The Review of Symbolic Logic
Volume 18, Number 1, March 2025

NON-FACTIVE KOLMOGOROV CONDITIONALIZATION

MICHAEL RESCORLA
Department of Philosophy, University of California, Los Angeles

Abstract. *Kolmogorov conditionalization* is a strategy for updating credences based on propositions that have initial probability 0. I explore the connection between Kolmogorov conditionalization and *Dutch books*. Previous discussions of the connection rely crucially upon a *factivity assumption*: they assume that the agent updates credences based on *true* propositions. The factivity assumption discounts cases of *misplaced certainty*, i.e., cases where the agent invests credence 1 in a falsehood. Yet misplaced certainty arises routinely in scientific and philosophical applications of Bayesian decision theory. I prove a non-factive Dutch book theorem and converse Dutch book theorem for Kolmogorov conditionalization. The theorems do not rely upon the factivity assumption, so they establish that Kolmogorov conditionalization has unique pragmatic virtues that persist even in cases of misplaced certainty.

§1. Conditioning on a probability zero proposition. Bayesian decision theory models how the credences of an idealized rational agent evolve over time. Suppose that the agent has initial credences P and subsequently transitions to new credences P_{new}. The agent has *conditionalized on* E when

$$(1)\quad P_{new}(.) = P(.|E),$$

where $P(H|E)$ is the initial conditional probability of H given E. A *conditioning proposition* is a proposition on which the agent has conditionalized.

Philosophers usually confine attention to scenarios where the conditioning proposition E has non-zero initial probability. Conditional probabilities then satisfy the *ratio formula*

$$P(H|E) = \frac{P(H\&E)}{P(E)},$$

which yields the following recipe for conditionalizing on E:

$$(2)\quad P_{new}(H) = \frac{P(H\&E)}{P(E)}.$$

When an agent updates her credences through (2), I will say that she engages in *ratio conditionalization*. Using ratio conditionalization, we can model a wide range of credal reallocation scenarios.

Received: March 29, 2023.
2020 *Mathematics Subject Classification*: 03A05, 03A10.
Key words and phrases: conditionalization, regular conditional distribution, subjective probability, Bayesian decision theory, Dutch books.

© The Author(s), 2023. Published by Cambridge University Press on behalf of The Association for Symbolic Logic.

doi:10.1017/S1755020323000345

https://doi.org/10.1017/S1755020323000345 Published online by Cambridge University Press

The ratio formula is ill-defined when $P(E) = 0$. Yet cases where $P(E) = 0$ arise frequently within scientific applications. To illustrate, let X be a random variable with continuum many values. Orthodox probability theory demands that $P(X = x) = 0$ for all but countably many values x. Conditionalization on propositions of the form $X = x$ occurs routinely in statistics [14], physics [47], robotics [46], economics [23], and virtually every other scientific discipline that uses Bayesian inference. For example, a navigator might estimate her position based on her distance from a familiar landmark; an astronomer might estimate when a comet will reach Earth given its current velocity; and so on.

Say that an agent engages in *null updating* when she updates her credences based on a proposition E such that $P(E) = 0$. Since null updating plays an important role in scientific practice, any complete foundation for Bayesian decision theory should address it. We must look beyond the ratio formula for a more general analysis of conditional probability.

The good news is that Kolmogorov [24] already offered a satisfying analysis in the same treatise where he laid measure-theoretic foundations for probability. Kolmogorov's analysis hinges upon the notion of a *regular conditional distribution* (rcd). In the decades since Kolmogorov's discussion, rcds have come to figure indispensably in probability theory [2, 34] and Bayesian statistics [11, 14]. The bad news is that philosophical inquiry into rcds has lagged behind mathematical and scientific practice. For many decades, rcds received almost no philosophical attention save for the occasional hostile dismissal (e.g., [16, 18, 19, 30, 41, 42]).

Lately, some philosophers have explored rcds in a more sympathetic vein (e.g., [8–10, 15, 20–22, 27, 28, 31, 35, 36, 40]). A running theme is that rcds support a kind of conditionalization, sometimes called *Kolmogorov conditionalization*. Recent work has explored how Kolmogorov conditionalization works and why one might find it an attractive credal update strategy. The present paper contributes to this initiative by probing the connection between Kolmogorov conditionalization and *Dutch books*.

A *Dutch book* is a set of acceptable bets with a guaranteed net loss. You are *Dutch bookable* if it is possible to rig a Dutch book against you. Dutch bookability is a very undesirable property. When you are Dutch bookable, a sufficiently clever bookie can pump you for money by offering you a series of bets that you accept. Thus, it is somehow bad to be Dutch bookable and somehow good to render yourself immune to Dutch books.

Dutch books have long played a foundational role in Bayesian decision theory. Ramsey [33] and de Finetti [6] independently proved that, under ancillary assumptions, agents whose credences violate the probability calculus axioms are Dutch bookable. Lewis proved a Dutch book theorem for conditionalization. He showed that one can rig a diachronic Dutch book (a Dutch book with bets at different times) against an agent who follows any update rule other than conditionalization. Although Lewis did not initially publish his proof, it became well-known due to Teller's [45] exposition, and Lewis [26] eventually published his own treatment. Skyrms [43] proved a converse theorem: conditionalizers who obey the probability calculus axioms are not Dutch bookable.

Lewis and Skyrms confined attention to learning scenarios where the conditioning proposition has non-zero initial probability. For that reason, their theorems have fairly limited scope when compared with scientific applications of the Bayesian framework. [36] extends the Lewis–Skyrms theorems to numerous learning scenarios

https://doi.org/10.1017/S1755020323000345 Published online by Cambridge University Press

that feature null updating. The extended theorems show that, in those learning scenarios, Kolmogorov conditionalization is the unique update strategy that immunizes the agent from Dutch books.

Unfortunately, the aforementioned diachronic Dutch book theorems rely heavily upon a questionable *factivity assumption*: they assume that the agent updates her credences based only on *true* propositions. The factivity assumption discounts cases of *misplaced certainty*, i.e., cases where the agent invests credence 1 in a falsehood. Misplaced certainty arises routinely in scientific and philosophical applications of the Bayesian framework. A navigator who estimates her position based on her distance from a landmark may mismeasure the distance; an astronomer estimating a comet's arrival time based upon its velocity may miscalculate the comet's velocity; and so on. Mistakes may arise through carelessness, deception, faulty measurement, or various other factors. Given the omnipresent potential for misplaced certainty, a complete treatment should not presuppose the factivity assumption [37]. A complete treatment should be *non-factive*.

In this paper, I will generalize the results from [36] to a non-factive setting. I will consider a class of learning scenarios where the conditioning proposition may be false and may have non-zero initial probability. I will show that, in those learning scenarios, Kolmogorov conditionalization is the unique update strategy that guards against Dutch books. Thus, conditionalization's pragmatic virtues persist when we allow non-factive null updating.

Section 2 reviews the basics of rcds. Section 3 discusses how rcds support Kolmogorov conditionalization. Section 4 states and proves the non-factive Dutch book theorem for Kolmogorov conditionalization. Section 5 concludes that Kolmogorov conditionalization has desirable pragmatic features distinguishing it from rival credal reallocation strategies.

§2. Regular conditional distributions. Probability theory as formalized by Kolmogorov studies a probability space $(\Omega, \mathcal{F}, P)$, where Ω is a set, $\mathcal{F}$ is a σ-field over Ω, and P is a probability measure on $\mathcal{F}$. Elements of Ω are *outcomes*. Elements of $\mathcal{F}$ are *events*. For purposes of this paper, we use $(\Omega, \mathcal{F}, P)$ to model an agent's credences. Events serve as the objects of credence. The real number assigned by P to event $A \in \mathcal{F}$ measures the agent's degree of belief in A. Since events serve as the objects of credence, they play one role traditionally assigned by philosophers to *propositions*. However, we need not assume that events have all the features traditionally ascribed to propositions.

