The Distribution Lottery Paradox

How are counterfactuals and probability related? One answer to that question is the Lockean thesis. It says that high condition probability is both necessary and sufficient for the truth of a counterfactual. In other words:

Locke: $A \mathbin{\Box\!\!\rightarrow} B$ iff $C(B|A) > t$.

for some suitably high $t \geq 1/2$. Both directions of the thesis are plausible and we seem to rely on them in everyday reasoning.

For example, suppose the lotto has recently hit a billion dollars and the drawing is happening—right now!—on the TV. Ole sees the winning ticket drawn and feels a pang of disappointment. Had he put his life savings into tickets, he would have won, he tells Lena, who is watching the lotto with him. She tells him this is absurd. It is *false* that had he put his life savings into lotto tickets, he would have won. Had he done that, he would have *lost*. Now Lena would seem to be right on both points. Even if Ole had put his life savings into tickets, his odds of winning would have been just a smidge over zero. Lena's denial that Ole would have won thus follows by Locke ($\rightarrow$). Her assertion that he would have lost follows by Locke ($\leftarrow$).

Besides relying on principles connecting counterfactuals and chance, we also rely on various counterfactual inference rules. One of those rules is distribution.

Distribution: $A \vee B \mathbin{\Box\!\!\rightarrow} C \vdash (A \mathbin{\Box\!\!\rightarrow} C) \vee (B \mathbin{\Box\!\!\rightarrow} C)$.

Had either Ole or Lena won the lotto, they would have bought a fishing boat. It then follows that either (a) had Ole won the lotto, they would have bought a fishing boat or (b) had Lena won the lotto, they would have bought a fishing boat. Those claims cannot *both* be false. One of them has to be true. That such reasoning preserves truth would seem obvious.

This paper gives a new lottery paradox for counterfactuals. It turns out that given distribution and *either* direction of the Lockean thesis, we can prove a flat contradiction. Something has to give. After sketching the paradox in section one, we consider similar lottery paradoxes from Hawthorne (2005) and Leitgeb (2014). We show that distribution is the strongest of three paradoxes. Besides having having weaker requirements and stronger conclusions, it most clearly demonstrates the basic conflict between the Lockean thesis and the possible worlds approach to counterfactuals.

1. Distribution

The lotto happens once a week and is hosted by Sven, a famous celebrity.[1] Unfortunately, while walking to work one morning, Sven is flattened by a pickled herring truck. He just never saw it coming. The entire country is shocked by the tragedy and, so, the lotto for that week is never held, out of respect for Sven and the perfectly coiffed hair he stood for. We might still wonder, though, what would have happen if Sven had made it to the studio? What would have happened if he had drawn a ticket?

While investigating his tragic death, the local police discover that the Swedish mafia had some interest in Sven. There was a very small chance that had he drawn a ticket, they would have kidnapped him on the way home. But every week for the last ten years, the mafia has decided to put off the kidnapping and, so, the police conclude that had Sven made it to the studio, he would not have been kidnapped. That said, the lotto also has an *enormous* number of tickets. So many tickets, in fact, that the odds of Sven drawing any particular ticket would have been lower than the odds of him being kidnapped afterwards.

Now we can prove a contradiction. Below, each proposition T_i is the proposition that ticket i wins. K is the proposition that Sven is kidnapped. Premise (1) then says that had Sven either drawn a ticket or been kidnapped, he would not have been kidnapped. Premise (2) says that the odds of kidnapping would have been greater than the odds of any particular ticket winning.

[1]This basic setup is borrowed from Leitgeb (2014), but the reasoning we use to get to paradox is completely different. We will say more about this in section four.

