Schoenfield (2017) Conditionalization Does Not (In General) Maximize Expected Accuracy
Handout for Epistemology Reading Group, Adrian Liu, Feb 11, 2025

1 | Introduction

Implicit in the literature on Bayesian updating is the idea that information processing happens in two stages. **First**: *exogenous*, non-inferential information possession. Metaphorically, we can think of this stage as involving the world 'flinging' some information at the agent. Call the flung stuff *exogenous evidence.* **Second**: *inferential* belief revision, which may lead the agent to possess further information *endogenously*.

Question: How do you rationally revise beliefs in response to exogenous evidence? Bayesians say it is to conditionalize on the evidence:

$$p_{new}(\,\cdot\,) = p_{old}(\,\cdot\mid E) := \frac{p_{old}(\,\cdot\wedge E)}{p_{old}(E)}. \tag{Conditionalization}$$

Greaves and Wallace (G&W) defend conditionalization via accuracy-maximization:

1. RatAcc: The rational update procedure is the one that maximizes expected accuracy according to any strictly proper scoring rule.
2. Conditionalizing on one's evidence is the update procedure that maximizes expected accuracy according to any strictly proper scoring rule.
3. Therefore, conditionalizing on one's evidence is the rational update procedure.

Schoenfield argues that, *in general*, the second premise is not true:

1. In general, the update procedure that maximizes expected accuracy according to any strictly proper scoring rule is conditionalization*:

$$p_{new}(\,\cdot\,) = p_{old}(\,\cdot\mid\langle\text{I learned that } E\rangle) \tag{Conditionalization*}$$

2. Central Thesis: RatAcc implies that in general the rational update procedure is conditionalization*, and not conditionalization.

She further argues that epistemologists who endorse RatAcc are committed to the existence of a class of propositions concerning an agent's situation, such that, for any rational agent S, these propositions will be true of S iff she is certain of them.

2 | Setup

Scoring Rules: A scoring rule is a function $\mathbf{A} : C_S \times S \to [0, 1]$ that takes a credence function $c \in C_S$ and a state $s \in S$, where $A(c, s) = t$ means that A thinks that credence function c has accuracy t at state s. Relative to a probability function p over S, we can take the expected accuracy of a credence function c across any subset of states $P_i \subseteq S$:

$$\mathbb{E}\mathbf{A}^{p}_{P_i}(c) =_{\text{def}} \sum_{s\in P_i} p(s)\mathbf{A}(c, s). \tag{$\mathbb{E}\mathbf{A}$}$$

A scoring rule is *strictly proper* if the value above is maximized at $c = p$. This captures the idea of expected accuracy being "according to p's estimate.

Learning Experiences and Update Procedures: Tomorrow you wake up and some proposition P might be flung at you as your evidence or, more boringly, there is some proposition "that you exogenously learn" (we can represent having no evidence flung at you as the same as having the trivial proposition $\top$ flung at you).

Let $L(P)$ be the proposition $\langle$I learned that $E\rangle$. Then if $X_1, X_2, \ldots, \top$ are the propositions you might learn (call the set of them X), then $L(X) = \{L(X_1), L(X_2), \ldots, L(\top)\}$, the set of propositions that describe what you might learn, partition the set of states: in every possibility you learn exactly one thing, so exactly one of the $L(X_i)$ is true.

A [learning] *update procedure* is a function U_L which maps each proposition X_i an agent might learn to a credence function $U_L(X_i)$, "with the intended interpretation that an agent conforming to U_L adopts $U_L(X_i)$ as her credence function if and only if the proposition she learns upon undergoing the learning experience is X_i."

The *expected accuracy* of an update procedure U_L relative to a credence function p can be calculated as the expected accuracy of $U_L(X_i)$ for the states in each X_i (according to p), weighted by the likelihood (again according to p) that a subject learns X_i:

$$\mathbb{E}\mathbf{A}^p_{\text{update}}(U_L) =_{\text{def}} \sum_{L(X_i) \in L(X)} \mathbb{E}\mathbf{A}^p_{L(X_i)}(U_L(X_i)) \qquad \text{(Update } \mathbb{E}\mathbf{A}\text{)}$$

3 | Results

1. G&W (Greaves and Wallace, 2006): Let S be a set of states and let $P = \{P_1, \ldots, P_n\}$ partition S. Let $\mathcal{F}$ be the set of functions that assign a credence function over S to each P_i. Then if $\mathbf{A}$ is a strictly-proper scoring rule, the function F that maximizes

$$\mathbb{E}\mathbf{A}^p_{\text{function}}(F) =_{\text{def}} \sum_{P_i \in P} \mathbb{E}\mathbf{A}^p_{P_i}(F(P_i)) \qquad \text{(Function } \mathbb{E}\mathbf{A}\text{)}$$

is the function $F(P_i) = p(\cdot \mid P_i)$.

Now let an *experiment* be a learning experience where you will exogenously learn proposition X_i if and only if X_i is true, i.e. $L(X_i) \leftrightarrow X_i$. Then:

2. CondMax: Suppose you know you will perform an experiment X. Then the update procedure that maximizes Update $\mathbb{E}\mathbf{A}$ is

$$U_L(X_i) = p(\cdot \mid X_i). \qquad \text{(conditionalization)}$$

Schoenfield shows that an agent will be certain that $L(X_i) \leftrightarrow X_i$ if and only if:

> Partitionality: The propositions that the agent assigns non-zero credence to exogenously learning partition the agent's possibility space.
>
> Factivity: The agent is certain that if she learns P, P is true.