In a measure-theoretic setting, the ratio formula mentions intersection of events rather than conjunction of propositions:

$$P(A|B) = \frac{P(A \cap B)}{P(B)},$$

which is not well-defined when $P(B) = 0$. Kolmogorov offers a generalized treatment designed to accommodate cases where $P(B) = 0$. The basic idea is to associate B with a *collection* of suitably related events (some of which may also have probability 0). Rather than conditionalize on B taken by itself, one conditionalizes upon the associated collection.

Formally speaking, Kolmogorov's theory centers on a subset $\mathcal{G} \subseteq \mathcal{F}$, where $\mathcal{G}$ is itself a σ-field. The idea is to define probabilities conditional on information (possibly

https://doi.org/10.1017/S1755020323000345 Published online by Cambridge University Press

including misinformation) regarding whether the true outcome belongs to each $G \in \mathcal{G}$. Kolmogorov's theory can be glossed in different ways depending on how we interpret the rather vague phrase "information." In what follows, I will present my own preferred interpretation and elucidate Kolmogorov's approach accordingly.[1]

2.1. Kolmogorov certainty acquisition scenarios. Suppose the agent gains new certainties regarding sub-σ-field $\mathcal{G}$. She assigns probability 0 or 1 to each event $G \in \mathcal{G}$, and on that basis she reallocates credence over the rest of $\mathcal{F}$. For $\omega \in \Omega$, let $\delta_\omega : \mathcal{G} \to \mathbb{R}$ be the function defined by

$$\delta_\omega(G) = \begin{cases} 1 & \text{if } \omega \in G, \\ 0 & \text{if } \omega \notin G, \end{cases} \qquad \text{for each } G \in \mathcal{G}.$$

Call δ_ω a *certainty profile* over $\mathcal{G}$, and call ω an *index* of δ_ω. Note that δ_ω is a probability measure. δ_ω models a scenario where, for every $G \in \mathcal{G}$, the agent becomes certain that G does or does not obtain. A notable special case arises when the agent's certainty profile δ_ω tracks the truth:

(3) For each $G \in \mathcal{G}$, $\delta_\omega(G) = 1$ iff the true outcome belongs to G.

Since I am studying non-factive conditionalization, I will not assume that (3) prevails. I will consider learning scenarios where the agent acquires a new certainty profile δ_ω that may or may not track the truth. I call these *Kolmogorov certainty acquisition scenarios*. In a Kolmogorov certainty acquisition scenario, the agent becomes certain as to whether the true outcome belongs to each $G \in \mathcal{G}$, but her new certainties may be misplaced.[2]

A single certainty profile can have many different indices. If outcomes ω and ν belong to precisely the same members of $\mathcal{G}$, then they index the same certainty profile:

$$\delta_\omega = \delta_\nu.$$

In practice, it is often more convenient to deal with indices rather than certainty profiles. But our main concern is certainty profiles, not indices.

Many scientifically important situations are naturally modeled as Kolmogorov certainty acquisition scenarios. To illustrate, let $X : \Omega \to \mathbb{R}$ be a random variable. Suppose the agent becomes certain that X has value r. Thus, she assigns credence 1 to the event $\{\omega : X(\omega) = r\}$. If the expression "$r$" is suitably informative, she is positioned to become certain of numerous additional events. Assuming that "r" belongs to any standard notational scheme for the real numbers, she should be willing to affirm or deny that X's value falls between a and b, for any $a, b \in \mathbb{Q}$. So she should be willing to assign probability 1 or 0 to each event

$$X^{-1}(a, b) \qquad \text{for any } a, b \in \mathbb{Q}.$$

Let $\sigma(X)$ be the σ-field generated by these events, i.e., $\sigma(X)$ results from starting with the events $X^{-1}(a, b)$ and closing under complementation and countable union. The agent's new certainties over events $X^{-1}(a, b)$ determine a unique certainty profile over $\sigma(X)$.

[1] Kolmogorov himself favored a frequentist rather than subjectivist interpretation of probability. So he would not have elucidated his mathematical framework in the subjectivist way that I will elucidate it.

[2] Lee [25] also discusses how sub-σ-fields can model non-factive doxastic states.

https://doi.org/10.1017/S1755020323000345 Published online by Cambridge University Press

There are uncountably many events in $\sigma(X)$. Most of those events a typical agent will never consider. Even when an agent explicitly represents an event $G \in \sigma(X)$, she may not compute which probability she should assign to G given her new certainties regarding the events $X^{-1}(a, b)$. So δ_ω does not plausibly model the explicit certainties of any agent who remotely resembles a normal human. We should instead regard δ_ω as modeling an agent's *implicit certainties*. The agent does not extrapolate from her certainty that $X = r$ to new certainties over each event in $\sigma(X)$, but she is *positioned* in principle to do so. δ_ω captures a large collection of certainties that the agent is positioned to acquire, given her newfound certainty that $X = r$.

2.2. Update rules. In a Kolmogorov certainty acquisition scenario, circumstances influence the agent's credences by fixing new certainties over $\mathcal{G} \subseteq \mathcal{F}$. Her new certainties over $\mathcal{G}$ provide her sole basis for reallocating credence over the rest of $\mathcal{F}$. How should we capture these intuitive ideas in mathematical terms? Initially, one might suggest a function that takes certainty profiles as inputs and yields new credal reallocations as outputs. This suggestion is not very mathematically tractable, so Kolmogorov employs a more roundabout procedure. Rather than treat credal reallocation as a function of the certainty profile, he treats it as a function of an *index* for the certainty profile. He considers a function $C : \mathcal{F} \times \Omega \to \mathbb{R}$, where $C(A, \omega)$ is the new credence to be assigned to $A \in \mathcal{F}$ in light of certainty profile δ_ω. I notate $C(A, \omega)$ as $C(A|\ \omega)$. I will sometimes notate $C(.|\ \omega)$ as C_ω.

What constraints should we place upon C? Since we are using probability measures to model credences, we demand that

(4) $C_\omega : \mathcal{F} \to \mathbb{R}$ is a probability measure *for each* $\omega \in \Omega$.

Kolmogorov imposes a crucial additional constraint. He demands that, for each $A \in \mathcal{F}$, the one-place function $C(A|.) : \Omega \to \mathbb{R}$ be $\mathcal{G}$*-measurable*:

(5) $C(A|.)^{-1}(-\infty, a] \in \mathcal{G}$ *for each* $a \in \mathbb{R}$.

Call any function C that satisfies conditions (4) and (5) an *update rule for* $(\Omega, \mathcal{F})$ and $\mathcal{G}$. As I will now explain, condition (5) ensures that the agent's newfound implicit certainties over $\mathcal{G}$ dictate the new credences to be allocated over $\mathcal{F}$.