(1)	$(\bigvee_i T_i) \vee K \mathrel{\Box\!\!\rightarrow} \neg K$	premise
(2)	$\bigwedge_i Ch(K \mid T_i \vee K) > t$	premise
(3)	$\bigvee_i (T_i \vee K) \mathrel{\Box\!\!\rightarrow} \neg K$	1, substitution
(4)	$\bigvee_i (T_i \vee K \mathrel{\Box\!\!\rightarrow} \neg K)$	3, distribution
(5)	$\bigvee_i Ch(\neg K \mid T_i \vee K) > t$	4, Locke ($\rightarrow$)
(6)	$\bot$	2, 5, probability

Alternatively, we can prove a contradiction using the other direction of the Lockean thesis. Just replace the last two lines with the following:

(7)	$\bigwedge_i (T_i \vee K \mathrel{\Box\!\!\rightarrow} K)$	2, Locke ($\leftarrow$)
(8)	$\bot$	4, 7, CLNC

The distribution paradox uses three auxiliary inference rules. The first is **substitution**, which tells us that we can replace the antecedent of a counterfactual with a logical equivalent.[2] The second is **probability**. Here, all we are relying on is the claim that because $t \geq 1/2$, $Ch(B|A) > t$ and $Ch(\neg B|A) > t$ are jointly inconsistent when A is consistent. This follows given any reasonable theory of conditional probability. The third rule is the counterfactual law of non-contradiction or **CLNC**. It says that $A \mathrel{\Box\!\!\rightarrow} B$ and $A \mathrel{\Box\!\!\rightarrow} \neg B$ are jointly inconsistent when A is consistent.[3] We also used classical propositional logic throughout.

2. Agglomeration

Our distribution paradox may remind you of another lottery paradox for counterfactuals from Hawthorne (2005). He points out that Locke ($\leftarrow$) is inconsistent with a natural rule of inference called agglomeration.

Agglomeration: $A \mathrel{\Box\!\!\rightarrow} B, A \mathrel{\Box\!\!\rightarrow} C \vdash A \mathrel{\Box\!\!\rightarrow} B \wedge C$

Had Ole won the lottery, we would have bought a boat. Had Ole won the lottery, he would have bought a cabin on the lake. So had Ole won the lottery, he would have both bought a boat and a cabin on the lake.

There are different ways of running the agglomeration argument. The simplest argument is to observe that in a fair lottery, the probability

[2] $A \mathrel{\Box\!\!\rightarrow} C \vdash B \mathrel{\Box\!\!\rightarrow} C$ when $A \vdash B$ and $B \vdash A$.

[3] The CLNC says that $A \mathrel{\Box\!\!\rightarrow} B, A \mathrel{\Box\!\!\rightarrow} \neg B \vdash \bot$ when A is consistent. But consistent in what sense? One option is to use counterfactual consistency. That is, $A \mathrel{\Box\!\!\rightarrow} B, A \mathrel{\Box\!\!\rightarrow} \neg B \vdash \bot$ when $\neg(A \mathrel{\Box\!\!\rightarrow} \bot)$. This requires an additional premise saying that for any ticket i, there would have been no logical contradiction, had Sven either drawn ticket i or been kidnapped. For my own part, I also accept a stronger version of the CLNC, one that says that $A \mathrel{\Box\!\!\rightarrow} B, A \mathrel{\Box\!\!\rightarrow} \neg B \vdash \bot$ whenever A is logically consistent. In that case, no additional premise is required.

of any ticket losing is high, conditional on the drawing being held. Locke (→) and agglomeration then tell us that had the lottery been held, all of the tickets would have lost. But by stipulation, a fair lottery is one in which some ticket would have been drawn, had the lottery been held. So by the CLNC, we have a contradiction.

