

A frequentist analysis of White’s Coin Puzzle

Shivaram Lingamneni

April 24, 2014

Abstract

White’s Coin Puzzle is a paradox for unsharp credences, or more generally, for the suspension of judgment about probabilities. I argue that a frequentist analysis dissolves the paradox; through a frequentist lens, the puzzle can be seen to argue for suspended judgment and against White’s position.

1 Frequentism

Frequentism is the idea that probabilities have to do with frequencies. If this sounds vague, there is a reason for it: there is no universally accepted view among frequentists about the precise relationship between the two notions. Indeed, Hájek [1996, 2009] argues convincingly that all existing accounts of the relationship are defective — so there is a need for additional foundational work on the subject. Nevertheless, even saying exactly what frequency probabilities are, frequentists are clear on what they are not: subjective degrees of belief or confirmation, meaningful in the single case even without a notion of repeated trial. A key implication of the frequentist interpretation of probability is that there is not necessarily any fact of the matter about how to assign individual Bayesian credences. If there is a fact of the matter, the way to get at it is via frequencies.

A frequency is defined relative to a particular *reference class* of repeated trials, and it is the ratio of successes to trials in that class. Since the reference class may be ambiguous or even nonexistent (in the case of a necessarily single-case event), frequentism means being comfortable with suspended judgment about probabilities, or more generally with the idea that there may not be obligatory rules for reasoning about probabilities. A natural question to ask is to what extent this suspended judgment can be interpreted in the formalized epistemological and decision-theoretic framework of *unsharp* or *mushy* credences. Where conventional Bayesianism holds that degrees of belief must be described by real numbers between 0 and 1, the unsharp credence view allows them to be described by intervals $[l, h]$, where $0 \leq l \leq h \leq 1$; in particular, the assignment $[0, 1]$ is a maximally noncommittal credal state, representing a complete suspension of judgment.¹ But it seems clear that even if unsharp credences do not perfectly represent or capture frequentist

¹See Gärdenfors and Sahlin [1982] for a seminal work on unsharp credences, and Joyce [2010] for a more recent survey.

attitudes, any argument that they are inherently irrational — notable examples include Elga [2010] and White [2009] — is also impugning the rationality of frequentism.

Roger White’s aforementioned paper advocates the Principle of Indifference as the correct alternative to unsharp credences; rather than suspending judgment about a question, one should adopt sharp credences that are uniformly distributed among the possible answers. For example, if one is completely uncertain about a proposition q , the Principle of Indifference yields $P(q) = P(\bar{q}) = \frac{1}{2}$.² However, from a frequentist point of view, White’s framing of the debate is problematic. Throughout the paper, White treats the Principle of Indifference as an *obligation* of rationality; he views it as a compulsory constraint on epistemic agents facing symmetric evidence. Thus, when he discusses the unsharp credence view, he also sees it as a candidate for an epistemic obligation. However, it seems more natural for the unsharp credence view to allow for choice — sometimes you may wish to apply the Principle of Indifference and sometimes you may wish to suspend judgment. The problem is then to seek clarification (but perhaps not *algorithmization*) about the kinds of situations each is applicable to.

A relatively simple example is White’s “Chocolate Box” scenario (183):

Four out of five chocolates in the box have cherry fillings, while the rest have caramel. Picking one at random, what should my credence be that it is cherry-filled? Everyone, including the staunchest opponents of POI, seems to agree on the answer $\frac{4}{5}$. Now of course the chocolate I’ve chosen has many other features, for example this one is circular with a swirl on top.

Although you have information about the proportion of cherry over all chocolates, you have no information about the proportion of cherry among circular chocolates with a swirl (alternately, about possible correlations between cherry and circularity). According to a view of unsharp credences in which they are actually obligatory, it is therefore obligatory to have only an unsharp credence that the chosen chocolate is cherry, contrary to the intuition that the probability is still $\frac{4}{5}$.

However, for a frequentist, this kind of problem is not a difficult edge case, but a fairly routine epistemic situation. Consider the following scenario:

John has been diagnosed with cancer. The five-year survival rate for his cancer is 73%, that is to say, we have a large sample of people with his diagnosis and we know that 73% of them survived over a five-year period. We know also that John’s occupation exposes him to high concentrations of a chemical that is suspected to be carcinogenic. Unfortunately, we do not have a robust sample for people with his diagnosis who were exposed to the chemical; we only have three case histories of individuals, all of whom survived for at least five years. What is the probability that John will survive for five years?