$C(A|\ \omega)$ is the credence to be assigned to A when the agent newly acquires certainty profile δ_ω. If an index ω for the agent's new certainty profile belongs to the event

$$C(A|.)^{-1}(-\infty, a],$$

then the new credence to be assigned to A is $\leq a$. If an index ω for the agent's new certainty profile belongs to event

$$C(A|.)^{-1}(a, \infty),$$

then the new credence to be assigned to A is $> a$. Assuming C satisfies the $\mathcal{G}$-measurability condition (5), both events belong to $\mathcal{G}$. The agent's certainty profile δ_ω must therefore assign one of the events probability 1. We consider each case:

- If δ_ω assigns probability 1 to $C(A|.)^{-1}(-\infty, a]$, then $\omega \in C(A|.)^{-1}(-\infty, a]$. Thus, the new credence to be assigned to A is $\leq a$.
- If δ_ω assigns probability 1 to $C(A|.)^{-1}(a, \infty)$, then $\omega \in C(A|.)^{-1}(a, \infty)$. Thus, the new credence to be assigned to A is $> a$.

https://doi.org/10.1017/S1755020323000345 Published online by Cambridge University Press

So the agent's new certainty profile dictates whether the new credence to be assigned to A is $\leq$ or $> a$. By contrast, suppose C violates (5). Then the agent's certainty profile does not dictate whether the new credence to be assigned to A is $\leq$ or $> a$. Her implicit certainties do not speak to whether the new credence should be $\leq$ or $> a$. Hence, $\mathcal{G}$-measurability formalizes the thought that the agent's new credal assignment over $\mathcal{F}$ is embedded in her implicit certainties over $\mathcal{G}$.

As an added bonus, $\mathcal{G}$-measurability entails that C induces a well-defined mapping from certainty profiles to credences. When C is $\mathcal{G}$-measurable, its output does not depend on the particular index through which we identify a certainty profile:

$$\delta_\omega = \delta_v \rightarrow C_\omega = C_v.$$

To see this, let us prove the contrapositive. Assume that

$$C_\omega \neq C_v.$$

Pick $A \in \mathcal{F}$ such that $C(A|\ \omega) \neq C(A|\ v)$ and let

$$H =_{df} \{\rho \in \Omega : C(A|\omega) = C(A|\rho)\}.$$

$C(A|.) : \Omega \rightarrow \mathbb{R}$ is $\mathcal{G}$-measurable, and H is the inverse image of $\{C(A|\omega)\}$ under $C(A|.)$, so $H \in \mathcal{G}$. Since $\omega \in H$ and $v \notin H$, it follows that

$$\delta_\omega(H) = 1\ \&\ \delta_v(H) = 0$$

and hence that $\delta_\omega \neq \delta_v$. Thus, C_ω and C_v are identical whenever ω and v index identical certainty profiles. This observation vindicates Kolmogorov's decision to study update rules that take indices rather than certainty profiles as arguments. We want to model credal reallocation in light of new certainty profiles, so the particular index through which we identify a certainty profile should not matter.

2.3. The integral formula. Kolmogorov supplements (4) and (5) with an additional constraint upon C. The constraint is now usually called the *integral formula*:

$$(6)\ \ P(A \cap G) = \int_G C(A|\omega)dP(\omega) \qquad \textit{for each } A \in \mathcal{F} \textit{ and } G \in \mathcal{G}.$$

The integral formula generalizes the law of total probability.

When $C(A|.)$ satisfies (5) and (6), it is a *conditional probability for A given $\mathcal{G}$*. A function $C : \mathcal{F} \times \Omega \rightarrow \mathbb{R}$ is a *regular conditional distribution (rcd) for P given $\mathcal{G}$* iff it satisfies (4)–(6). I will often use the notation $P_\mathcal{G} : \mathcal{F} \times \Omega \rightarrow \mathbb{R}$ to denote an rcd for P given $\mathcal{G}$.

As an important special case, suppose that $\mathcal{G}$ is generated by a countable partition $E_1, \ldots, E_n$ of Ω, where $P(E_i) > 0$ for each i. Then it is not hard to show that there exists a unique rcd for P given $\mathcal{G}$, defined by

$$P_\mathcal{G}(A|\omega) = P(A|E_i) = \frac{P(A \cap E_i)}{P(E_i)} \quad \textit{if } \omega \in E_i.$$

In this way, Kolmogorov's theory subsumes the special case where the ratio formula dictates conditional probabilities.

There always exists a conditional probability for A given $\mathcal{G}$ [2, p. 430]. This conditional probability is unique up to measure 0: any two conditional probabilities for A given $\mathcal{G}$ must agree everywhere except possibly on a set of P-measure 0. Existence of rcds is less straightforward. If the probability space is pathological, then an rcd for

https://doi.org/10.1017/S1755020323000345 Published online by Cambridge University Press

P given $\mathcal{G}$ may not exist [2, p. 443]. When the probability space is nice enough, there exists an rcd for P given $\mathcal{G}$. For any Borel set $A \subseteq [0, 1]$, let $\mathcal{B}(A)$ consist of the Borel subsets of A. A measurable space $(\Omega, \mathcal{F})$ is *Borel* iff there exists a Borel set $A \subseteq [0, 1]$ and a bijection f from $(\Omega, \mathcal{F})$ to $(A, \mathcal{B}(A))$ such that f and f^{-1} are measurable. The following theorem is basic to the discipline [12, p. 418]:

THEOREM: *Let $(\Omega, \mathcal{F})$ be a Borel space, P a probability measure on $(\Omega, \mathcal{F})$, and $\mathcal{G} \subseteq \mathcal{F}$ a sub-σ-field. Then there exists an rcd for P given $\mathcal{G}$.*

Virtually all probability spaces used in Bayesian decision theory are Borel. For example, every Polish space is Borel. For further discussion of rcd existence, see [34].

§3. Kolmogorov conditionalization. Imagine an agent who begins with credences given by $(\Omega, \mathcal{F}, P)$. Suppose there exists $P_{\mathcal{G}}$, an rcd for P given $\mathcal{G}$. Suppose the agent adopts the following credal reallocation policy:

(7) *Respond to new certainty profile δ_ω over $\mathcal{G}$ by adopting new credences $P_{\mathcal{G}}(.|\omega)$ over $\mathcal{F}$.*

Suppose the agent subsequently acquires certainty profile δ_ω and, on that basis, adopts new credences $P_{\mathcal{G}}(.|\omega)$ over $\mathcal{F}$. Then I say that the agent *uses $P_{\mathcal{G}}$ to conditionalize on δ_ω*. When an agent uses an rcd to conditionalize, I say that she engages in *Kolmogorov conditionalization*.

Kolmogorov conditionalization plays a foundational role in Bayesian statistics [11, 14]. It also figures prominently in many applications of the Bayesian framework, such as within cognitive science [1] and economics [23, 29].

My main thesis in the paper is that Kolmogorov conditionalization is an appealing update strategy with distinctive virtues. However, an honest appraisal must grant that it also faces significant challenges.

The first challenge concerns *existence*. If there does not exist an rcd for P given $\mathcal{G}$, then an agent with initial credences given by P cannot Kolmogorov conditionalize. Luckily, rcds exist in a wide range of cases, including virtually all cases likely to arise in scientific practice.

The second challenge concerns *uniqueness*. Setting aside the special case given by the ratio formula, unconditional probabilities do not determine unique conditional probabilities: when an rcd exists, there exist infinitely many rcds. Then there are infinitely many credal reallocation policies (7) to choose from. A Kolmogorov conditionalizer must select one such policy. Whereas ratio conditionalization mandates a unique credal update rule, Kolmogorov conditionalization allows infinitely many distinct credal update rules. This is not a fatal problem, but it suggests that we may want to supplement Kolmogorov conditionalization with additional constraints upon credal updates.

The third challenge, which requires more extended discussion, centers upon a phenomenon called *impropriety*. Section 3.1 explains the phenomenon. Sections 3.2–3.4 explore the challenge that it poses to Kolmogorov conditionalizers.

