Ryan Bergner
9-28-22
Philosophy of Statistics
Prof. Biran Zaharatos

The Bayesian Probabilistic Interpretation :

Why Frequentist "Rigor" Skews Realistic Thought

I. Introduction and Background: The Frequentist vs. Bayesian Debate

For centuries, the idea of probability has been used to answer questions about the chances of occurrences of events. However, more broadly, 'probability' is the method human beings have to think about the world, our lives, and how we attempt to grapple with the innate desire to know the unknown. In an attempt to develop accurate, objective, and reproducible methods of conducting such inductive reasoning, philosophers have debated the validity of numerous schools of thought, including but not limited to the debates over the proper use of inductive inference (or whether or not inductive reasoning should be used as means to base any factual claim), finding more robust methods for deciphering coincidental correlations from causal and the possibility of mathematically claiming causality, and of course - the great debate of the Bayesian and Frequentist interpretations of probability.

Ia. The Reasoning to Adopt a Probabilistic Interpretation

To illustrate the importance and notoriety of the ever-lasting debate of the "better" probabilistic interpretation, this particular point of disagreement among advocates of Bayesian and Frequentist interpretation has always taken many forms - from either shining brightly at the forefront of the world of probability to quietly, annoyingly reminding theorists of its presence in the background. This is the case for good reason, as the very idea of probabilistic interpretation is what allowed the founders of probability to lay a framework to create the statistical methods we have come to rely on. However, more importantly, the belief and utilization of a probabilistic interpretation provide a framework for how humans should logically make decisions when faced with limited information. Probabilistic interpretations, in general, are the methods mathematicians have created over the past century that define what we should do to obtain our best guess at objective reality. The real debate is centered around the idea of human epistemology: Is a belief reflective of the world around them? Or is it simply informing how one adjusts their current mindset in an attempt to interpret reality as they see it over time?

While the difference between both is notable in practice, it is clear that not accepting a probabilistic interpretation is to not benefit from the meaning of any particular observation on an individual, sometimes themselves, to construct a reality based on subjective opinions formed in the absence of any factual substance. As most scientists, theorists, and logically motivated humans can agree, the worldview above (and overall negative, skeptical attitude towards the use of probability interpretations in general) is certainly not an optimal way to make decisions, and almost certainly guaranteed to open the door for certain humans to be manipulated by others at some point. These characters have long aimed to change peoples' factless but realistic and mostly understandable perspectives to benefit their own personal interests, such as politicians, journalists, and the leaders of nations.

1b. Introduction of the Subjective Bayesian Interpretation

Per the previous assertion, the consideration (or lack thereof) of a certain probabilistic interpretation, namely the Subjective Bayesian, Objective Bayesian, and Frequentist probabilistic interpretations, will direct an individual (thereby semi-directing or guiding society) toward what we call "the truth." However, whichever interpretation is employed on an individual level may lead to a different conclusion that can take different forms.

I claim the subjective variant of the Bayesian Interpretation of probability to be the method that, if adopted by an individual, leaves that individual with the most excellent chance of gaining an accurate, unbiased view of most situations in which they have limited information. In this context, when we refer to the subjective Bayesian - the literal meaning of subjective must be taken to mean "subjectively immune." That is, no matter what prior beliefs are held regarding a subject, we will eventually epistemically converge to whatever belief is the most advantageous to be held to best suit one's needs at the time based on the current situation.

However, there is a case for some instances proving frequentist inference to be useful in the context of non-hypothetical events or well-defined processes. These claims will be articulated and argued later in the body of this work.

II. Accepting The Bayesian Interpretation

To define the theory and logical underpinnings of the Bayesian Interpretation, which was initially, one must understand the advantages of a subjective approach when interpreting the chances of potential outcomes of a random event. The classic quote from Margaret de Valois summarizes the underlying argument for the bayesian interpretation just short of perfectly:

"There is no greater **fool** than he who thinks himself wise; no one wiser than he who suspects he is a **fool**." - Margaret de Valois

To frequentists, whose methods have dominated probability and statistics for the last century, "reality" is a set of universally held beliefs, and hence the decisions we come to are thought to define the real world. However, the quote exemplifies the core belief those advocating for the Bayesian Interpretation hold: humility in the confidence of their predictions with the pairing of encouraging mental flexibility actually leads to the most accurate answer. However, to accurately accomplish this Probabilistic interpretation, in general, is the practice of attempting unseeable or otherwise unknown occurrences. Statistics (and statistical thinking in general) is supposedly based on coming to conclusions based on observations in the form of data. Essentially, the Bayesian probabilistic interpretation is a realistic, more flexible way of thinking that always assumes that any particular observation could be somehow corrupted. The only thing that can mitigate this corruption is to observe an ever-increasing number of instances.

