<u>**CREDENCE AND SYMMETRY**</u>

Frank Arntzenius

SOME UNFINISHED THOUGHTS!

It was once hoped that the problem of induction could be solved by figuring out what the unique logical probabilities are which determine the credences that rational people should have both prior to the acquisition of any evidence, and conditional upon evidence. This Carnapian program is generally perceived as having failed, although versions of this program (objective Bayesianism, Williamson's 'evidential probabilities') have made a come-back recently. Today I want to discuss the central problem of this program, namely that of dealing with symmetries in infinite cases. I am going to draw pretty much the opposite conclusion from the one that usually is drawn. The conclusion that usually is drawn is that in infinite cases rational credences are underconstrained: there are lots of different equally good candidates for rational credences in the infinite case. I think this is a mistaken conclusion: rather, in infinite cases symmetries often over-constrain rational credences, i.e. sometimes there are no probability distributions that fit all the symmetries that one could reasonably impose. I am going to conclude that sometimes the epistemic states of rational people can not be standard probabilities, or even probabilities at all.

Let me start by noting that one can not dismiss cases in which one has credences distributed over infinitely many possibilities, on the grounds that infinities are weird and pathological. Here is why not. The basic idea for figuring out rational credences on the Carnapian program is an indifference principle. For instance, if your evidence is compatible with Harry being in Pittsburgh and Harry not being in Pittsburgh you should be indifferent between the two possibilities, so your rational credence in Harry being in Pittsburgh should be 0.5. Of course the problem now is that you also don't know whether Harry is in Cleveland or not, so by indifference you should have credence 0.5 that he is in Cleveland, and you don't know whether he is in Cincinatti or not, so by indifference you should have credence 0.5 that he is in Cincinatti. But Pittsburgh, Cleveland and Cincinatti are mutually exclusive possibilities, so we have arrived at an inconsistency. The natural response is that you should distribute your credences uniformly over the most fine-grained division into possibilities: each possible exact location of Harry is equally likely. But of course that will be a space which has infinitely many possibilities, indeed uncountably many. And of course, people have then worried that there are lots of different ways of making each location equally likely, i.e. lots of different probability distributions such that the probability of each exact location is 0. I will return to this issue of underdetermination. In the meantime however, it should be clear that if there is any hope for the program of figuring out rational credences, the case in which there are infinitely many possibilities can not be ignored.

What I am now going to do is present you with a number of cases in which it will turn out to be impossible to have standard probabilities satisfy certain plausible symmetry constraints. I will then tentatively suggest a solution to these puzzles, and attempt to draw some general conclusions. Let me start with a puzzle due to Cian Dorr.

There is an infinitely long road with houses all along it. Every house is either blue or red. If you walk along the road the pattern you see is 3 red, 1 blue, 3 red, 1 blue, ...Dorothy has no idea as to which house she is in. What should her credence be that it is red? The obvious answer is 0.75.

But now comes the twist. Each night Dorothy is switched to a new house. Over time the sequence of colors of Dorothy's lodgings is: 3 blue, 1 red, 3 blue, 1 red, ... Should her credences be aligned with the spatial frequencies or the temporal frequencies?

Here is an argument for not going with the spatial frequencies:

Symmetry: Cr_{Mon}(Red Mon)= Cr_{Tue}(Red Tue)

Reflection: Cr_{Mon}(Red Tue)= Cr_{Tue}(Red Tue)

If you were to got with spatial frequencies then

Cr_{Mon}(Red Mon)=0.75.

So Cr_{Mon}(Red Tue)=0.75.

So Cr_{Mon}(Red Mon & Red Tue)>=0.5.

But the temporal sequence is 3 blue, 1 red, 3 blue, 1 red, ...

So Cr(Red Mon & Red Tue)=0.

So credences can not correspond to spatial frequencies.

So, plausibly, credences should correspond to temporal frequencies.

However, here is another twist. Suppose that every house is always filled with exactly one person, and that every person is switched in such a way that each person spends 3 consecutive nights in a blue house, then 1 night in a red house, then 3 nights in a blue house, and so on. Now we can give an argument against setting one's credences in accord with the temporal relative frequencies.

Suppose that Dorothy and Mary are neighbours. Then

Symmetry: Cr_{Dor}(Dor Blue)= Cr_{Mary}(Mary Blue)

Solidarity: Cr_{Dor}(Mary Blue)= Cr_{Mar}(Mary Blue)

If one were to go with the temporal frequencies then Cr_{Dor}(Dor Blue)=0.75.