²A paradigm case for such a proposition, one with an unambiguous, knowable truth value, but concerning which the typical person has complete ignorance: “Charibert I of France reigned before Chilperic I.”

Although the two settings are not precisely parallel, they are similar in a key respect: we have a background reference class with a known frequency (random draws of chocolates from the box, patients with a particular diagnosis) and a more specific reference class with an unclear frequency (random draws that result in a circular chocolate, patients with the diagnosis and the chemical exposure). In order to be able to say anything at all about John, it seems necessary to a frequentist to be comfortable ignoring the more specific reference class and reverting back to the original probability of 0.73 — so too in the case of the chocolate box. Moreover, I disagree with White’s contention [2009, 170] that this kind of reasoning must imply the acceptance of a generalized Principle of Indifference.

White’s “coin puzzle” [2009, 175], the central thought experiment of his paper, has a still more ambitious goal: to show that unsharp credences lead not simply to counterintuitive consequences, but to outright paradox. The coin puzzle is an exciting development in the debate over unsharp credences, and previous responses to it — notably Joyce [2010], Rayo [2011], Dodd [2013], and Bradley [2013] — have done much to advance the conversation. Nevertheless, I think they all miss the most natural resolution of the puzzle, because they do not question a core framing assumption of the debate: that credal assignments, whether sharp or unsharp, can be meaningfully evaluated in the single case. My purpose here is to claim that when the coin puzzle is read through a frequentist lens — i.e., with the idea that any credal assignment must be grounded in a frequency notion — the paradox is dissolved and suspended judgment is vindicated.

2 The coin puzzle

Here is White’s statement of the puzzle:

You haven’t a clue as to whether q . But you know that I know whether q . I agree to write “ q ” on one side of a fair coin, and “ \bar{q} ” on the other, *with whichever one is true going on the heads side* (I paint over the coin so that you can’t see which sides are heads and tails). We toss the coin and observe that it happens to land on “ q ”.

Let P denote your credence function before the reveal (i.e., before seeing q or \bar{q} come up), and P' your credence function afterwards. Let H denote the event that the coin lands heads. White notes that the following statements are jointly inconsistent:

1. $P(q)$ is indeterminate, i.e., before seeing the flip, you have no precise credence that q . (One natural formalization of this state as an unsharp credence is $P(q) = [0, 1]$. This can be read as “my credence in q is somewhere between 0 and 1.”)
2. $P(H) = \frac{1}{2}$, i.e., before seeing the flip, you have a precise credence of $\frac{1}{2}$ that the coin will land heads.

3. $P'(q) = P'(H)$. This should be true because after seeing the flip, q is true if and only if the coin landed heads.
4. $P(q) = P'(q)$. This should be true because seeing the flip provided no information about whether q is in fact true. (Note that this would be false for a biased coin.)
5. $P(H) = P'(H)$. This should be true because seeing the flip provided no information about whether the coin landed heads. (Note that this would be false if you had meaningful information about p , in particular a sharp credence of anything other than $\frac{1}{2}$.)

Put these together and we derive $P(q) = P'(q) = P'(H) = P(H) = \frac{1}{2}$, contradicting claim 1. White concludes that we should deny claim 1, instead applying the Principle of Indifference to obtain a sharp probability of $\frac{1}{2}$ for q . That is to say, we should set $P(q) = P'(q) = \frac{1}{2}$.

3 A dialogue

Imagine the following dialogue between Simplicius, a frequentist, and Salviatus, who agrees with White:

Simplicius: This coin puzzle is vexing. But in the end, I don't actually know anything about the proposition q — certainly I don't know any frequencies for how often q_i might be true in an iterated game. So I'm going to suspend judgment about $P(q)$.

Salviatus: OK, but unsharp credences are bad, because they cause *dilation*, that is to say, denial of claim 5 instead of claim 1. If your $P(q)$ is the unsharp interval $[0, 1]$, you must then update $P'(H)$ to $[0, 1]$ as well, instead of its previous value of $\frac{1}{2}$. This is paradoxical because receiving new information has seemingly destroyed some of your knowledge.