The disadvantage of the simple argument is that it targets only one direction of the Lockean thesis. With the right setup, though, we can target both directions at once. Take the original case of Sven and suppose there are n parallel universes, all with an identical copy of Sven getting flattened by an identical herring truck. When there was only one universe, we said that had Sven_1 made it to the studio, he would not have been kidnapped. But now since the universes are causally isolated, it follows that had *all* the Svens made it to the all the studios, Sven_1 would not have been kidnapped. Why would adding causally isolated parallel universes change anything? But then by parallel reasoning, for all i, had all the Svens made it to the studio, Sven_i would not have been kidnapped. We then observe that every Sven_i would have had some chance of being kidnapped and that the those kidnappings are probabilistically independent. As long as we choose a large enough n, it follows that the probability of at least one Sven being kidnapped, given that they all survive the herring trucks, is high.

Now we prove absurdity. Let D be the proposition that all the Svens make it to the studio and draw a ticket. Let each proposition K_i be the proposition that Sven_i is kidnapped. Premise (1) then says that for every Sven_i, had all the Svens made it to the studio, that Sven_i would not have been kidnapped. Premise (2) says that the probability of at least one Sven being kidnapped, had they all made it the studio, is high.

(1)	$\bigwedge_i (D \mathrel{\Box\!\!\rightarrow} \neg K_i)$	premise
(2)	$Ch(\neg \bigwedge_i K_i \vert D) > t$	premise
(3)	$D \mathrel{\Box\!\!\rightarrow} (\bigwedge_i \neg K_i)$	1, agglomeration
(4)	$Ch(\bigwedge_i K_i \vert D) > t$	3, Locke (→)
(5)	$\bot$	2, 4, probability

Replacing the last two lines lets us target the other direction of the Lockean thesis.

(6)	$D \mathrel{\Box\!\!\rightarrow} \neg \bigwedge_i \neg K_i$	1, Locke (←)
(7)	$\bot$	3, 6, CLNC

We thus have an agglomeration paradox that targets both direction of the Lockean thesis.

3. Rational Monotonicity

The basic setup for the distribution paradox is borrowed from Leitgeb (2014). Both his paradox and ours use the sad story of Sven and the herring to motivate premises (1) and (2).[4] Besides those, Leitgeb has a premise (3) not used in the distribution paradox.

The other major difference is that Leitgeb relies on an inference called rational monotonicity.

$$\textbf{RM: } A \mathrel{\Box\!\!\rightarrow} C, \neg(A \mathrel{\Box\!\!\rightarrow} \neg B) \vdash A \wedge B \mathrel{\Box\!\!\rightarrow} C$$

RM is completely natural, especially when put in its disjunctive form. Had Ole and Lena gone to the cabin, they would have gone fishing. It then follows that either (a) had they gone to the cabin, it would not have rained or (b) had they gone to the cabin and it *had* rained, they would have gone fishing anyway. RM tells us that reasoning of this form preserves truth.

When proving Leitgeb's paradox, it is helpful to have some abbreviations. Think of S as the proposition that Sven survives the herring truck and L_i as the proposition that every ticket *besides* ticket i loses. These propositions are then defined so that $S \wedge L_i$ is logically equivalent with $T_i \vee K$.[5] Premise (1) then says that had Sven survived the herring truck, he would not have been kidnapped. Premise (2) says that the chance of him being kidnapped would have been greater than the chance of any particular ticket winning. Premise (3) is the new premise. It denies that for every ticket, that ticket would have lost.

(1)	$S \mathrel{\Box\!\!\rightarrow} \neg K$	premise
(2)	$\bigwedge_i Ch(K \mid T_i \vee K) > t$	premise
(3)	$\bigvee_i \neg(S \mathrel{\Box\!\!\rightarrow} \neg L_i)$	premise
(4)	$\bigvee_i (S \wedge L_i \mathrel{\Box\!\!\rightarrow} \neg K)$	1, 3, RM
(5)	$\bigvee_i (T_i \vee K \mathrel{\Box\!\!\rightarrow} \neg K)$	4, substitution
(6)	$\bigvee_i Ch(\neg K \mid T_i \vee K) > t$	5, Locke ($\rightarrow$)
(7)	$\bot$	2, 6, probability

We thus have a conflict between one direction of the Lockean thesis and rational monotonicity.[6]

[4] To be fair to Leitgeb, his host is not name Sven and there is no herring truck, but the cases are structurally similar.