So Cr_{Dor}(Mary Blue)=0.75.

So Cr_{Dor}(Dor Blue & Mary Blue)>=0.5.

But the spatial pattern is 3 red, 1 blue, 3 red, 1 blue, ..

So Cr_{Dor}(Dor Blue & Mary Blue)=0.

So credences can not correspond to temporal frequencies. Now what?

Let's try to get a better grip by going through some other puzzles. Consider the following variation on Cian Dorr's puzzle. There are infinitely many houses on a road. Each house is either red, blue or green. Each house contains three walkie talkies: a red one, a blue one and a green one. Each walkie talkie connects to a unique other walkie talkie of the same color. For any connected pair of walkie talkies, one of them is in a house that has the same color as the walkie talkie, one of them is in a house of a different color. Now suppose that Dorothy talks to Rosie on a red walkie talkie, to Belle on a blue one, to Grace on a green one. Then:

Symmetry: Cr_D (Dor red)=Cr_R(Ros red).

Solidarity: Cr_D(Dor red)= Cr_R(Dor red)

So Cr_D (Dor red)=Cr_R(Ros red)=1-Cr_M (Dor red)=1-Cr_D (Dor red)=1/2.

Exactly the same reasoning applies to the other two colors.

So Cr_D (Dor blue)=½, and Cr_D (Dor green)=½.

But the three color are mutually incompatible. So this is impossible. So

either Solidarity fails or Symmetry fails, or epistemic states are not probabilities.

Now, in this example we did not need Reflection to get a problem. Symmetry seems a very plausible principle. So one might suggest that Solidarity should be abandoned. Let me now give an example which suggests that if that is so Reflection should also be abandoned.

Suppose Dorothy lives for ever, but every night she switches houses. The houses come in three colors: red, blue, green. Every night she finds three letters in her house: one on red paper, one on green paper, one on blue paper. The notes say: “There exists exactly one identical copy of this note. You have received it, or will receive it some other night.”

Now, If Dorothy satisfies Symmetry and Reflection, then her credences can not satisfy the axioms of probability theory. The point here is that Reflection just Solidarity with respect to her future/past selves, so that if Solidarity has to go, then Reflection will have to go too.

Option 1: Abandon either Symmetry, or Reflection and Solidarity, or all three.

Option 2: Abandon the idea that epistemic attitudes satisfy the axioms of probability. But what instead?

Let’s try some more puzzles before discussing which option is best. The following puzzle is due to Tim Williamson. Suppose one has two fair coins A and B, and a machine is about to start tossing them. You know the machine is going to toss them infinitely often, in Zeno-like fashion, so that all the tosses will take place in a finite time. You are offered 3 contracts. What are your preferences among the following contracts

1) Eternal life if A always lands heads

2) Eternal life if B always lands heads

3) Eternal life if A from its second toss on always lands heads

Here is some apparently very plausible reasoning that leads to incoherence. Contract 3 pays out in strictly more possible cases than contract 1, so contract 3 is preferable to contract 1. Contract 1 and 2 are equally good since the map that maps toss n of coin A onto toss n of coin B preserves all probabilities. Similarly contracts 2 and 3 are equally good since the mapping that maps toss n+1 of A to toss n of B preserves all probabilities. Imagine, e.g. that the first toss of B is done at the same time as the second toss of coin A, and from then on they are tossed simultaneously.

Let me put the puzzle a different way. It would seem that one should have the same credence in A always landing heads as in B always landing heads. It also seems that one should have the same credence in B always landing heads as in A always landing heads after toss 2, since these tosses are all going to occur simultaneously, and each toss has the same probabilities of outcomes. But it also seems that if one's credence in A landing heads from toss 2 onwards is ϵ then one's credence A always landing heads has to be $\frac{1}{2}\epsilon$. But these three claims are jointly inconsistent unless $\epsilon=0$. Now, one might think that the proper conclusion indeed is that one's credence in any of these infinite sequences of outcomes should be 0. But then one's epistemic attitude would not encode the epistemically important fact that the one proposition is true in strictly more possible cases than the other.