Simplicius: I'm not necessarily committed to representing my suspended judgment in this way. In particular, I don't think that the reveal makes me *forget* that half of all coinflips land heads. This fact might constitute some other kind of knowledge than a Bayesian credence about any individual flip.

Salviatus: That's all very well, but how would you bet? Will you say, via some kind of ad-hoc rule, that your $P'(H)$ is $\frac{1}{2}$ regardless? This seems unprincipled at best and inconsistent at worst.

Simplicius: OK, as far as *behavior* goes, I'll bite the bullet. I will bet as though I were using the rule of maximin expected utility, or MMEU [Gärdenfors and Sahlin, 1982], over a dilated credence interval $P'(H) = [0, 1]$, which will effectively prevent me from buying or selling any bets on H after the reveal. When I do this, I can't be Dutch Booked, nor will I miss out on

any *certain* gain.³ So you shouldn't be able to say that I'm irrational.

Salviatus: White [2009, 179] calls this “conservative betting”, and it's still irrational. If you play an iterated version of the coin game, in which a distinct unknown proposition q_i is used in every round, and you're offered bets on H after the reveal at a price of $\frac{1}{3}$, you won't be able to buy them. However, these bets are clearly advantageous — even though none of these bets individually gives you a certain gain, over the long run buying them will almost surely earn you a lot of money.

Simplicius: Well, I'm glad that you seem to be admitting that there's nothing wrong with my *single-case* dilation. But I would, in fact, prefer to win money in this long-run game — so I'll concede that if dilation and MMEU forbid me to buy these bets under any circumstances, then they must not capture the ideal behavior. Let's ask the question directly: if I want to win money in the long-run betting game, what should I do? Having sharp credences seems neither necessary nor sufficient.

4 Analysis

As an extreme example, consider an iterated coin game in which the uncertain propositions q_i are always true. (Let “Reveal” denote the side of the coin that comes up, either q_i for the true side or \bar{q}_i for the false side.)

| | | | | | | | | |
|--------|-------|-------------|-------|-------------|-------|-------------|-------|-------------|
| Flip | H | T | H | T | H | T | H | T |
| Truth | q_1 | q_2 | q_3 | q_4 | q_5 | q_6 | q_7 | q_8 |
| Reveal | q_1 | \bar{q}_2 | q_3 | \bar{q}_4 | q_5 | \bar{q}_6 | q_7 | \bar{q}_8 |

Before the reveal, only one reference class is available for H , the class of all coinflips. The relative frequency of heads in this class is $1/2$. Moreover, there is no *place selection function* [Salmon, 1971] that picks out a subclass with a different relative frequency. No matter what you know or how clever you are, you can't predict in advance a subsequence of flips that will have either more or fewer heads than the background proportion of $1/2$.

After seeing the reveal, there are two reference classes for H , the class of all coinflips (frequency $1/2$ as before), and the class of all coinflips that land with the q_i side up. This second class is more complicated. In it, H is true iff q_i is true, so $\text{freq}(H) = \text{freq}(q_i)$, that is to say, the relative frequency of true propositions q_i . This frequency is unknown to Simplicius but knowable in principle, which intuitively justifies a picture in which he has $P'(H) = [0, 1]$. For example, in the case given previously the q_i are always true, so the relative frequency of H in this class is 1. Furthermore, the class may admit place selection functions — there might be some subsequence

³MMEU says to value each alternative in a decision problem by its minimum expected utility according to any probability within the agent's credence interval, then choose the alternative that maximizes this quantity. Therefore, MMEU allows arbitrage, e.g., buying a bet on H for 0.4 and another bet on T for 0.4 for an overall certain profit of 0.2, despite neither of these bets being worthwhile isolation according to $P(H) = [0, 1]$.

of them such that someone knows the answers outright, or has calibrated probabilistic knowledge of the answers.