3.1. Propriety. Intuitively speaking, the probability of an event conditional on itself should be 1. And it is indeed an easy theorem that $P(E|E) = 1$ when $P(E) > 0$. What about cases where $P(E) = 0$? Say that an update rule C for $(\Omega, \mathcal{F})$ and $\mathcal{G}$ is *proper at* ω iff

(8) If $\omega \in G$, then $C(G|\omega) = 1$ *for all* $G \in \mathcal{G}$.

https://doi.org/10.1017/S1755020323000345 Published online by Cambridge University Press

Equivalently, C is proper at ω iff

$$C_\omega \mid \mathcal{G} = \delta_\omega.$$

Here $f \mid d$ is the restriction of function f to domain d. C is *proper* iff it is proper at all $\omega \in \Omega$. Thus, an rcd $P_\mathcal{G}$ is proper iff

$$P_\mathcal{G}(.|\omega) \mid \mathcal{G} = \delta_\omega \qquad \textit{for each } \omega \in \Omega.$$

Propriety is the natural extension of the formula $P(E|E) = 1$ to the rcd formalism. Given Section 2's interpretation of the formalism, it is a highly intuitive desideratum: $P_\mathcal{G}$ encodes a rule for updating credences based upon new certainties over $\mathcal{G}$, so $P_\mathcal{G}$ should preserve those certainties.

Unfortunately, the desideratum is not always satisfiable. Blackwell and Dubins [3] prove the following:

THEOREM: *If $\mathcal{F}$ is a countably generated σ-field, and $\mathcal{G} \subseteq \mathcal{F}$ is sub-σ-field that is not countably generated, and P is a probability measure on $\mathcal{F}$, then no rcd for P given $\mathcal{G}$ is proper.*

As a simple illustration, let λ be Lebesgue measure on Borel subsets $\mathcal{B}$ of $[0, 1]$. Let $\mathcal{C}$ be the sub-σ-field generated by the countable subsets of $[0, 1]$. Thus, $S \in \mathcal{C}$ iff S is countable or S's complement is countable. Note that $\mathcal{B}$ is countably generated but that $\mathcal{C}$ is not. As Billingsley [2, p. 437] observes, the function $\lambda_\mathcal{C} : \mathcal{F} \times \Omega \to \mathbb{R}$ defined by

$$(9)\ \ \lambda_\mathcal{C}(A|\omega) = \lambda(A)$$

is an rcd for λ given $\mathcal{C}$. This function massively violates propriety. In particular, (9) entails

$$\lambda_\mathcal{C}(\{\omega\}\,|\omega) = 0.$$

Intuitively: the probability of an event conditional on itself should be 1, but in this case it is 0! Worse, Seidenfeld, Schervish, and Kadane [42] show that any rcd for λ given $\mathcal{C}$ must satisfy (9) for λ-almost all ω. Each rcd for λ given $\mathcal{C}$ is improper almost everywhere. For further theorems and examples along these lines, see [42].

Propriety becomes more achievable when $\mathcal{G}$ is countably generated, due to the following theorem [42, p. 1614]:

THEOREM: *If $\mathcal{G}$ is countably generated, and $P_\mathcal{G}$ is an rcd for P given $\mathcal{G}$, then $P_\mathcal{G}$ is proper at ω for P-almost all ω.*

This theorem covers cases of conditioning on the value of random variable X, because $\sigma(X)$ is countably generated. Even in the countably generated case, though, one cannot always eliminate the exceptional null event $\{\omega : P_\mathcal{G}$ is improper at $\omega\}$. Only under special assumptions does there exist an rcd that is proper at *all* outcomes ω, as shown by a theorem due to Blackwell and Ryll-Nardzewski [4]:

THEOREM: *Let $(\Omega, \mathcal{F})$ be a Borel space, P a probability measure on $(\Omega, \mathcal{F})$, and X a random variable on $(\Omega, \mathcal{F})$. There exists a proper rcd for P given $\sigma(X)$ only if $X(\Omega)$ is a Borel set.*

There are many random variables X for which $X(\Omega)$ is not a Borel set.

More generally, say that $\Phi : \mathcal{F} \to \mathcal{G}$ is a *selection homomorphism for $\mathcal{G}$ with respect to $\mathcal{F}$* iff **(a)** Φ respects complementation and countable union, and **(b)** $\Phi(G) = G$

https://doi.org/10.1017/S1755020323000345 Published online by Cambridge University Press

for every $G \in \mathcal{G}$. As Sokal [44] notes, the proof techniques from [4] easily extend to establish the following theorem:

THEOREM: *Let $(\Omega, \mathcal{F})$ be a Borel space, P a probability measure on $(\Omega, \mathcal{F})$, and $\mathcal{G} \subseteq \mathcal{F}$ a sub-σ-field. Then there exists a proper rcd for P given $\mathcal{G}$ iff there exists a selection homomorphism for $\mathcal{G}$ with respect to $\mathcal{F}$.*

Selection homomorphisms may not exist even when $\mathcal{G}$ is countably generated.

Some authors, disturbed by impropriety, deny that rcds furnish a good mathematical elucidation of conditional probability [3, 18, 42]. These authors maintain that conditional probabilities should always be proper. They argue on that basis that rcds cannot serve as conditional probabilities.

In this paper, I am not addressing how well rcds elucidate our pre-theoretic notion of conditional probability. I am addressing a slightly different question: do rcds help us model null updating? So I will not explore the challenge that improper rcds pose to the conceptual analysis of conditional probability. Instead, I will explore the challenge that they pose to Kolmogorov conditionalizers.

3.2. Implications of impropriety. In a Kolmogorov certainty acquisition scenario, the agent acquires new certainties over $\mathcal{G}$ and on that basis reallocates credences over the rest of $\mathcal{F}$. Improper updates rules do not provide a good basis for the reallocation. For suppose that update rule C is improper at ω:

$$(10)\quad C_\omega \mid \mathcal{G} \neq \delta_\omega.$$

Then it is impossible to maintain certainty profile δ_ω while also maintaining credences C_ω. The agent cannot update her credences using C, because she cannot simultaneously assign different credences to the same event. Since proper rcds do not always exist even when rcds exist, there are situations where the agent cannot Kolmogorov conditionalize *even though an rcd $P_{\mathcal{G}}$ exists*. If $P_{\mathcal{G}}$ is improper, then the agent cannot follow the credal reallocation policy (7).

Ideally, we would have liked a credal reallocation strategy that works for all possible Kolmogorov certainty acquisition scenarios. It is disappointing that Kolmogorov conditionalization does not give us everything we want. But just how disappointing?

Begin with cases where $\mathcal{G}$ is not countably generated. A thinker who mentally represents putative membership information for $\mathcal{G}$ must be able to entertain uncountably many propositions simultaneously. This is a highly infinitary mental capacity. On that basis, Easwaran [9, p. 143] argues that $\mathcal{G}$ is "irrelevant for probability as degree of belief" because it "can only be grasped by minds that are far more complicated than the ones that we normally attribute subjective probabilities to." I agree. Cases where $\mathcal{G}$ is not countably generated may have mathematical interest, but they are not well-suited to model agents remotely like us. We should not fret that some of them fall outside the scope of Kolmogorov conditionalization.

Relatedly, scientific and philosophical applications of Bayesian decision theory make serious use only of conditioning sub-σ-fields that are countably generated. Scientific applications emphasize learning scenarios that can be modeled with countably generated conditioning sub-σ-fields. To the best of my knowledge, these are the only conditioning sub-σ-fields that play a significant role in Bayesian statistics, probabilistic

https://doi.org/10.1017/S1755020323000345 Published online by Cambridge University Press

robotics, economics, Bayesian cognitive science, or any other field that studies credal reallocation.[3]

Let us turn to cases where $\mathcal{G}$ is countably generated. In these cases, each rcd $P_{\mathcal{G}}$ is proper almost everywhere. More precisely, if we define

$$Prop =_{df} \{\omega : P_{\mathcal{G}} \text{ is proper at } \omega\},$$

then $Prop^c$ has probability 0. At first blush, it might seem that the challenge posed by impropriety is not so worrisome here. If you have credence 0 that impropriety will occur, then you can rest assured that conflicts between your new certainties and your update rule will almost never arise. You can adopt credal reallocation policy (7) and be certain that it will not break down due to a conflict between certainty profile δ_ω and update rule $P_{\mathcal{G}}$.