[5] $S = (\bigvee_i T_i) \vee K$ and $L_i = (\neg T_1 \wedge \ldots \wedge \neg T_{i-1} \wedge \neg T_{i+1} \wedge \ldots \neg T_n)$ for all i.

[6] This formulation of Leitgeb's paradox improves on his original in a couple ways. First, we get by without the rule of right weakening ($A \mathrel{\Box\!\!\rightarrow} B \vdash A \mathrel{\Box\!\!\rightarrow} C$ when $B \vdash C$). Second, we avoid the use of double negation elimination ($\neg\neg A \vdash A$). This eliminates the possibility of escaping the paradox by accepting intuitionism.

4. Two Problems

The agglomeration paradox is essentially the observation that the Lockean thesis is incompatible with multipremise closure. This point is familiar from the discussion of lottery paradoxes in the case of belief. There, we learn that the epistemic version of the Lockean thesis entails that you can know that A and know that B without being in a position to know that $A \wedge B$. The agglomeration paradox shows that the same goes for counterfactuals. One of the reasons that the distribution paradox and the rational monotonicity paradox are important is that they shows that multipremise closure is not the only route to contradiction. You can also get there by other means.

As compared to the RM paradox, the distribution paradox fixes what I see as two decisive problems. The first is that the RM paradox relies on a problematic additional premise. The second is that rational monotonicity is simply invalid. We will consider each of these points in turn.

Think of Leitgeb's additional premise three as the claim that for some ticket, had the lotto been played, that ticket might have won. What could possibly be wrong with assuming that? Here is the problem: By the laws of probability, premise two implies that the odds of any individual ticket losing are high. Given Locke (←), it follows that for any ticket, had the lotto been played, that ticket would have lost. But this is just the negation of premise three! Premise two and premise three are inconsistent given the full Lockean thesis. Leitgeb has thus given us a paradox for half the Lockean thesis only assuming that we have already rejected the other half. This is less than ideal.

This curious feature of Leitgeb's paradox poses a further problem. Premise one says that had Sven survived, he would not have been kidnapped. Now, some may claim that we only accept this premise because we think that the conditional probability of kidnapping is low. We are implicitly relying on Locke (←) when we accept premise one, in other words. I don't agree with this view myself, but the point is that for Leitgeb's paradox, the objection would be fatal if true. He would be relying on the *acceptance* of Locke (←) to get premise one, then relying on the *denial* of Locke (←) to get premise three.[7]

Even if we reject Locke (←), and so have no reason to deny premise three, there is still the question of what reason we have to accept it. Why

[7]The distribution paradox also has premise one and so the same complaint could be raised. In that case, though, the results are not catastrophic, since there is no premise three and so no need to deny Locke (←).

not think that for every ticket, had the lotto been played, that ticket would have lost? Here is the best reason I can think of: Suppose that every ticket is such that it would have lost. It then follow that every ticket would have lost. But we know that some ticket would have won, so it must not be that for every ticket, that ticket would have lost. But this reasoning relies on agglomeration. Meaning that either the RM paradox relies on agglomeration, and so is not a distinct paradox, or it is unclear what reason there could be to accept the third premise.

Now Leitgeb might respond that we should accept premise three because it is part of common sense. Fair enough. The concern, though, is that *agglomeration* is also part of common sense. If common sense is only committed to premise three because it is already committed to agglomeration, then we do not, after all, have a distinct paradox.