For a frequentist, the denial of claim 5 (the dilation of $P'(H)$ from the sharp $\frac{1}{2}$ to some indefinite value such as $[0, 1]$) is a feature, not a bug; it is the proper response to the appearance of a reference class ambiguity. Moreover, the ambiguity has actual implications for betting behavior in the iterated coin game:

Proposition 1 *The following is a necessary and sufficient condition to make money in long-run betting on H_i in the iterated coin game: you must buy and sell bets within a reference class such that the prices are favorable with respect to the long-run frequency of H in that class. That is to say, you must buy at a price lower than the frequency, and sell at a price higher.*

The claim is in essence trivial, but proving it requires choosing an interpretation of the term “frequency”. For example, we could take it to mean an observed frequency over a finite sequence of bets (Hájek [1996] calls these “finite frequencies”), for example buying 10 \$1 bets on heads for \$0.4 each and winning 5 of them. In this setting, the claim is immediate: the condition of having bought or sold at a favorable price with respect to the frequency (0.5) is equivalent to having winnings (\$5) greater than or equal to the price paid (\$4). If we take it to mean the limit of the frequency over an infinite sequence of bets, it is similarly easy to see that the condition is equivalent to having the profit (i.e., winnings minus prices) go to infinity.

It is tempting to interpret “frequency” here as the probability of heads $P(H)$ and then try to prove the claim using probability theory. Indeed, if we do this and model each bet as an independent trial of heads with probability $P(H) = p$, then the condition is equivalent to each bet having expected profit $E[B] = c > 0$, with standard deviation σ . Then the Central Limit Theorem tells us that over a sequence of n bets, the profit will be approximately normal with mean cn and standard deviation $\sigma\sqrt{n}$. Since $\sigma\sqrt{n} \ll cn$ for large n , this means the profit will be positive with probability approaching 1 as n approaches infinity. But this seeming clarification is actually misleading. In the original reference class of all heads, each H_i is independent and this model is applicable. But we are also interested in reference classes of bets produced after the reveal, and in such a class, the H_i may not be independent at all — if the q_i are correlated, and the result of the reveal affects whether the bet takes place, then the H_i may be correlated as well. And this prevents us from providing any guarantees about the rate of convergence of $\text{freq}(H_i)$ in the class to the background frequency. The earlier formulation of the proposition in terms of limiting relative frequency is the best we can do.

Proposition 1 is shallow in that it is only a success condition, not a decision rule for achieving that success condition; it simply describes the situations in which an agent comes out ahead. Nevertheless, it illuminates Simplicius’s question of how to bet — a betting strategy should be judged by its ability to bring about this success condition. Armed with this result, Simplicius can continue the debate.

5 Resolving the paradox

Simplicius: You're committed to the Principle of Indifference for $P(H)$, combined with the standard betting interpretation of that credence. Therefore, after the result of the flip is revealed, you will *buy and sell* bets on H at a price of $\frac{1}{2}$.⁴

Salviatus: Correct.

Simplicius: Actually, according to our analysis, this seems to get you into trouble. Let's say a third party (call him Sagredus) actually knows the truth values of the q_i , and you are committed to *buying and selling* bets at a price of $\frac{1}{2}$ after the reveal, every time. Then Sagredus can buy from bets on heads from you in exactly the cases where q_i comes up on the coin and he knows that q_i is true; he will then be buying at a price of $\frac{1}{2}$ in a reference class with a frequency of 1, and you will lose money on every bet. Moreover, if Sagredus has calibrated probabilistic knowledge for any identifiable subsequence of the q_i , you are still exposed to long-run loss. For example: there's a stream of French history questions, Sagredus remembers enough French history to answer 70% of them correctly, and furthermore correctly estimates his calibration at .7. Then, whenever the coin lands q_i on a French history question and Sagredus has a hunch that \bar{q}_i , Sagredus buys a bet on tails for 0.5 (favorable according to his calibrated degree of belief $P(\bar{q}_i) = 0.7$), and makes money in the long run according to our analysis. But Sagredus won't be able to take advantage of me, because I'm not allowed to buy and sell bets with him after the outcome of the flip is revealed.

Salviatus: This is interesting, but it seems like you're changing the subject. The Principle of Indifference is intended as a constraint on your internal belief states; after all, White explicitly disclaims [2009, 163] that it is supposed to inform us about objective probabilities. Certainly it can't account for the information other people may or may not possess! So this doesn't seem to show that I'm irrational, only that I can be exploited by agents who know more than I do, which was obvious to begin with.