Let me now turn to a related puzzle. Suppose we are going to throw a dart at a dartboard, and we are aiming at the center of the board. Suppose that there is a line that runs from the center of the dartboard straight up (vertically), and that we are interested in the angle between that line and the straight line drawn from the center of the dartboard to the location that the dart lands on. Imagine that we, intuitively speaking, think each angle is equally likely. So, for instance, our credence that the dart will land at an angle between 45 and 90 degrees is 0.25 (this being 1/4 of the total range of angles.) (Let's not worry as to how we would classify a throw that lands exactly on the center.) What credence should we have that the dart will land at some completely precise angle, say 22 degrees exactly? One might be inclined to say, simple, one's credence in that has to be 0. However, it seems to me that the inclination to say so is because one has been indoctrinated by the formalism of probability theory. One realizes that any non-zero number is going to lead to a violation of the axioms of probability, since by uniformity of one's credences over all possible exact angles, and additivity of probabilities, any other number would imply that for any finite number n one could find a finite set of possibilities such that one's credence that the angle would be in that set would exceed n, and that violates the axioms of probability theory (not to even mention that the probabilities of infinite sets would have to be infinite.)

However, there are a couple of reasons not to like the idea that one should set one's credence in exact angles equal to 0. One motive not to do so, expressed by David Lewis and many others, is that one thinks that probability 0 should be identified with impossibility. The idea is something like this: probability 0 means that there is no chance that it could happen, that one is certain that it will not happen, i.e. that one judges it to be impossible. As it stands I am not very sympathetic to this line of reasoning. A perfectly good response would be: Uh, no, sorry, probability 0 does not mean that it is impossible. As to whether probability 0 means that one is certain that it will not happen, that depends on what one means by 'certain'. If certainty means 'credence 1' then yes, but then certainty does not imply a judgment of necessity. On the other hand if certainty implies a judgment of necessity, then credence 1 does not imply certainty.

Here is a line of reasoning that I find a bit better. There are no credences smaller than 0, so if one represented epistemic attitudes just as credences one could not distinguish between things in which one has credence 0 and things that one thinks impossible. But rational people should make such distinctions. This, I think, is a good argument, but as it stands it only implies that one's epistemic attitude to a proposition is not fully characterized by one's credence in it, one needs to add a little tick or a little x attached to a box labeled "possibly p". Moreover, one could argue that this part of one's attitude should not be considered part of one's epistemic attitude towards p, rather it should fall than under one's epistemic attitude towards a different proposition, namely the proposition "possibly p". So, I am still not convinced by this argument. Here is what I find a much more convincing argument.

Surely one would want to say it is more plausible that the dart will land at exactly 22 degrees than that it will land either at exactly 22 degrees or at exactly 44 degrees. And in so far as one initially thought it plausible that one's epistemic attitudes satisfy the axioms of probability theory, one presumably will find it plausible that such judgments about the plausibility of various possibilities, once one has restricted attention to a particular set of possibilities (conditionalised on a particular set of possibilities), should also satisfy the axioms of probability theory.

A very natural way to do this is to allow credences to take infinitesimal values. Now, I whole-heartedly concur with this motivation to do something about the standard representation of one's epistemic states. But I will argue that the move to infinitesimal probabilities was not exactly the right move. Let me parenthetically say, that some have argued against the use of infinitesimals on the grounds that they violate countable additivity and hence one is subject to a countable Dutch book if one has infinitesimal credences. I do not consider that to be a good argument against the adoption of infinitesimal credences. We could go over this in discussion time, but the short version of my argument is that the countable Dutch book argument relies on the assumption that if one finds each of a countable collection of offered bets a good deal, then one should find the whole collection a good deal. However, I claim that in the infinite case this is an invalid inference, so I reject the argument against infinitesimal credence from countable

Dutch books. Note also that violations of countable additivity lead to non-conglomerability, where this means that the unconditional probability can have a value outside the range of values of the probabilities conditional upon all the elements of a partition. Non-conglomerability is odd, but non-conglomerability occurs more generally for continuous distributions, so will presumably have to be dealt with anyhow.

No, I think there is a different reason to reject (unique) infinitesimal credences, the basic reason being that infinitesimal credences can not satisfy certain very plausible symmetries. Here is what I mean. One can show that there exists no probability distribution which attributes non-zero, infinitesimal, probability to every point on a circle which is also invariant under rigid rotations. It is in fact fairly easy to see why this is the case.

