want instead are examples of a special kind: **an intrinsically unlikely claim, which is special in some way. Those are the testimonies that are applicable to the resurrection.**

The right testimony: examples

We don’t want testimonies like “I’m a safe driver” or “I have a car” - too likely.

We also don’t want testimonies like “I was making a right turn onto 5th street, going 26.5 mph, at 6:47pm on June 21st, 2020” - unlikely, but not special.

Instead, we want a driving testimony like this one, to apply to the resurrection.

We also don’t want testimonies like “I sometimes play the lottery” - too likely.

Nor do we want “I picked 4, 5, 21, 23, 34, and 45” - unlikely, but not special.

Instead, we want a lottery testimony like this one, to apply to the resurrection.

Calculating the Bayes factor from the examples

Bayes factor of the right kind of human testimony - like the ones shown in those videos - **is about 1e8**. That's the number we need to use for the testimonies in 1 Corinthians 15.

I have about a dozen calculations like these in my blog post - empirically sourced human testimonies which are applicable to the resurrection. **They all converge around 1e8 as their Bayes factor.**

A second way to quantify human testimony

The Bayes factor, by definition, is how well the testimony is explained by the claim being true, vs. how well it's explained by the claim being false.

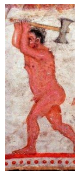
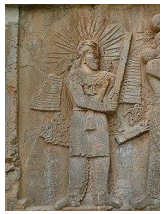
For the resurrection testimonies in 1 Corinthians 15, if Jesus did really rise from the dead, that explains the testimony quite well.

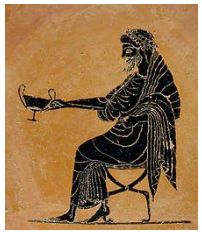
But if he didn't really rise from the dead, what are the chances for such resurrection stories coming from a non-resurrection?

"What are the chances" is not an idle question. We can answer this numerically and empirically, through the historical record. **If these testimonies are weak, then it should be easy to find equivalent false positives in history.**

The historical false positives

Try this experiment right now: get out your phones, and search for the most convincing non-Christian resurrection stories you can find. We want the best that non-Christians have to offer, in creating a “false positive” of a resurrection story.


Who testified to the supposed resurrection? Were any of witnesses prominent historical figures? What exactly constituted their testimony? Where they made within the lifetime of the witnesses? With corroborating historical evidence surrounding their claim? **Do any of them match up to what’s in 1 Corinthians 15?**

Calculating the Bayes factor from the false positives

There are zero false positives in history that match up to 1 Corinthians 15.

There are zero false positives that match up to even the weakest of the six testimonies in 1 Corinthians 15.

Meaning, it's virtually impossible to create such a false positive, even for the weakest of the six. If there has been a billion (1e9) people in history whose deaths have been recorded, that means that the chances of a false positive are about 1e-9. That's how bad the non-resurrection hypothesis is at explaining the testimonies.

Then the Bayes factor for the weakest of the six testimonies is about 1e9.

Putting it all together

So did the resurrection happen?

Bayes' rule:

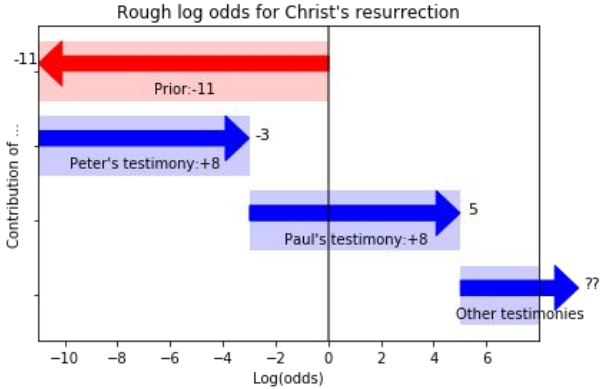
Prior odds X Bayes factor = Posterior odds