On further examination, this reassuring line of reasoning is mistaken. Since misplaced certainty is possible, you can acquire new certainties δ_ω even though no index for δ_ω is the true outcome. You may therefore assign a *second-order credence* to the proposition

(11) $\exists\omega$(I will acquire new certainty profile δ_ω & $P_{\mathcal{G}}$ is improper at ω)

that differs from your *first-order* credence in $Prop^c$. Probability space $(\Omega, \mathcal{F}, P)$ may not formally model second-order credences of this kind, but you may nevertheless assign them.[4] In particular, you may assign positive credence to (11) while assigning credence zero to $Prop^c$. Even if you are certain that *Prop* will occur, you may think there is a serious chance that you will acquire future certainties as if $Prop^c$ occurs. For example, you may suspect that an evil demon will manipulate you into acquiring future certainties δ_ω, with $\omega \in Prop^c$. Thus, you may assign positive second-order credence to the possibility that your update rule conflicts with your new certainty profile.

Overall, then, impropriety seems harmless enough when $\mathcal{G}$ is not countably generated but rather disappointing when $\mathcal{G}$ is countably generated. Still, we should keep this setback in perspective. Although Kolmogorov conditionalization is not always usable, it is a very general strategy that can be implemented in numerous Kolmogorov certainty acquisition scenarios.

To see just how far its reach extends, consider the following setup ([11, pp. 26–28], [14, pp. 5–7]), which is general enough to accommodate virtually every scientific application of Bayesian inference. A "parameter space" $(X, \mathcal{X})$ models possible states of the world. A "sample space" $(E, \mathcal{E})$ models possible information the agent will

[3] One of the main examples discussed by Seidenfeld, Schervish, and Kadane [42] is the tail σ-field $\mathcal{A}$ for infinitely many flips of a fair coin. $\mathcal{A}$ contains all events that do not depend upon any finite initial sequence of coin tosses. For example, the event *the 100th coin toss is heads* does not belong to $\mathcal{A}$, while the event *all but finitely many coin tosses are heads* does belong. $\mathcal{A}$, which is not countably generated, generates extreme impropriety when used as a conditioning sub-σ-field. Since tail σ-fields arise naturally in probability theory, one might argue that $\mathcal{A}$ is a conditioning sub-σ-field that occurs in scientific practice yet that is not countably generated. However, while $\mathcal{A}$ figures in probability theory, it does *not* figure in scientific applications that use probability theory to model credence. Thus, I stand by my assertion in the main text: the only conditioning sub-σ-fields that figure in the scientific study of credal reallocation are countably generated sub-σ-fields.

[4] See [13, 39] for discussion of how to model higher-order credence.

https://doi.org/10.1017/S1755020323000345 Published online by Cambridge University Press

receive. $\mathcal{X}$ is a σ-field over X. $\mathcal{E}$ is a σ-field over E. The parameter space and sample space combine into a joint space $(X \times E, \mathcal{X} \otimes \mathcal{E})$, where

$$\mathcal{X} \otimes \mathcal{E} =_{df} \sigma(\mathcal{X} \times \mathcal{E})$$

is the σ-field generated by all measurable rectangles of the form

$$A \times B \qquad\qquad A \in \mathcal{X}, B \in \mathcal{E}.$$

The agent's credences are given by a probability measure P over $\mathcal{X} \otimes \mathcal{E}$. Let $\mathcal{G}$ be the sub-σ-field containing all events

$$X \times B \qquad\qquad B \in \mathcal{E}.$$

$\mathcal{G}$ is the canonical embedding of $\mathcal{E}$ within $\mathcal{X} \otimes \mathcal{E}$. It is easy to show that there exists a selection homomorphism for $\mathcal{G}$ with respect to $\mathcal{X} \otimes \mathcal{E}$ [44]. Thus, assuming $(X \times E, \mathcal{X} \otimes \mathcal{E})$ is Borel, a proper rcd $P_{\mathcal{G}}$ exists. The agent can use $P_{\mathcal{G}}$ to conditionalize. Upon gaining new certainties over $\mathcal{G}$, she can use $P_{\mathcal{G}}$ to reallocate credence over the rest of $\mathcal{X} \otimes \mathcal{E}$. This analysis indicates that, although impropriety may arise in principle, it seldom if ever arises in scientific practice.

3.3. Non-credal interpretations. Given the challenge posed by impropriety, it is worth briefly examining whether the challenge persists if we reinterpret the "information" based on which the agent conditionalizes. I have adopted a *credal* interpretation: I glossed "information" in terms of a change in the agent's credences over $\mathcal{G}$. Specifically, I interpreted the agent as gaining new certainties regarding $\mathcal{G}$, codified through a certainty profile δ_ω, and I asked how the agent should reallocate credences over the rest of $\mathcal{F}$ in light of her newfound certainties over $\mathcal{G}$. But we might instead interpret "information" in *non-credal* terms. We might posit an epistemic state that impacts the agent's credences from outside the credal system.

In principle, there are many possible non-credal interpretations one might explore. Here are two examples from the recent literature:

- In [36], I interpreted the agent as receiving membership *knowledge* regarding $\mathcal{G}$. On this interpretation, $\delta_\omega(A) = 1$ means that the agent knows that the true outcome belongs to A, while $\delta_\omega(A) = 0$ means that the agent knows that the true outcome does not belong to A. The question then becomes how the agent should reallocate her credences over $\mathcal{F}$ in light of her new knowledge regarding $\mathcal{G}$.
- Meehan and Zhang [28] interpret the agent as receiving *evidence* regarding $\mathcal{G}$, where evidence is guaranteed to be veridical. The question then becomes how the agent should reallocate her credences over $\mathcal{F}$ in light of her new evidence regarding $\mathcal{G}$.

Both interpretations are factive. In particular, knowledge is factive: if an agent knows that the true outcome belongs to A, then the true outcome belongs to A.

Suppose we regard the agent as conditionalizing based on membership knowledge regarding $\mathcal{G}$. Knowledge is not certainty: an agent may know that E without being certain that E. It is possible for an agent to know that the true outcome belongs to $G \in \mathcal{G}$ without assigning credence 1 to G. Hence, impropriety no longer poses the challenge raised in Section 3.2. Yet one might worry that impropriety now poses a slightly different challenge. Maybe it is *possible* to know that E while setting $P(E) < 1$, but is

https://doi.org/10.1017/S1755020323000345 Published online by Cambridge University Press

it *rational* to do so? An improper update rule may look unsuited for conditionalizing on knowledge, given that it leads the agent to assign credence < 1 to what she knows. A similar dialectic arises on Meehan and Zhang's [28, p. 729] evidentiary interpretation: it is *possible* to receive evidence E and set $P(E) < 1$, but is it *rational* to do so? Meehan and Zhang claim that it is not. They endorse a norm that they call *Evidential Certainty*: "one should be certain of one's total evidence."

I am skeptical of Evidential Certainty and other similar norms. I also see no reason to suspect that knowledge rationally mandates certainty. For present purposes, though, I set these matters aside. The key point is that, however they play out, impropriety poses a less serious challenge to non-credal factive interpretations than to Section 2's credal interpretation.