Suppose we have a circle of radius 1, which therefore has a circumference of length 2π. Consider the following countable set of points. Point 1 is the point that I get to when I start at the top and go distance 1 along the circumference in a clockwise direction. Point 2 is the point that I get to when I take another step of distance 1 in the clockwise direction. And so on. Now consider the collection all such points for all finite n. This is a countable collection. Suppose each such point has infinitesimal probability ϵ. What is the probability of the set? Well, intuitively it is the sum of countably many ϵ's. However, that sequence does not converge. Still, we know it has to be some infinitesimal number λ which is an order of magnitude larger than ϵ (i.e. for any finite integer n, $n\epsilon<\lambda$) and (at least) an order of magnitude smaller than 1 (since one can break up the whole circle into uncountably many disjoint subsets each of which is just a rotated version of this set). Now, think of what happens when we rotate this whole set 1 unit distance clockwise. Point 1 will rotate onto point 2, point 2 onto point 3, more general point n will rotate onto point n+1.

So once we have rotated it, the rotated set will be exactly equal to the original set minus point 1. By finite additivity the probability of the rotated set must be λ-ϵ, while the probability of the original set was λ. So the probability distribution can not be strictly invariant under rotations. One can prove that it can be so ‘up to infinitesimal differences’. (Note however that, as we will see in a minute, there is a relative sense in which ‘infinitesimal differences’ can be as large as you like.) In short, if you allow infinitesimal probabilities, then these probabilities can not satisfy certain extremely natural symmetries. And, since I believe that rational people should, at the very least, be allowed to hang on to these symmetries, they should be allowed to resort to something other than infinitesimal credences.

First let me clarify my remark that even though the invariance under rotations can be up to an infinitesimal, nonetheless, in a relative sense, the breaking of the invariance can be as large as one likes. Here is why I said this. One can show that there exists a set of points S in a Euclidean infinite plane, such that if one rigidly shifts S one unit distance to the right, one gets set TS, and if one rotates S one unit counterclockwise one gets set RS, where the intersection of TS and RS is empty and the union of TS and RS equals S. Now, by additivity the sum of the probability of TS and RS must equal the probability of S. And this means that if S has non-zero probability then translation and rotation invariance is broken in a pretty bad way: the difference between the probabilities of either S and TS, or of that between S and RS has to be at least ½ the probability of S: if you either translate S or rotate S rigidly, you lose at least half of its probability! Now, if S only has infinitesimal probability that means you only lose an infinitesimal amount, so we still have invariance under rigid transformations ‘up to an infinitesimal’, but in a relative sense this loss is pretty big: you are losing at least half the probability merely by shifting or rotating it!

[Here is how to do the above. Start with a point at the origin of the Euclidean plane. Consider the following two operations: step one unit to the right, rotate one radian counterclockwise, i.e. multiply by e^i=exp(i). S is the set of points that one can reach by any finite sequence of such operations from the origin. Moving z one step to the right: z->z+1. Rotating it one radian z->e^i z So each point in S is of the form $\sum n_j \exp(im_j)$ where n_j and m_j are integers (note that not all points of this form are in the set). Now define set TS: shift S one unit to the right. RS: rotate S by one radian. Clearly each point in the set can be reached either by taking a step to the right from a point in the set, or by rotating a point that is in the set (since the original set is generated from the origin in that way, and the origin itself can be reached by rotating the origin). So the union of the RS and TS equals S. So all we need to show is that TS and RS have no points in common. Well suppose they did. Then there would be some point such that p=Rx and p=Ty where x and y are in S. Since p=Ty it corresponds to an expression of the form

$$n_1+n_2\exp(im_2)+n_3\exp(im_3)+\ldots\ldots\ldots \quad =n_1+n_2(\exp(i))^{m2}+n_3(\exp(i))^{m3}+\ldots.$$

Since p=Rx it corresponds to an expression of the form

$$k_1\exp(il_1)+k_2\exp(im_2)+\ldots\ldots\ldots. \quad =k_1(\exp(i))^{l1}+k_2(\exp(i))^{l2}+\ldots.$$

Since there are different expressions for the same point we should have:

$$n_1+n_2(\exp(i))^{m2}+n_3(\exp(i))^{m3}+\ldots.=k_1(\exp(i))^{l1}+k_2(\exp(i))^{l2}+\ldots.$$

So subtracting the right hand side from the left hand side should give 0. But subtracting the one from the other gives a polynomial equation (with integer coefficients and exponents, and not all coefficients are 0 since n_1 is not equal to 0), i.e. an algebraic equation with exp(i) as its solution.