Prior odds = 1e-11

Bayes factor (say, Peter's) >= 1e8

Bayes factor (Paul, who's independent of Peter) >= 1e8

Posterior odds, after just those 2 testimonies >= 1e5



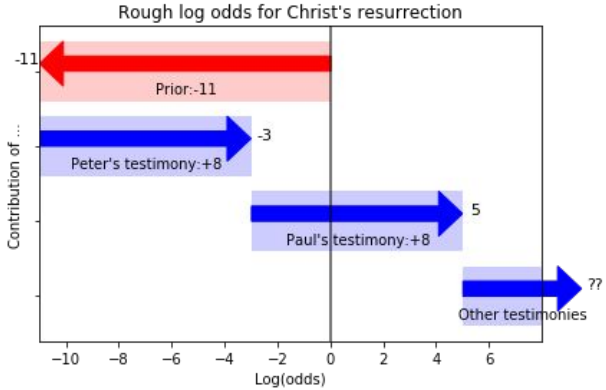
Objections, double checks, and robustness

What about dependency factors?

We have to be careful here. We CAN’T treat all 6 sets of testimonies in 1 Corinthians 15 as independent.

Assuming independence is a common mistake - and it leads to an absurdly large combined Bayes factor (1e48) and posterior odds (1e37). We can’t be that certain.

But **Peter and Paul are sufficiently independent, and their testimony is enough to get us safely well over even odds.** The remaining testimonies can then just be used as backup.



What about EXTREME dependency factors?

What about some kind of crazy conspiracy theories, or some runaway feedback effect, which generated the crazy amount of evidence for the resurrection?

Even these can be ruled out probabilistically.

As a rule of thumb, no metric of human activity - even the ones intrinsically likely to “runaway” or “go viral” - cannot have the first place winner beat the runner up by a factor of 6. **So such a “first place” must belong to a different category altogether.**

Test the rule of thumb, with metrics like these: who is the richest person, what is the most popular YouTube video, who is the greatest athlete, etc.

"But what if..." speculations

“What if the disciples went to the wrong tomb?”

“What if Jesus just swooned?”

“What if the disciples just hallucinated the post-resurrection appearances?”

All of this amounts to ignoring the existing evidence, and making stuff up instead. Our empirically based calculations already take all these possibilities into account, and it shows that they’re unlikely.

Only evidence moves the odds. Speculations do nothing.

“But what about…” observations

“But what about all the other stories about all the other people who are said to have risen from the dead, like Krishna, Asclepius, Romulus, or Hebo?”

We’ve discussed this in the “historical false positives” - but these weak cases actually strengthen the case for the resurrection. It is the failure of these cases to measure up to the bar set by Jesus’s resurrection that makes Christ’s case so strong.

The more such weak examples the skeptics come up with, the STRONGER the case for Jesus’s resurrection becomes.

Additional evidence we haven’t considered

More evidence from people other than Peter or Paul, or not mentioned in 1 Corinthians 15 at all.

Built-in anti-conspiracy features: e.g. Mary Magdalene being the first to see the resurrected Christ.

You can run the kind of analysis I did here for the resurrection, but expand it to miracles in general. Again, the resurrection blows all non-Christian miracles out of the water.

What would falsify my conclusion?

My result comes from empirical data, and the mathematics of Bayes' rule.

This makes it VERY robust. It is simply immune to much of the "what if / what about" type of objections. It makes very generous concessions to the skeptics. It takes into account even the craziest, most conspiratorial possibilities.

In order to falsify my conclusions, one would need additional empirical data (which the skeptic's don't have), or point out an error in the math.

But barring that, we have our conclusion.

Conclusion

Jesus almost certainly rose from the dead

With the final, posterior odds well in excess of 1e5!

Questions, comments, and further reading

All this is a very condensed summary of the full content of the post in my blog, where I go into things with much more depth and rigor.

Any questions and comments can be emailed to me at

naclhv at gmail dot com