IIa. Examples of Advantages & Contrasting to Frequintism

The overarching reasoning to support the bayesian interpretation over that of the frequentist stem from the realistic nature and simple advantages of observing data on an ordered basis in the most detailed manner possible, considering a distribution's shape in addition to its expected value and variance parameters. This viewpoint aligns with the Bayesian interpretation at its core in that we are grounding our updates in observable instances and considering those observances as they are - observances - as opposed to reality.

The Bayesian interpretation is not only advantageous in general – but far superior when dealing with dynamic decision-making. One of many specific types of examples includes competitive games involving sampling without replacement from a relatively small and static sample space.

Consider the example of a high-stakes gambler playing a competitive game of poker. For simplicity and clarity, the match style is Texas Hold em', in which players are dealt 2 cards, and 3 cards are selected to be the first cards "on the river." Directly after, two betting cycles occur, and once five cards have been placed on the river, all remaining players in the game see which of their combinations (seven choose five) produce the strongest hand, and the player who has the strongest hand wins thereby receiving all of the money bet that round. The general goal of the game is to "play" or bet money when you believe you have the best hand - an occurrence only realized after the remaining players' cards are flipped. The full set of information is completely, and perhaps more importantly, uniformly available to all players. Furthermore - if a "tie" occurs and two players have the same strength of hand at the end of the round, the money goes to the player with the higher hand

Therefore, the winner of the game is usually the player who can be deceptive and convince others they have weak hands when they are actually strong – and vice versa when the player has a strong hand. They are also usually defined by consistently and correctly making decisions based on two mathematically defined, dynamic factors:

1. The probability the player has the hand <u>they will eventually bet,</u> given the fact they don't know what the remaining unflipped cards on the river are

2. The probability the player will win using the hand after all card faces are revealed, given the possible hands of the other players

If the player wanted to increase their chances of winning this game, they could opt to play with a probabilistic interpretation to gain a better sense of the moves they need to make to give them the best chance of winning. Suppose the player used the frequentist method to see whether to play a certain hand - thereby assuming they have the ability to predict an inherent characteristic of the poker process; considering all the possibilities of hands dealt and in-game decisions made to either bet, call, or fold that proved to be victorious eventually. One may conclude the best way to win the game would be to calculate the chances of being dealt the strongest hand. If the player obtains the strongest hand, he has the best chance of winning - right? That's what the dataset showing the proportion wins a certain hand receives when the cards are flipped would confirm. As a result, a player might inform their strategy by only betting when they have one of the strongest hands, such as a royal flush (Ace, King, Queen, Jack, 10 [of the same suit]), a three-of-a-kind, or a full house. After all, these are the hands most likely to win against the other hands - it seems like a decent strategy.

However, an important consideration to note is that the overwhelmingly vast number of poker hands are not won using the strongest two-card hands. There exists a problem with the overarching philosophy of the player's approach - they are quite literally attempting to draw an empirical conclusion to something that is inherently random at its core. To learn to predict any outcome based on the data of past poker games is objectively foolish because, like all processes in at least some way, they are random. To consider the deck of cards being dealt to have any “inherent” or predictable properties would be akin to the player being incredibly predictable to the other players at the table, purposefully leading themselves objectively – so why would we even attempt to define it? Drawing a conclusion about such an unpredictable outcome may make an initially foreign idea seem more interpretable and palatable for the collective human consciousness - but what purpose does it hold for anything besides leading us down a self-sabotaging path? Why would we not attempt to see the behavior of the process and instead use massive assumptions to skip to the final, almost certainly incorrect answer?

In reality, accurately betting on strong hands, given the randomly presented evidence they gain throughout the round is the mental framework of every successful poker player in the world. The prove this, observe the bayesian can navigate the following situation:

> Statistically, the two-card hand with the lowest probability of winning across all poker games is a seven-two with both cards having different suits. When compared to any possible river, the number of winning combinations is the lowest when a player has this hand irrespective of cards on the river. A mindset utilizing frequentist principles would conclude the best decision would be to fold the hand every single time automatically.
>
> An important aspect to note however is that in Texas Hold Em, the cost to fold any other non-winning hand is zero, even if the hand is strong before the cards are revealed. Furthermore, a good poker player could “bluff” in this situation - pretending they had a pair of aces on another statistically strong hand by betting aggressively. This tactic intimidates the other players into folding their own hands - even though their hands may have been statistically stronger initially.
>
> Consider the one vs. one Hold Em’ game in which another player at the table draws a great hand of an ace-two of the same suit. With a frequentist perspective, that player would be prone to aggressive betting because many rivers will eventually create a winning combination. But what if further into the round, the river was shown to be:

- River: {“Seven of Hearts,” “Seven of Diamonds,” “Two of Clubs,” *unflipped*, *unflipped*}

Again, the two players have the following sets of cards:

- Player one: {“Seven of Spades,” “Two of Hearts,”} = Poor hand under frequentist interpretation

- Player two: {“Ace of Spades,” “Two of Spades”} = Strong hand under frequentist interpretation

Player two would probably be overly excited to bet after seeing the two clubs on the flip, reasoning they at least have the possibility of a two-of-a-kind with the possibility of getting a two-of-a-kind. This is a decent position to be in - but its safety is rarely guaranteed. With the information available, the probability of flipping another ace is low: Player two already has one of the four aces, and Player one could possibly have one or both of the other aces. A pair of twos is a mediocre hand: When other players don’t have good hands, the bet usually pays off. But, a two-of-a-kind of twos is a low value and would be beaten by any other two-of-a-kind hand - and the occurrence of a higher value of the same hand from any other player is surprisingly probable.

If we look at the state of Player one, they already have a full house - one of the best hands in poker (two-of-a-kind and three-of-a-kind simultaneously). This hand literally beats any other hand besides a four-of-a-kind, straight flush, or a royal flush: three hands that are all extremely improbable to occur. If Player one can act ‘cool,’ he could convince Player two they had a bad hand, causing Player two to overbet. In the end, this would eventually cause Player two to lose a large sum of money to Player one in an embarrassing fashion.

In summary, the point is that random processes cannot be reduced to an extrapolation without considering the steps in the process and how each step incrementally changes the behavior of some aspect of the experiment. This is the case for any process that requires probabilistic interpretation such as evaluating future stock prices or evaluating future demand for a specific company or industry. They are all inherently random, so we can not act as if they are not random.

III. Criticisms of The Frequentist Interpretation

On the other side of the aisle stand the frequentists, whose theories have been spearheaded by numerous forefathers, such as renowned British Mathematician Ronald Fisher, whose work sought to create methods the accuracy of popularly used statistical methods such as the F-distribution, ANOVA tests, the Student's t-distribution test, and Linear Discriminant Analysis (Hald, 53). These methods were designed with a heavy emphasis on the mathematical rigor that served as the backbone for frequentist interpretation. It was this mathematical rigor that convinced propelled Fisher to the forefront of statistics. The overarching reasoning behind the interpretation is that probabilistic calculation, given a certain level of confidence, leads to an answer that describes aspects of the real world. Due to this notion of frequentist models uncovering empirical truths, the scientific community, over the course of the last three centuries, has largely adopted and accepted the use of these methods in totality.

To a frequentist, finding the most accurate interpretation relies on two main philosophical principles to make their claims: the world is well defined and can be well defined up to a point by humans. However, this sort of worldview is somewhat arrogant in the same as religion and other belief systems we have come to know as non-factual and presumably sub-optimal guides: Frequentists essentially put themselves in the shoes of an all-seeing, god-like being to make a prediction that seems like a god-like command or inherent wisdom to any regular individual; after all, most of the world doesn't study statistical thinking and probabilistic interpretation regularly - as a result, they will always turn to the experts for guidance.

They make a claim a business would <u>want</u>; in our profit-driven economy, leaders don't care so much about accuracy as they do about efficiency and velocity. Polynomial and linear extrapolation of a stock price or any demand-related, macroeconomic metric is almost always snake oil: If the guess is eventually right, the data scientist will be praised by his superiors for being a ‘genius,’ but when an unforeseen event such as the covid-19 pandemic takes place, the data scientist can have an excuse to be wrong because they never had the data. However, just because a data scientist doesn’t have the data doesn’t mean they can’t consider all possibilities and realize some similar event is eventually bound to happen. The whole practice leads statistics down a path of accepting mediocre answers to please those who benefit from presenting “good metrics” in exchange for praise and salary bonuses. In business settings, where the correctness of one’s predictions oftentimes cannot be verified by a superior, invalid and fraudulent claims are rampantly made due to the ease of manipulating frequentist parameters (e.g: removing rows of data, p-hacking) - but the statistical practitioner can seldom be criticized because many people in the company treat them as some risk-mitigating, magical prophet.

They make a claim a fraud scientist would <u>want</u>; “if we can pinpoint a single value with enough confidence, we can basically extract constants out of nature, such as pi or Euler’s number.” But, the threshold for discovering such constants is far greater than what can be provided with frequentist statistics or any statistics at the current moment.