Fortunately, we can build an RM paradox that simply eliminates premise three. Here is how it goes: Consider the population of Minnesota. As it stands, there are less than ten million people. But suppose there had been at least ten million people. How many people would there have been? Minnesota is a big state, but there would have been fewer than a hundred million—there are limits to how big the population would have been. This is premise one. Furthermore, had there been at least ten million people, the probability of there being *exactly* ten million would have been low. Odds are there would have been a few more.[8] But notice that this is not just true for ten million—it also goes for all the other populations between ten million and a hundred million too. This is premise two. The proof then proceeds by induction.[9]

[8] If you have any doubts, think about it in terms of bets. Suppose you learn that a certain state has at least ten million people. Would you accept a bet at even odds that it had less than ten million and one?

[9] In order to use substitution on line (7), we need to make $n \wedge (n+1)$ logically equivalent to $n+1$. This can be done by building those propositions out of enormous, but finite, conjunctions and disjunctions. Let each n be a conjunction $A_n \wedge B_n$. A_n is a giant conjunction saying of $m < n$ that Minnesota does not have m people. B_n is giant disjunction saying of all $m \leq 100m$ and less than a hundred million that Minnesota has one of those populations. That $n \wedge (n+1)$ logically equivalent to $n+1$ is then easily verified.

(1)	$10^7 \mathrel{\Box\!\!\rightarrow} \neg 10^8$	premise
(2)	$\forall n\ Ch(n+1\vert n) > t$	premise
(3)	$\forall n\ \neg Ch(\neg(n+1)\vert n) > t$	2, probability
(4)	$\forall n\ \neg(n \mathrel{\Box\!\!\rightarrow} \neg(n+1))$	3, Locke ($\rightarrow$)
(5)	$n \mathrel{\Box\!\!\rightarrow} \neg 10^8$	hypothesis
(6)	$n \wedge (n+1) \mathrel{\Box\!\!\rightarrow} \neg 10^8$	4, 5, RM
(7)	$n+1 \mathrel{\Box\!\!\rightarrow} \neg 10^8$	5, substitution
(8)	$10^8 - 1 \mathrel{\Box\!\!\rightarrow} \neg 10^8$	1, 5, 7, induction
(9)	$\bot$	4, 8, QPL

This gives us we wanted. We have a version of the RM paradox that avoids the problematic extra premise. Unlike Leitgeb's paradox, it is compatible with Locke (←) and there is no reason to think that our acceptance of the premises depends on agglomeration.

The problem now is that we face a challenge to RM itself. Consider the claim that had the population of Minnesota been at least ten million, it would have been exactly ten million. This is something we deny by ordinary standards—we generally think that counterfactuals about things like population have some give. You might think that had the population been at least ten million, it would have been less than eleven million, but no one seriously thinks it would have been less than ten million and one. Now, maybe one reason that we think counterfactuals have some give is that we accept Locke and derive them from the probabilities. But this is not essential. For example, we accept that had the population of Minnesota been at least ten million, it *might* have been at least ten million and one. Line four then follows by duality and parity of reasoning.[10] Even without duality, our denial of the would counterfactual may simply be a basic expression of our conviction that the counterfactuals have some give. For my own part, I am far more willing to deny the Lockean thesis than I am to accept that had the population of Minnesota been at least ten million, it would have been less than ten million and one. My denial of that claim simply does not depend on the probabilities.

This means that lines two and three of our revised RM paradox are idle wheels. We can eliminate them and take line four as a premise. But in that case, we have an argument that gives us independent grounds for rejecting RM. Even *ignoring* facts about probability, rational monotonicity still ensnares us in paradox. Myself, I think this just shows that RM is invalid. You might not agree but, even if not, what we learn is that RM is controversial, making it less than ideal for building lottery paradoxes.

[10]Duality says that $\vdash A \mathrel{\Diamond\!\!\rightarrow} B \equiv \neg(A \mathrel{\Box\!\!\rightarrow} \neg B)$.

Distribution, in contrast, is completely uncontroversial. The major debate in the literature is whether $A \vee B \mathrel{\Box\!\!\rightarrow} C$ entails that *both* $A \mathrel{\Box\!\!\rightarrow} C$ and $B \mathrel{\Box\!\!\rightarrow} C$ are true, not whether it entails that *at least one of them* is true. This just strikes everyone as obvious.