But exp(i) is transcendental so it is not the solution to some algebraic equation.]

One more example and then I will suggest answers to the puzzles. The last example concerns unbounded spaces. Suppose you think the universe is infinitely large. You might still want to consider questions such as “How likely do you think that the brightest star is within 1 million light-years of us”, or “If it is withing 1 million light years, how likely do you think it is within 0.5 million light-years?” And you might think a rational person should be allowed to have a credence that is uniform over all of space. Well, if space is unbounded, and if credences are not allowed to be infinitesimals, one can not do this. To be precise: suppose that one has a notion of rigid transformation such that one can break up the space into countably many disjoint regions which are rigid transformations of each other. Then one can not have a countably additive probability distribution over this space which is uniform relative to this notion of rigid transformation.

[There is, of course, within physics a standard way of representing something like a uniform credence, namely by the so-called Lebesque measure. A measure satisfies the same axioms as probabilities, except that the measure of the entire space of possibilities can be infinite, rather than 1. (If it is any finite number one can always renormalize it to 1, so the gain in generality amounts just to allowing the value ‘infinity’.) If one does this, then measure conditional upon finite size subsets of the total space will be well-defined.]

OK, enough examples. What are the morals? Well, I suggest that rational people, at the very least, should be rationally permitted to have epistemic attitudes that are invariant under all of the symmetries of my examples. So, one must allow rational people to have epistemic attitudes that are not standard probabilities. I have no general recipee for what structure epistemic attitudes should be allowed to have, but I have some suggestions for each of the examples that I have given.

Let me start with the dart on the unit circle example. Let's start with some non-standard probability distribution, one which includes infinitesimal probabilities, and then impose rotation invariance by brute force. That is to say, let one's epistemic state be represented by an equivalence class of non-standard probability distributions [ρ(x)], where if non-standard probability distribution ρ(x) is in the equivalence class, then so is $\rho^{\alpha}(x)$, where $\rho^{\alpha}(x)$ is the non-standard probability distribution that you get when you rotate ρ(x) by angle α. To be precise the rotated probability distribution $\rho^{\alpha}(S)$ is defined by the following: $\rho^{\alpha}(S)=\rho(S^{-\alpha})$, where $S^{-\alpha}$ is the set you get by rotating set S over angle minus α. (I am neutral as to whether one should close this set up under convex combinations.)

Now, what this does is that it makes certain probabilities unique, and others highly non-unique. Let me give examples. If one start with a non-standard probability distribution which attaches infinitesimal probability ϵ to each individual point, then, since rotations put points on points, each $\rho^{\alpha}(x)$ will attach that same value ϵ to each point, so the probability that one attaches to each individual point is unique. However, now consider the set of points 1,2,3,.... that I constructed in order to explain why any particular non-standard measure can not be fully rotation invariant. It's measure will change when one rotates the measure by various steps, so that one will now get a set of measures, which differ by $n\epsilon$ for all finite integers n. So the probability that attaches to this set will be highly non-unique, but it will have a unique order of magnitude.

How about conditional probabilities? Well, whenever the unconditional probabilities are the same for all the probability distributions in the set, so, of course, are the conditional ones. However, the conditional ones can also be unique if the unconditional ones change in such a way as to preserve their ratios. Consider say two distinct sets S1 and S2 that are generated by taking one point and then taking all the points that you get by moving one step to the right. One could have $\rho(S1/S1 \text{ or } S2)=0.5$ and this could be preserved under all rotations of $\rho(x)$.

Consider another example: consider such a set S, and then consider the subset SE that consists of the generating point and all the points that are an even number of steps 'to the right' of the generating point, and the subset SO consisting of the points that are an odd number of steps to the right of the generating point. Suppose $\rho(SE)=\rho(SO)=\mu$. Then $\rho(SO/SE \text{ or } SO)=0.5$. Now consider the probability measure ρ^{-1} that we get by rotating ρ one step 'to the left'. Then $\rho^{-1}(SE)=\rho(SE^1)=\rho(SO)=\mu$, and and $\rho^{-1}(SO)=\rho(SO^1)=\rho(SE \text{ minus the generating point})=\mu-\epsilon$. So $\rho^{-1}(SO/SE \text{ or } SO)=(\mu-\epsilon)/(2\mu-\epsilon)=\mu/(2\mu-\epsilon)-\epsilon/(2\mu-\epsilon)$ which is within an infinitesimal (of order ϵ/μ) of 0.5. More generally, $\rho^n(SO/SE \text{ or } SO)$ will differ at most infinitesimally from 0.5.