The question that’s been stuck on my mind since entering this never-ending debate that almost everyone had seemed to ignore, or at least never question, is, “What did Robert Fisher <u>want</u>? What motivated him to create all of these statistical methods that the entire world relies on today?” The truth is that Robert Fisher was one of the most well-known eugenists of his time and the statistical methods he made were, in part, to benefit himself and his own agenda. A eugenist who believed in and wanted to “prove” the inherent racial superiority of white people may be motivated to make rigid claims about the real world – even when they have no basis to do so. Who was there to check Robert Fisher on his biased views in the 1920s? I suppose the only constituents with an opportunity to criticize his work would be…other similarly-minded, racist white males in groups such as those in “The Cambridge Eugenics Society” an organization that Fisher himself founded. But, the promises of Fisher’s work probably felt like pure heroin to other lazy, fraud scientists who only aimed to hold intellectual superiority over others (Britannica). Imagine how embarrassed Fisher would be to know there are countless examples of non-white individuals objectively beating out white people for job positions, in academia, and just about every other context in western society regarding intelligence.

He would probably retort, saying the data wasn't present at the time. Well, then why even make the claim in the first place? If you knew you would be wrong eventually, what purpose does making such an empirical (and dangerous) claim serve you besides seeming like an all-seeing oracle to others? I'll allow you, the reader, to come to your own conclusion about that point.

The problem with Frequintest Interpretation is quite the fundamental assumptions of the theory fall short in terms of accurately making claims regarding the real world. To underscore every notion spoken of previously, there have been numerous instances of frequentist mindsets that have made what turned out to be wildly inaccurate claims and/or predictions, either under ignorance or with the intention or desire to manipulate.

In summary, the frequentist method, when used to predict unknown outcomes based on the language of cheap journalists and politicians: those who seek to win some short-term popularity contest by convincing and manipulating less-aware, regular people rather than doing the hard work of showing them the truth – thereby empowering them. The specifically-defined output of a frequentist conclusion may be more understandable to the layperson, but it is certainly not to correctly guide their decision-making. The interpretation was found in an attempt to reduce our human discomfort with ambiguity, but use a fairly weak justification for doing so. The truth is that the very idea of "truth" is created by humans. Therefore, it makes no sense to place such universally-rooted certainty in the beliefs we as humans hold.

IV. Rebuttals to Frequintist Criticisms of the Baysian Interpretation

The school of frequentism often objects to the conclusions of those who hold such flexible opinions of the word as impeding scientific progress by not giving a well-defined stepping stone to prove other scientific theories. It argues if we cannot come to a specific, conclusion, then we cannot have an empirically defined basis for the purpose of future scientific experiments, thereby hindering scientific progress.

I refute this disgusting claim in its entirety: scientists have a responsibility to convey their findings for the purpose of benefitting society, not themselves. The problem with this view is that it causes frequentists to see the absence of their hypothesis being realized as a "problem."

This is absolutely, unequivocally not a problem unless the statistician is solely concerned with their personal interests. They are simply guilty of not collecting the amount of information needed about a process to say anything meaningful about it. When you don't find the answer you want, you can't adjust your model to make yourself seem correct.

Sometimes, this is just an inevitable occurrence - predicting the future or unknown event is not a simple task and is not supposed to be easy. This uncertainty can be conveyed with a Bayesian interpretation, considering the fact they could be wrong and additional flexibility may need to be considered to correctly model a process. Making massive assumptions can never and will never be a replacement for counting even though the former involves much less work and the latter requires much patience to complete, especially in the early 1900s. However, the presence of high-powered computers of today makes counting large combinatorial objects fairly simple. Due to this fact, I claim there is no reason to use anything the Subjective Bayesian Probabilistic Interpretation.

Works Cited

Books

Clayton, Aubrey. Bernoulli's Fallacy: Statistical Illogic and the Crisis of Modern Science. Columbia University Press, 2021.

Hald, Anders. A History of Parametric Statistical Inference from Bernoulli to Fisher, 1713-1935 (Sources and Studies in the History of Mathematics and Physical Sciences). 2007th ed., Springer, 2008.

Articles

Britannica, The Editors of Encyclopaedia. "Sir Ronald Aylmer Fisher." Encyclopedia-Britannica, 25 Jul. 2022, https://www.britannica.com/biography/Ronald-Aylmer-Fisher. Accessed 11 October 2022.

Hájek, Alan. "Fifteen Arguments Against Hypothetical Frequentism." Erkenntnis, vol. 70, no. 2, 2009, pp. 211–35, https://doi.org/10.1007/s10670-009-9154-1.