Here is why. On a factive interpretation, the agent receives veridical information δ_ω indexed by the true outcome ω. I argued in Section 3.2 that we may restrict attention to countably generated $\mathcal{G}$, so that C is proper at ω for P-almost all ω. Given this restriction, the agent is certain that her update rule will be proper on whatever outcome ω indexes her new veridical information. On the knowledge interpretation, for example, she is certain that she will receive new knowledge encapsulated by a δ_ω such that C is proper at ω. Since the agent regards impropriety as vanishingly unlikely to occur, she need not find it particularly worrisome. For that reason, non-credal factive interpretations look much better positioned than Section 2's credal interpretation with respect to impropriety.

However, my goal in this paper is to explore null updating from a *non-factive* perspective that allows misplaced certainty in conditioning propositions. No treatment that interprets "information" in factive terms can capture all the credal updates we would like to capture: factive treatments neglect numerous credal updates that occur routinely in scientific practice. So interpretations centered around knowledge or non-credal factive evidence are too limited in scope for my purposes.

One might pursue a *non-credal, non-factive* interpretation of the agent's "information" regarding $\mathcal{G}$. For example, one might articulate a non-factive conception of "evidence" and treat the agent as gaining evidence regarding $\mathcal{G}$. I am not sure what this conception would look like or how it would adjudicate the challenge posed by impropriety. Pursuing such questions would inevitably embroil us in many epistemological complexities.

In contrast, the credal interpretation is straightforward and requires only notions that we already need when doing Bayesian epistemology. The credal interpretation also lets us study credal updates on their own terms, ignoring whatever non-credal states prompt the updates. So I think we do well to explore the credal interpretation, which is the task I have undertaken in the present paper. Once we adopt the credal interpretation, the challenge raised in Section 3.2 arises.

3.4. Alternative update rules. To gain a better perspective on the challenge, let us compare Kolmogorov conditionalization with alternative credal reallocation strategies one might pursue.

I confine attention to strategies based on update rules. Imagine an agent who adopts the following credal reallocation policy:

(12) *Respond to new certainty profile δ_ω over $\mathcal{G}$ by adopting new credences C_ω over $\mathcal{F}$,*

where C is an update rule but not an rcd, i.e., C satisfies (4) and (5) but does not satisfy (6). (12) faces a problem parallel to that faced by (7): C may be improper at ω. The

https://doi.org/10.1017/S1755020323000345 Published online by Cambridge University Press

theorems proved by Blackwell and Ryll-Nardzewski [4] are more general than stated in Section 3.2. The general theorems state:

THEOREM: *Let* $(\Omega, \mathcal{F})$ *be a Borel space and* X *a random variable on* $(\Omega, \mathcal{F})$. *There exists a proper update rule for* $(\Omega, \mathcal{F})$ *and* $\sigma(X)$ *only if* $X(\Omega)$ *is a Borel set*.

THEOREM: *Let* $(\Omega, \mathcal{F})$ *be a Borel space and* $\mathcal{G} \subseteq \mathcal{F}$ *a sub-*σ*-field. There exists a proper update rule for* $(\Omega, \mathcal{F})$ *and* $\mathcal{G}$ *iff there exists a selection homomorphism for* $\mathcal{G}$ *with respect to* $\mathcal{F}$.

When C is improper at ω, it is impossible to maintain certainty profile δ_ω while simultaneously maintaining credences C_ω. Thus, discarding the integral formula does not eliminate the challenge posed by impropriety.

A theorem proved by Ramachandran [32] shows that, under very natural assumptions, discarding the integral formula does not ameliorate impropriety at all:

THEOREM: *Let* $(\Omega, \mathcal{F}, P)$ *be a probability space and* $\mathcal{G} \subseteq \mathcal{F}$ *a sub-*σ*-field. There exists a proper rcd for* P *given* $\mathcal{G}$ *iff*: (i) *there exists an rcd for* P *given* $\mathcal{G}$ *that is proper at* ω *for* P*-almost all* ω; *and* (ii) *there exists a proper update rule for* $(\Omega, \mathcal{F})$ *and* $\mathcal{G}$.

I argued in Section 3.2 that we may confine attention to cases where $\mathcal{G}$ is countably generated, so that condition (i) of the theorem is satisfied. In such cases, a proper rcd exists iff a proper update rule exists. Abandoning the integral formula yields no progress whatsoever regarding impropriety.

As this dialectic reveals, the disappointment engendered by impropriety is not specific to rcds. It arises for all update rules, whether or not one adopts the integral formula as a constraint on conditional probabilities. Using an update rule other than an rcd is no remedy.

Some readers may suspect that we should look beyond update rules. Perhaps we should lift the requirement (4) that new credences be codified by a probability measure? Or lift the measurability requirement (5)? Neither option seems promising to me:

- The only controversial element in (4) is the requirement that credences be countably additive. Some authors contend that we should abandon countable additivity, requiring only that credences be *finitely* additive (e.g., [7, 19, 42]). However, countable additivity figures crucially in probability theory and in many scientific applications of Bayesian decision theory. Abandoning it would require massive revisions to scientific practice. I doubt that a finitely additive approach can replicate the explanatory and pragmatic success achieved by the standard countably additive framework.
- The measurability requirement (5) is highly plausible, as argued in Section 2.2. Measurability formalizes the intuitive thought that new credal assignments are embedded in the agent's implicit certainties regarding $\mathcal{G}$.

For these reasons, I restrict attention to update rules. Given the restriction to update rules, impropriety poses no special problems for rcds.

§4. Dutch book theorems. I will now show that rcds have distinctive pragmatic benefits that privilege them over other update rules. I will prove a *Dutch book theorem* and *converse Dutch book theorem* tailored to Kolmogorov conditionalization.

As mentioned in Section 1, Lewis [26] proves a Dutch book theorem for ratio conditionalization, and Skyrms [43] proves a converse theorem. Lewis and Skyrms

https://doi.org/10.1017/S1755020323000345 Published online by Cambridge University Press

assume that the conditioning proposition is *true* and that it has *non-zero initial probability*. They do not address scenarios where misplaced certainty is possible or where null updating occurs. Subsequent work generalizes the Lewis–Skyrms theorems as follows:

- [38] extends the Lewis–Skyrms theorems to a non-factive setting. This paper proves a Dutch book theorem and converse Dutch book theorem concerning a class of learning scenarios where misplaced certainty is possible. The theorems do not cover null updating scenarios.
- [36] extends the Lewis–Skyrms theorems to a setting where null updating can occur. This paper proves a Dutch book theorem and converse Dutch book theorem tailored to *Kolmogorov knowledge acquisition scenarios* (i.e., scenarios where the agent gains membership knowledge for a sub-σ-field $\mathcal{G}$ and updates her credences on that basis).[5] Since knowledge is factive, the theorems do not cover scenarios where the conditioning proposition is false.

[38] lifts the restriction to true conditioning propositions. [36] lifts the restriction to propositions with positive initial probability. In what follows, I lift both restrictions simultaneously. I prove a Dutch book theorem and converse Dutch book theorem tailored to Kolmogorov certainty acquisition scenarios.

In Section 4.1, I formalize key notions such as *bet*, *bookie strategy*, and *Dutch book*. Section 4.2 reviews literature that applies these formalized notions within a factive setting. Section 4.3 proves a non-factive Dutch book theorem and converse Dutch book theorem.

4.1. Dutch books formalized. Consider an agent who at time t_1 has credences modeled by probability space $(\Omega, \mathcal{F}, P)$ and at time t_2 gains new certainties δ_ω over $\mathcal{G} \subseteq \mathcal{F}$. We want to compare possible strategies for reallocating credence based on the new certainties. We assume that the agent uses a proper update rule $C : \mathcal{F} \times \Omega \to \mathbb{R}$, satisfying constraints (4), (5), and (8). She responds to new certainties δ_ω by forming new credences C_ω. Thus, her credal reallocation policy is given by (12). Assumption (4) guarantees that C_ω is a probability measure. If $C(A|.)$ satisfies the integral formula for all $A \in \mathcal{F}$, then C is an rcd for P given $\mathcal{G}$. I do not assume that $C(A|.)$ satisfies the integral formula.