This last fact also suggests a slightly different approach to the problem of rotation invariance. Rather than using infinitesimal probabilities, and making one's epistemic state rotation invariant by assuming that one's epistemic state corresponds to an equivalence class of such non-standard probability distributions which are related by rigid rotations, we could assign probability 0 to sets of Lebesque measure 0, but allow conditionalisation on sets that have Lebesque measure 0. I.e. we could allow Popper-Renyi primitive conditional probabilities, while demanding that these conditional probabilities be invariant under rigid rotations.

Let me now say a little bit more about the relation of epistemic states and decision theory. One might worry how to do decision theory when one's credences correspond to equivalence classes of probability distributions rather than a single probability distribution. The natural way to generalise decision theory to such a circumstance is to say Act A is rationally mandated when you have dominance, i.e. when expected utility EU(A) is maximal for all probability distributions in one's equivalence class [ρ]. But, how about one's preferences among two acts when it is neither true that EU(A)>EU(B) for all ρϵ [ρ], nor true that EU(A)<EU(B) for all ρϵ [ρ], nor true that EU(A)=EU(B) for all ρϵ[ρ]? Well, then neither act will be rationally mandated, and either is permitted. But one should be careful not to say that in such a case one is indifferent between the two acts, at least not if one wants indifference and preference to satisfy the usual (transitivity) axioms. Here is why.

Consider the coin tossing case. Suppose that one wants to have an epistemic attitude towards the coin tosses which is invariant under rigid translations of coin toss results in time. So, one wants it to be the case that one's epistemic attitude towards any finite, or infinite, sequence of results, such as ....HTHHT....., is the same no matter at what sequence of times the outcomes HTHHT occur. It follows that one's epistemic attitude towards "Heads only from toss 1 of coin A onwards" has to be the same as one's epistemic attitude towards "Heads only from toss 2 of coin A onwards". A way to ensure this is to start with some non-standard probability distribution ρ over infinite sequences of coin tosses which assigns, say, infinitesimal probability ϵ to heads from toss 1 onwards, and then generate an equivalence class [ρ] of probability distributions by including all the probability distributions that one gets by finite shifts of ρ. (I.e. start with a non-standard probability distribution ρ over all sets of sequences of outcomes S that are infinite in both directions of time. Then generate the new ones ρ^n by stipulating $\rho^n(S)=\rho(S^{-n})$, where S^{-n} is what you get when you shift sequence S, or set of sequences S, by −n places, where n is finite positive or negative integer.)

Now let me consider the two-coin case. One might start with a non-standard probability distribution according to which coin A and B are independent (i.e. a product distribution) which is moreover such that ρ(coin A has outcomes.....HTTH....)=ρ(coin B has outcomes.....HTTH.....), when the outcomes HTTH occur at exactly the same times for A and B. And one might demand translation invariance, where one allows different translations for coin A and coin B. I.e. one might generate one's equivalence class of probability distributions by saying that every probability distribution of form ρ^{nm} must in the equivalence class where $\rho^{nm}(S_A\&S_B)=\rho(S_A{}^{-n})\rho(S_B{}^{-m})$. Now, for all n,m it will be the case that ρ^{nm}(A heads from toss 2 on)>ρ^{nm}(A heads from toss 1 on). But the relative magnitude of ρ^{nm}(A heads from toss 1 on) and ρ^{nm}(B heads from toss 1 on) will vary, and so will the relative magnitude of ρ^{nm}(A heads from toss 2 on) and ρ^{nm}(B heads from toss 1 on). Now recall the three contracts:

1) Eternal life if A always lands heads

2) Eternal life if B always lands heads

3) Eternal life if A from its second toss on always lands heads

Suppose that one were to infer from the non-invariance of the relative magnitudes to indifference. Then one would be indifferent between contracts 2 and 1 and between contracts 2 and 3. But one would prefer contract 1 to contract 3. So one would violate transitivity of preferences. So I suggest that non-invariance of relative probability magnitude corresponds not to epistemic indifference, but rather to an absence of an epistemic preference, i.e. I am suggesting that epistemic preferences, and prudential preferences, can be incomplete.