Simplicius: But our arguments are symmetric: just as your argument against me required an agent who was mysteriously misinformed about $P(H)$ (and thus selling bets on H at a price of $\frac{1}{3}$), my argument against you requires an agent who is better-informed than you about $P(q_i)$. Which of these is more realistic in practice? While your hypothetical agent is a fanciful dullard, it seems clear that you should be aware of the possibility that mine might exist. After all, by your own admission, you know nothing about the q_i , so you should be aware that someone might know *something* about them.

⁴White's argument seems to elide this aspect of the betting interpretation, so it should be emphasized that all standard accounts of the betting interpretation commit the agent to actually making book, not just buying and selling bets at will from other bookies. Otherwise, how could the agent ever be Dutch Booked? Of course, if it is *compulsory* for the agent to buy and sell when given an offer favorable according to his degrees of belief, the two are equivalent.

Salviatus: All right, but now it seems like we're even. I can exploit an ill-informed agent, but I can be exploited by a well-informed one. You can neither exploit nor be exploited. So which of us is better off? It seems like a question of taste.

Simplicius: Actually, we now have an exact characterization (proposition 1) of how to win this long-run betting game. And according to this characterization, I can exploit the misinformed agent (call him Duncie) without being exploited by Sagredus. Notice that if Duncie is guaranteed to offer me a bet on H for $\frac{1}{3}$ *every time* after the reveal, then this is in some sense logically equivalent to him offering the bets before the reveal — at which point I have not yet dilated and I'm allowed to buy them. Here's the precise form of this argument: if Duncie offers the bet every time, then he is unable to manipulate the reference class in which I buy and sell the bet to make $\text{freq}(H)$ anything other than $\frac{1}{2}$. But if he can decide whether or not to offer the bet after the flip, then he can manipulate me into long-run loss after all, by offering the bets only when he has calibrated probabilistic knowledge that $P(q_i) < \frac{1}{2}$.⁵

Salviatus: To be clear, this betting behavior you're endorsing based on proposition 1 no longer seems to coincide exactly with MMEU over $[0, 1]$.

Simplicius: That's correct. Suppose Duncie doesn't offer the bet every time, but only exercises his ability to withhold the bet to a very slight degree — say, $\frac{1}{100}$ times over the class of all flips. Thus, his post-reveal offers are not *logically* equivalent to pre-reveal offers, and he might in fact be using knowledge of the q_i to his advantage. However, since the frequency of heads over the class of offered bets must still lie in the interval $[\frac{1}{2} - \frac{1}{100}, \frac{1}{2} + \frac{1}{100}]$, I know that he hasn't been able to skew the reference class in his favor relative to his price of $\frac{1}{3}$, and the bets are still worthwhile for me to buy. I can't see a way to understand this behavior using MMEU, because it seems necessary to view the bets as being offered and evaluated post-dilation, and yet they are worthwhile even under maximal $[0, 1]$ -unsharpness about the true frequency of q_i . They are worthwhile because of the frequency characteristics of the way Duncie is offering them, and this information is outrunning our information about each bet in the single case.

Salviatus: Your behavior seems paradoxical because you value the same bet on H differently before and after the reveal, even though the reveal didn't inform you about H at all.

Simplicius: Yes, in some sense bets are only valuable to me in virtue of their long-run properties. This is a common argument against frequentism and it needs a reply, although this isn't the time. Again, we're discussing

⁵An analogous argument also dismisses White's charge [2009, 180] that if the dilating agent manages to buy a bet on H for $\frac{1}{3}$ before the reveal, he will wish to revoke or cancel it afterwards. Since until the reveal, the flips admit no place selection functions, this bet is guaranteed to be in a reference class with $\text{freq}(H) = \frac{1}{2}$. Thus the bet is still worthwhile.

the question of how to win a long-run betting game, so it seems only natural that the answer should hinge on questions of long-run behavior.

Salviatus: It also sounds like your betting behavior is coming apart from the epistemic justification of dilation. In particular, if the outcome of the flip is not revealed to you, but only to Sagredus, you should still “dilate” in a behavioristic sense, that is to say, you cannot buy and sell bets with him. But it seems like your own epistemic state didn’t change at all.

