Deference and Infinite Frames

Three Worked Examples

Brian Weatherson

2024-05-04

Abstract

Dmitri Gallow showed that it is impossible to defer to two experts, and plan to linearly mix their views if they disagree. Snow Zhang showed that if there are only finitely many possibilities for the experts’ credences, it is impossible to defer to both and plan to have any strict mixture of the two views if they disagree. This note shows that if we drop the assumption that only finitely many values are possible for the expert credences, this kind of mixing is possible.

Dmitri Gallow (2018) proved that there is no triple of probability functions C, A, B satisfying the following constraints.

1. ∀*a*: C(p | A(p) = *a*) = *a*;
2. ∀*b*: C(p | B(p) = *b*) = *b*;
3. C(A = B) < 1;
4. For some λ ∈ (0,1), ∀*a*,*b*: C(p | A(p) = *a* ∧ B(p) = *b*) = λ*a* + (1-λ)*b*.

That is, C can’t defer to both A and B individually (with respect to p’s probability), think that A and B might disagree, and in the event they do disagree, plan to take a fixed linear mixture of A’s probability and B’s probability.

Snow Zhang recently proved a result that mostly generalises Gallow’s result, though it does weaken it in one crucial respect. (I’m describing here a simplification of Zhang’s result, which also generalises the number of possible experts.) She shows that it is impossible for A, B and C to satisfy the following five constraints.

1. ∀*a*: C(p | A(p) = *a*) = *a*;
2. ∀*b*: C(p | B(p) = *b*) = *b*;
3. C(A = B) < 1;
4. For any *a*,*b*: C(p | A(p) = *a* ∧ B(p) = *b*) is strictly between *a* and *b*.
5. For some finite set of values S, C(A(p) ∈ S ∧ B(p) ∈ S) = 1.

This note shows that the last constraint is essential; it is possible to satisfy the first four constraints without it. I’ll show this by constructing a model where the first four constraints are satisfied. In this model there will uncountably many values that A(p) and B(p) could take. It’s an open question whether Zhang’s result holds if we weaken 5 to say that S is countable.

I’ll describe the model in words first, then describe it algebraically. Let X, Y and Z be normal distributions with mean 0 and variance 1. In symbols, each of them is (0,1). So the sum of any two of them has distribution (0,2), and the sum of all three has distribution (0,3). Let p be the proposition that this sum, X + Y + Z, is positive. Let C be a credence function that knows[[1]](#footnote-20) nothing about X, Y, Z except what’s stated in this paragraph, so C(p) = ½.

C knows some things about A and B. Both of them know everything C knows about X, Y, Z, and each are logically and mathematically omniscience. One of them knows the value of X, and one of them knows the value of X + Y. A fair coin was flipped. If it landed heads, then A knows X and B knows X + Y; if it landed tails, it was the other way around. C knows about this arrangement, but doesn’t know how the coin landed. Let H be the proposition that it landed heads.

Since both A and B know everything C knows plus something more, C should defer to them. If C knew which knew X + Y and which only knew X, they would defer to the one who knew X + Y. They don’t know this, but conditional on knowing the values of A(p) and B(p), they can go close to figuring it out.

Assume for now that the coin landed heads, so H is true. We’ll work out the joint density function for A and B. Then we can work out the same density function conditional on , and from those two facts work out the posterior probability of H. Call this value *h*. Conditional on A(p) = *a*, and B(p) = *b*, C’s probability for p should be (1-h)*a* + *hb*. That’s because conditional on A(p) = *a*, B(p) = *b* and H, C’s probability for p should be *b*, while conditional on A(p) = *a*, B(p) = *b* and , C’s probability for p should be *a*. The short version of what follows is that since *h* is a function of *a* and *b* and is always in (0,1), it follows that C obeys constraint 4.

Given H, we can work out the value of X from A(p) = *a*. In what follows, φ(*x*) is the cumulative distribution for the standard normal distribution, i.e., for (0,1), and φ-1 is its inverse. If X = *x*, then p is true iff Y + Z > -*x*. Since Y + Z is a normal distribution with mean 0 and variance 2, i.e., standard deviation , the probability of this is φ(). So *x* = φ-1(*a*).

Given H, that X = φ-1(*a*), and B(p), we can work out what Y must be as well. If B(p) = *b*, that means that the probability that Z > -(X + Y) is *b*. Since Z just is a standard normal distribution, that means that X + Y is φ-1(*b*), and hence Y is φ-1(*b*) - φ-1(*a*).

Now we can work out the joint density function for *a* and *b* conditional on H. Given H, A(p) = *a* and B(p) = *b* just when X = φ-1(*a*) and Y = φ-1(*b*) - φ-1(*a*). And if we write d(*x*) for the density function for the standard normal distribution[[2]](#footnote-21), the joint distribution has density d(

1. I’m speaking metaphorically here; really the person whose credence function is C knows these things. But the point of this setup is to picture where the model below is coming from, so this way of speaking is harmless enough. [↑](#footnote-ref-20)
2. i.e., d(*x*) = . [↑](#footnote-ref-21)