Next let me consider the case of unbounded (infinite volume) spaces. Well, in that case, my prima facie suggestion is obvious: allow that one's epistemic state corresponds to an unbounded measure. There is another way of doing this: one can consider sequences of measures L_n that converge to Lebesque measure. E.g. on the real line confine the lebesque measure to larger and larger segments [-n,n]. Then one can define a sequence of probability measures on the real line: $P_n=L_n/n$. This will converge to a finitely additive probability measure that is uniform on the real line in the sense that each finite size region will get probability 0. Now finite additivity leads to non-conglomerability, and to countable Dutch books, but as I have said, I am not worried about that. However, there is another problem: what probability distribution one ends up with depends on what sequence of measures converging to the Lebesque measure that one takes. For instance suppose one took a sequence of Lebesque measures confined to segments of form [-n, 2n]. Then for each P_n one would have $P_n(0, \infty)=2P_n(-\infty, 0)$, so that this would also be true of the P that one converges to. This is odd, e.g. not reflection invariant even though the Lebesque measure is reflection invariant.

So one might suggest: just stick with Lebesque measure. But, one might even want one's epistemic attitudes to have more structure than that. For instance, one might want one's epistemic attitude to encode differences between certain infinite sets: surely the brightest star in the universe is less likely to be in an infinite quadrant Q than in an infinite hemisphere H, when Q is a subset of H. But this is not represented in the Lebesque measure over all of space. Moreover, certain infinite volumes can be rigidly transformed into subsets of other infinite volumes, while the inverse is not possible. Any infinite quadrant can be rigidly transformed into a subset of any infinite hemisphere, while the inverse is not possible. (In some cases both is possible: take two infinite hemispheres A and B: one can always rigidly transform A into a subset of B, and one can always rigidly transform B into a subset of A). In general, surely one should be allowed to attach a probability to infinite cones spreading out from points which is proportional to the solid angle of the cone.

Perhaps therefore the thing to do is as before: take one of the finitely additive probability distributions, and generate an equivalence class by demanding that all rigid transformations of P also be in the class. Interestingly this generates more possibilities of epistemic attitudes than just the lebesque measure. For instance one could have an epistemic attitude that corresponds just to a symmetrically expanding probability distribution, or one could have one that consists of two asymmetrically expanding probability measures, one that expands twice as fast to right as to left, and the reflection of that. This also shows that demanding symmetry doesn’t fix one’s epistemic attitude. So one can still be a subjectivist in the face of symmetry.

Finally let me turn to the puzzles that I started with, Cian Dorr’s puzzle, and my variations on Cian Dorr’s puzzle. I suggest that rational people should be allowed to hang on to Symmetry, Reflection and Solidarity, so that their epistemic states can not be standard probability distributions. But what should they be instead? Well, that depends on what one takes to be the symmetries under which they should be invariant. The problems arose in the following way: one is initially inclined to align one’s credences with limiting relative frequencies. Then one discovers, to one’s chagrin, that limiting relative frequencies depend on the way in which one orders the data. If one were to think that any ordering is as good as any other, then the only relevant structure that one is left with is set-theoretic structure (cardinality structure and subset structure). So, a natural suggestion is that in such cases one's epistemic attitudes towards propositions should have the following structure. One should form equivalence classes of propositions, where the equivalence class that a proposition is in is just determined by the cardinality of the set of possibilities that makes the proposition true, and one should epistemically prefer proposition A to proposition B if the cardinality of cases in which A is true is higher than that of B. Within each equivalence class one could add a bit more structure in the following manner. One should epistemically prefer proposition A to proposition B if the set of possible cases in which B is true is a strict subset of the set of cases in which B is true. If it is true in exactly the same cases, one should be epistemically indifferent. Otherwise the propositions are epistemically incomparable. This is a very weak epistemic structure. But, what is one to do, when one's space of possibilities is so symmetric?

Let me end by summarizing what I have suggested. I have suggested that in certain cases the demand that one's epistemic attitudes satisfy certain symmetries can rule out that one's epistemic attitudes have a standard probabilistic representation. And I have suggested that the epistemic attitudes of rational people should be allowed to satisfy these symmetries. Exactly what the possible structures of epistemic states of rational people in all generality should be taken to be (in addition to being a partial ordering), I don't know. I just made a few suggestions for a few types of cases, but obviously, more thought is required.