Simplicius: This is accurate, and yes, there’s some tension between my view and standard accounts of unsharp credences, such as the “committee” model described by Joyce [2010]. If, in the case you describe, I can in fact be represented as having an unsharp credence $P(H) = [0, 1]$ after the reveal, then the credences in the interval may not necessarily represent my own epistemic states, but rather the epistemic states of possible betting adversaries such as Sagredus and Duncie.

Salviatus: Again, it seems like we’ve drifted off topic! This talk of betting adversaries is irrelevant to the problem that the Principle of Indifference purports to solve: whether your own credal states, given the information that you yourself possess, are rational.

Simplicius: I see the concern. But if that’s the only question at stake, then I don’t think I’m committed to answering it; after all, I hold that single-case credence in general is a pseudoproblem. So I think it’s significant that when you tried to convict me of irrationality, the scenario you tried to use against me was not only behavioral (concerned with my betting behavior rather than my internal states), but moreover involved *long-run* betting behavior. Now we’ve obtained an exact characterization of how not to lose these long-run betting games, and it validates my position and is unfavorable to yours. So either not losing at these games is important, in which case my position is superior, or it’s not important, in which case you have no grounds to criticize me.

Salviatus: Your argument against me still depends on the existence of Sagredus, who is better-informed about the q_i than I am. And this still seems unfair. What if, as White [2009, 182] suggests in a related scenario, “none of us have any clue” about the q_i , so Sagredus can’t exist? If the irrationality of the Principle of Indifference is justified by imagining adversaries with superior knowledge about the q_i , would perfect universal ignorance about the q_i make it rational again?

Simplicius: In fact, if someone knows that the relative frequency of q_i is any value other than $\frac{1}{2}$, then they already know enough to win in the long run against you. (For example, suppose they know that the limiting relative frequency of true q_i is $\frac{1}{3}$; then they should sell you a bet on H every time the q_i side comes up.) This would seem to imply that a necessary condition for “no one to have any clue” is that the limiting relative frequency of

q_i is exactly $1/2$. An exact characterization of perfect universal ignorance would most likely be “the q_i are an algorithmically random (i.e., Martin-Löf random) sequence with limiting relative frequency $1/2$ ”. (In particular, Martin-Löf randomness of a sequence is equivalent to the impossibility of any computably enumerable betting scheme making money by betting on it.) So if we knew that the q_i were algorithmically random, that would in fact justify betting on them at a price of $\frac{1}{2}$ before and after the reveal. But how could we ever know that this is the case? It seems as though in practice, we always have to be on our guard against an epistemically privileged agent like Sagredus — and a necessary condition for this is the suspension of judgment about $P(q_i)$.

Acknowledgements

I am greatly indebted to Lara Buchak for her comments on an earlier version of this argument.

References

- Seamus Bradley. Painted coins and imprecise credence. Circulated manuscript, 2013.
- Dylan Dodd. Roger White’s Argument against Imprecise Credences. *The British Journal for the Philosophy of Science*, 64(1):69–77, 2013.
- Adam Elga. Subjective probabilities should be sharp. *Philosopher’s Imprint*, 10(5), 2010.
- Peter Gärdenfors and Nils-Eric Sahlin. Unreliable probabilities, risk taking, and decision making. *Synthese*, 53(3):361–386, 1982.
- Alan Hájek. “Mises Redux” — Redux: Fifteen Arguments Against Finite Frequentism. *Erkenntnis*, 45(2-3):209–27, 1996.
- Alan Hájek. Fifteen arguments against hypothetical frequentism. *Erkenntnis*, 70(2):211–235, 2009.
- James M. Joyce. A defense of imprecise credences in inference and decision making. *Philosophical Perspectives*, 24(1):281–323, 2010.
- Agustin Rayo. A puzzle about ineffable propositions. *Australasian Journal of Philosophy*, 89(2):289–295, 2011.
- Wesley C. Salmon. *Statistical Explanation & Statistical Relevance*. University of Pittsburgh Press, 1971.
- Roger White. Evidential symmetry and mushy credence. In *Oxford Studies in Epistemology*. Oxford University Press, 2009.