The agent faces a bookie who can offer bets at both t_1 and t_2. Following standard practice in probability theory, I formalize a *bet* as a random variable. It is convenient to allow random variables that take values in the *extended real line* $\overline{\mathbb{R}} = [-\infty, \infty]$. Thus, a bet is an $\mathcal{F}$-measurable function $X : \Omega \to \overline{\mathbb{R}}$. Here $X(\omega)$ is the net gain for outcome ω. Write

$$E_\mu[X] =_{df} \int_\Omega X d\mu$$

for the expectation of random variable X with respect to probability measure μ. A bet X is *acceptable relative to* μ iff its expected value with respect to μ is non-negative:

[5] In [36], I called these *Kolmogorov learning scenarios*. My terminology here is more informative albeit more cumbersome.

https://doi.org/10.1017/S1755020323000345 Published online by Cambridge University Press

$$E_\mu[X] =_{df} \int_\Omega X d\mu \geq 0.$$

At t_2, the agent acquires updated credences C_ω based on new certainties δ_ω. Say that X is *acceptable given* δ_ω iff its expected value with respect to C_ω is non-negative:

$$E_{C_\omega}[X] = \int X dC_\omega \geq 0.$$

This definition is well-formed, since $\mathcal{G}$-measurability ensures that

$$\delta_\omega = \delta_\nu \rightarrow C_\omega = C_\nu.$$

I assume that our agent adopts a policy of accepting all acceptable bets that are offered and rejecting all unacceptable bets.

Lewis and Skyrms assume that the bookie learns at t_2 which conditioning proposition is true. In [36], I made a comparable assumption when extending the Dutch book results to Kolmogorov knowledge acquisition scenarios: I assumed that the bookie at t_2 learns membership facts for $\mathcal{G}$. These assumptions no longer seem appropriate once we broaden the allowed learning scenarios to include misplaced certainty. After all, if the bookie learns membership facts for $\mathcal{G}$ but the agent has no such luck, then the bookie's ability to exploit his superior knowledge for sure gain hardly indicates that the agent's credal reallocation policy is problematic. So a non-factive setting requires us to rethink our assumptions regarding the bookie's epistemic state at t_2. What now seems most appropriate is to assume that the agent and the bookie acquire the same certainty profile δ_ω. Whereas I assumed in [36] that agent and bookie acquire the same membership *knowledge* regarding $\mathcal{G}$, I now assume that they acquire the same *certainties* regarding $\mathcal{G}$. Thus, the bookie is no better and no worse off than the agent when it comes to membership facts about $\mathcal{G}$.

Based on his new certainty profile over $\mathcal{G}$, the bookie decides what bet (if any) to offer at t_2. The bookie does this using a bookie strategy, which maps certainty profiles δ_ω into bets. We can formalize these ideas through the measurable space $(\Omega \times \Omega, \mathcal{G} \otimes \mathcal{F})$. A *bookie strategy* is a $\mathcal{G} \otimes \mathcal{F}$-measurable function $Y : \Omega \times \Omega \rightarrow \overline{\mathbb{R}}$. For fixed ω, $Y(\omega, .) : \Omega \rightarrow \overline{\mathbb{R}}$ is the bet that the bookie offers upon acquiring new certainties δ_ω over $\mathcal{G}$. Since Y is $\mathcal{G} \otimes \mathcal{F}$-measurable, one can easily show that

$$Y(\omega, .) : \Omega \rightarrow \overline{\mathbb{R}} \text{ is } \mathcal{F}\text{-measurable} \qquad \textit{for every } \omega \in \Omega,$$

so that $Y(\omega, .)$ is indeed a bet according to our official definition. One can also show that

$$Y(., \nu) : \Omega \rightarrow \overline{\mathbb{R}} \text{ is } \mathcal{G}\text{–measurable} \qquad \textit{for every } \nu \in \Omega,$$

which ensures that

$$\delta_\omega = \delta_\rho \rightarrow Y(\omega, .) = Y(\rho, .).$$

Thus, a bookie strategy carries each certainty profile δ_ω to a well-defined bet, independent of the index ω. I will often abbreviate $Y(\omega, .)$ as Y_ω. To model scenarios where the bookie offers no bet in light of new certainty profile δ_ω, I set $Y_\omega(\nu) = 0$ for all inputs ν.

https://doi.org/10.1017/S1755020323000345 Published online by Cambridge University Press

The $\mathcal{G} \otimes \mathcal{F}$-measurability requirement ensures that the event $Y^{-1}(-\infty, a] = \{(\omega, v) : Y_\omega(v) \leq a\}$ belongs to $\mathcal{G} \otimes \mathcal{F}$. This event is notable, for the following reason:

(13) The bet offered by the bookie has net gain $\leq a$ iff $(\omega, v) \in Y^{-1}(-\infty, a]$, where ω is any index for the bookie's newly acquired certainty profile and v is the true outcome.

We may employ $Y^{-1}(-\infty, a]$ as a formal proxy for the proposition *The bet offered by the bookie has net gain* $\leq a$. $\mathcal{G} \otimes \mathcal{F}$-measurability ensures that the formal proxy belongs to $\mathcal{G} \otimes \mathcal{F}$.

We may use $\mathcal{G} \otimes \mathcal{F}$ to model the implicit knowledge of an observer who learns which new certainty profile over $\mathcal{G}$ the agent and bookie acquire at time t_2 *and* who learns which events in $\mathcal{F}$ occurred. Such an observer learns whether

$$\delta_\omega(G) = 1, \qquad \textit{for each } G \in \mathcal{G},$$

where ω is any index for the agent's new certainty profile. Equivalently, the observer learns whether $\omega \in G$, for each $G \in \mathcal{G}$. The observer also learns whether the true outcome v belongs to F, for each $F \in \mathcal{F}$. In principle, the observer can extrapolate from his knowledge regarding $\mathcal{G}$ and $\mathcal{F}$ to knowledge regarding $\mathcal{G} \otimes \mathcal{F}$. For each $D \in \mathcal{G} \otimes \mathcal{F}$, he gains implicit knowledge whether

$$(\omega, v) \in D,$$

where ω is any index for the agent's new certainty profile and v is the true outcome. Assuming $\mathcal{G} \otimes \mathcal{F}$-measurability, our observer gains implicit knowledge whether $(\omega, v) \in Y^{-1}(-\infty, a]$, where ω is any index for the agent's new certainty profile and v is the true outcome. In other words, he gains implicit knowledge whether the bet offered by the bookie has net gain $\leq a$. On the other hand, if $Y^{-1}(-\infty, a]$ does not belong to $\mathcal{G} \otimes \mathcal{F}$, then the observer does not gain implicit knowledge whether the selected bet has net gain $\leq a$. Hence, $\mathcal{G} \otimes \mathcal{F}$-measurability codifies the following intuitive thought: an observer who knows which new certainty profile the agent and bookie acquire *and* which events in $\mathcal{F}$ occurred thereby gains implicit knowledge whether the bet selected by the bookie has net gain $\leq a$.

We may now formalize Dutch bookability. I consider two formalizations, corresponding respectively to the intuitive notions *sure loss in Kolmogorov knowledge acquisition scenarios* and *sure loss in Kolmogorov certainty acquisition scenarios*. A ***factive Dutch book for*** $(\boldsymbol{\Omega}, \boldsymbol{\mathcal{F}}, \boldsymbol{P})$***, sub-σ-field*** $\boldsymbol{\mathcal{G}} \subseteq \boldsymbol{\mathcal{F}}$***, and update rule C*** is a pair (X, Y) such that, for all $\omega \in \Omega$:

(i) X is a bet that is acceptable relative to P.
(ii) Y is a bookie strategy.
(iii) Y_ω is acceptable given δ_ω.
(iv) $X(\omega) + Y_\omega(\omega) < 0$.