5. Possible Worlds

There is a clear sense in which distribution is embedded in the very idea of giving a possible world semantics for counterfactuals. The same is not true for either agglomeration or rational monotonicity. Staging the lottery paradox for counterfactuals using distribution, then, most clearly demonstrates the basic incompatibility between the Lockean thesis and the possible worlds approach to counterfactuals. It also most clearly demonstrates what is unique about the lottery paradox for counterfactuals.

We can make these observations precise. Giving a possible worlds semantics for counterfactuals means using what we will call accessibility models. These are triples $\langle W, R, I\rangle$ consisting of a non-empty set of worlds, a three-place accessibility relation on those worlds, and an interpretation. When $R(wvu)$, we say that v is accessible from w relative to u. We then define the counterfactual operators:

> $A \mathrel{\Box\!\!\rightarrow} B$ is true at u iff there is a world w at which A is true and $A \supset B$ is true at every v such that $Rwvu$.
>
> $A \mathrel{\Diamond\!\!\rightarrow} B$ is true at u iff for every world w at which A is true, there is a world v at which $A \wedge B$ is true such that $Rwvu$.

Lewisian semantics is then just a special case of accessibility semantics. It's the special case in which we require R to be a similarity ordering on worlds.

Now maybe you fully embrace Lewisian semantics. That's fine. What makes the more general accessibility models a useful tool is that we can use them to compare the strength of various rules of inference and determine *which* features of the similarity relation make which rules of inference valid.

Here is what we learn from accessibility models: Lewisian similarity semantics validates agglomeration because similarity is total and acyclic. It validates the CLNC because it is total and validates rational monotonicity because it is transitive. Why does it validate distribution? *Because the semantics is given in terms of an accessibility relation on possible worlds.* Any accessibility relation whatsoever would validate distribution.[11] Staging the counterfactual lottery paradox using distribution, then, most clearly demonstrates the essential incompatibility between the Lockean thesis and the possible worlds approach to counterfactuals.

This is also one of the ways in which the lottery paradox for counterfactuals is strikingly different than epistemic lottery paradoxes. Suppose we think about knowledge as a kind of epistemic necessity operator. The semantics is then given in terms of epistemic models $\langle W, R, I\rangle$ with a two-place accessibility relation. $K(A)$ is then true at w just in case A is true at all the worlds accessible from w. From this perspective, agglomeration is a deep feature of the possible worlds approach to knowledge.[12] So is the epistemic analogue of rational monotonicity.[13] On the other hand, distribution is not even a *feature* of knowledge, let alone a deep feature of the possible worlds approach. You can know that a coin will either land heads or tails without being in position to either (a) know that the coin will land heads or (b) know that the coin will land tails.[14] The distribution lottery paradox demonstrates the *unique* incompatibility between the Lockean thesis and the possible worlds approach to counterfactuals.

[11]These results are proven in appendix [ref].

[12]Since $K(A), K(B) \vdash K(A \wedge B)$ is valid on the class of all epistemic models.

[13]Since $K(A \supset C), \neg K(A \supset \neg B) \vdash K(A \wedge B \supset B)$ is valid on the class of all epistemic models.

[14]It is not completely clear what the analogue of distribution is in the case of knowledge. One option is $K(A \vee B) \vdash K(A) \vee K(B)$. The other is $K(A \vee B \supset C) \vdash K(A \supset C) \vee K(B \supset C)$. Letting A be the proposition that a coin lands heads, B the proposition that it lands tails, and C an arbitrary known falsehood, we get a counterexample either way.

References

Hawthorne, J. (2005). Chance and counterfactuals. *Philosophy and Phenomenological Research*, *70*(2), 396–405.

Leitgeb, H. (2014). A lottery paradox for counterfactuals without agglomeration. *Philosophy and Phenomenological Research*, *89*(3), 605–636.