So CondMax only implies that conditionalization maximizes expected accuracy in cases where Partitionality and Factivity are true. And both of these are substantive assumptions. If we remove the condition that $L(X_i) \leftrightarrow X_i$, we get:

3. Generalized CondMax: Suppose you know you will undergo a learning experience X. Then the update procedure that maximizes Update $\mathbb{E}\mathbf{A}$ is $U_L(X_i) = p(\cdot \mid X_i)$. Then the update procedure that maximizes Update $\mathbb{E}\mathbf{A}$ is

$$U_L(X_i) = p(\cdot \mid L(X_i)). \qquad \text{(conditionalization}^*\text{)}$$

Generalized CondMax is strictly more general than CondMax. In cases where $L(X_i) \leftrightarrow X_i$, plugging X_i in for $L(X_i)$ in the statement of Generalized CondMax returns CondMax.

Generalized CondMax and RatAcc together imply:

4. Cond*: The rational update procedure is conditionalization*. Upon learning P_i, an ideally rational agent will conditionalize on the proposition that she learned P_i.

Cond* and the definition of conditional probability imply:

4. LL: If one learns P_i, one is rationally required to be certain that one learned P_i.

(If one is rational, then when one learns P_i, $p_{new}(L(P_i)) = p_{old}(L(P_i) \mid L(P_i)) = 1$.)

And externalists are apt to dislike this sort of iteration principle.

4 | Getting Out of the Results

Bronfman (2014) argues: the rational update procedure isn't the accuracy-maximizing procedure *from the pool of possible update procedures*, but *from the pool of update procedures that the agent can competently execute*. Schoenfield has some smaller complaints, but ultimately objects that we are looking for an *idealized* update procedure.

Distinguish between agents who are idealized information *possessors* (they know all and only the truths) and agents who are idealized information *processors* (they may not know all the truths, but what they do know, they are perfect at updating on). If we're interested in idealized information *processing*, Schoenfield argues, then any operation performed on the proposition exogenously learned should be admissible. But then conditionalization* is allowed:

> like conditionalization, conditionalization* is simply an operation performed on the proposition exogenously learned. The operation is the following: if P is the proposition learned, take P, attach an L to it, and conditionalize on the resulting proposition: $L(P)$.

Question: doesn't conditionalizing on $L(P)$ require the agent to possess more information (namely $L(P)$) than just possessing P? An objector should say: just attaching an L to the proposition P doesn't give you the information that you learned that P.

5 | More Results

5. Super Generalized CondMax: Let R be a relation and U_R be a function from a set of propositions X to credence functions such that an agent A conforming to U_R adopts $U_R(X_i)$ whenever $R(A, X_i)$. If $R(A, X_i)$ is always true only for exactly one X_i, then the U_R such that conforming to U_R maximizes expected accuracy[1] is

$$U_R(X_i) = p(\cdot \mid R(A, X_i)) \qquad \text{(relation-conditionalization)}$$

6. Super-Duper Generalized CondMax (SD CondMax): Let $P = \{P_1, \ldots, P_n\}$ partition a set of states Ω. Let U_F be a function from P to credence functions such that an agent adopts $U_F(P_i)$ whenever P_i obtains. The U_F such that conforming to U_F maximizes expected accuracy is

$$U_F(P_i) = p(\cdot \mid P_i).^{2} \qquad \text{(function-conditionalization)}$$

Schoenfield argues that SD CondMax and RatAcc together imply:

7. Luminous Infallibility: There is a class of propositions concerning an agent's situation, such that, for any agent S, if S is rational, these propositions will be true of S if and only if she is certain of them.

Thus: "There is a sense, then, in which a defender of RatAcc can't help but adopt some version of the truth rule. For whatever one's theory of rationality is, one can partition the space of possible situations an agent might find herself in in such a way that the same doxastic state is rational in each cell of the partition. Perhaps, for example, a theorist partitions the space based on what the agent's phenomenology is: {She has phenomenology P1, she has phenomenology P2...} or what she learns: {She learns X_1, she learns $X_2, \ldots$} or what her evidence is: {She possesses E1, She possess E2...}. Call this partition, whatever it is, P."

Then SD CondMax implies that the accuracy-maximizing update procedure is to conditionalize on P_i whenever P_i is true. But if an agent conditionalizes on P_i whenever P_i is true, then they are certain of P_i whenever P_i is true. Thus:

> If RatAcc is true, then the propositions whose truth determines what credence function it is rational for an agent to adopt are propositions that a rational agent is luminously infallible about that is, they are propositions that she will be certain of if and only if they are true.

Schoenfield's final upshot: "the thought that rationality involves maximizing expected accuracy and such claims as Luminous Infallibility are intertwined."

1. The quantity maximized is

$$\mathbb{E}\mathbf{A}^{p}_{\text{relation}}(U_R) =_{\text{def}} \sum_{R(S,X_i) \in R(S,X)} \mathbb{E}\mathbf{A}^{p}_{R(S,X_i)}(U_R(X_i)) \qquad \text{(Relation } \mathbb{E}\mathbf{A})$$

2. This follows directly from G&W: F could be thought of as the function that describes the agent who adopts update procedure U_F. The quantity maximized is Function $\mathbb{E}\mathbf{A}$, with U_F substituted for F.