Condition (i) requires that X have nonnegative expectation at t_1. This condition ensures that the agent will accept bet X at t_1. Conditions (iii) and (iv) jointly ensure that a bookie who offers bet X and pursues bookie strategy Y will inflict a sure loss *assuming that the agent and the bookie learn true membership facts for* $\mathcal{G}$. A ***non-factive Dutch book for*** $(\boldsymbol{\Omega}, \boldsymbol{\mathcal{F}}, \boldsymbol{P})$***, sub-σ-field*** $\boldsymbol{\mathcal{G}} \subseteq \boldsymbol{\mathcal{F}}$***, and update rule C*** is a pair (X, Y) such that for all $\omega, v \in \Omega$:

https://doi.org/10.1017/S1755020323000345 Published online by Cambridge University Press

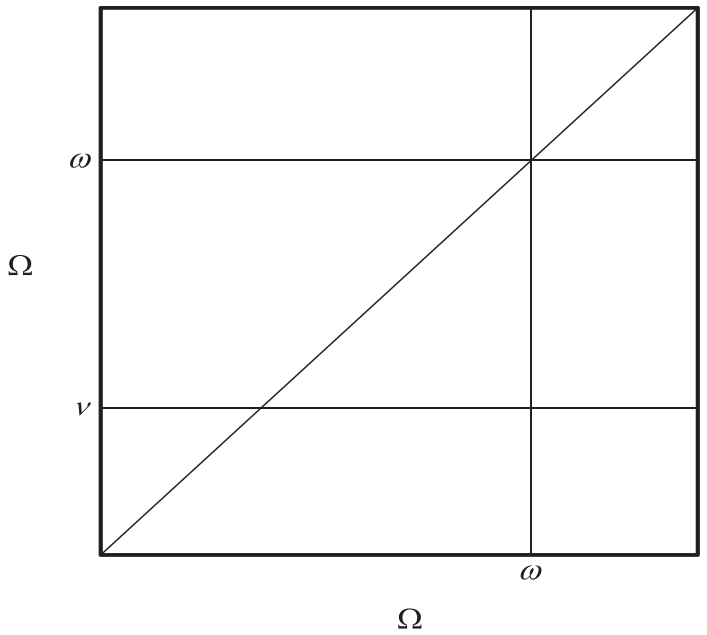


Figure 1. A visualization of $\Omega \times \Omega$. In general, there need not be any natural linear ordering of Ω. Nevertheless, the visualization is a useful heuristic. The horizontal axis corresponds to $(\Omega, \mathcal{G})$. Points on this axis index the agent's new certainty profile, which determines the bet selected by the bookie. When the agent acquires certainty profile δ_ω, the bookie offers bet Y_ω. The vertical axis corresponds to $(\Omega, \mathcal{F})$. Points on this axis reflect the true outcome. We specify how a gambling scenario unfolds by specifying an ordered pair (ω, ν), where ω is an index for the agent's new certainty profile and ν is the true outcome. Each ordered pair (ω, ν) determines a well-defined net gain $Y_\omega(\nu)$. A factive Dutch book inflicts a net loss along the diagonal. A non-factive Dutch book inflicts a net loss on all points.

(i) X is a bet that is acceptable relative to P.
(ii) Y is a bookie strategy.
(iii) Y_ω is acceptable given δ_ω.
(iv*) $X(\nu) + Y_\omega(\nu) < 0$.

Conditions (iii) and (iv*) jointly ensure that a bookie who offers bet X and pursues bookie strategy Y will inflict a sure loss *whether or not the agent and bookie become certain of true membership facts for* $\mathcal{G}$. A non-factive Dutch book ensures a net loss even in situations where agent and bookie acquire misplaced certainties regarding $\mathcal{G}$. See Figure 1.

4.2. Factive Dutch books. The Dutch book theorem and converse Dutch book theorem from [36] are tailored to Kolmogorov knowledge acquisition scenarios:

Factive Dutch Theorem for Kolmogorov Conditionalization: *Let* $(\Omega, \mathcal{F}, P)$ *be a probability space, let* $\mathcal{G} \subseteq \mathcal{F}$ *be a sub-*σ*-field, and let* C *be an update rule for* $(\Omega, \mathcal{F})$ *and* $\mathcal{G}$. *If* C *is not an rcd for* P *given* $\mathcal{G}$, *then there exists a factive Dutch book for* $(\Omega, \mathcal{F}, P)$, $\mathcal{G}$, *and* C.

https://doi.org/10.1017/S1755020323000345 Published online by Cambridge University Press

Factive Converse Dutch Book Theorem for Kolmogorov Conditionalization: *Let* $(\Omega, \mathcal{F}, P)$ *be a probability space, let* $\mathcal{G} \subseteq \mathcal{F}$ *be a sub-*σ*-field, and let* C *be an rcd for* P *given* $\mathcal{G}$. *Then there does not exist a factive Dutch book for* $(\Omega, \mathcal{F}, P)$, $\mathcal{G}$, *and* C.

The two theorems show that rcds offer unique pragmatic advantages in Kolmogorov knowledge acquisition scenarios.

In [36], I motivated the $\mathcal{G} \otimes \mathcal{F}$-measurability requirement on bookie strategies along lines similar to Section 4.1. I spoke of "membership information" rather than "certainty profiles," but otherwise my reasoning was essentially the same. I offered the event $Y^{-1}(-\infty, a]$ as a formal proxy for the proposition *The bet offered by the bookie has net gain* $\leq a$. $\mathcal{G} \otimes \mathcal{F}$-measurability then codifies the following intuitive thought: an observer who knows which new membership information the agent and bookie acquire *and* which events in $\mathcal{F}$ occurred thereby gains implicit knowledge whether the bet offered by the bookie has net gain $\leq a$.

Meehan and Zhang [28] argue that, assuming a factive conception of credal updates, the $\mathcal{G} \otimes \mathcal{F}$-measurability requirement on bookie strategies is misguided. Specifically, they deny that $Y^{-1}(-\infty, a]$ is the correct formal proxy for the proposition *The bet offered by the bookie has net gain* $\leq a$. In a factive setting, the bookie will offer bet Y_ω *where* ω *is the true outcome.* The bet's net gain is $Y_\omega(\omega)$. Accordingly, Meehan and Zhang urge that the correct formal proxy is

$$\{\omega : Y_\omega(\omega) \leq a\}.$$

They offer a revised notion of Dutch book, which requires

(14) $\{\omega : Y_\omega(\omega) \leq a\} \in \mathcal{F}$

rather than

$$Y^{-1}(-\infty, a] = \{(\omega, v) : Y_\omega(v) \leq a\} \in \mathcal{G} \otimes \mathcal{F}$$

in order for Y to count as a bookie strategy. They prove a Dutch book theorem and converse Dutch book theorem using the revised notion of Dutch book. The theorems establish that a Dutch book (in the sense favored by Meehan and Zhang) does not exist for update rule C iff C is an rcd that is proper almost everywhere. Meehan and Zhang conclude that rcds *per se* do not offer unique pragmatic advantages over alternative update rules; only rcds that are proper almost everywhere offer unique pragmatic advantages.

In my opinion, $\{\omega : Y_\omega(\omega) \leq a\}$ is not a good formal proxy for the proposition *The bet offered by the bookie has net gain* $\leq a$. Assuming a factive setting,

(15) The bookie will acquire membership information δ_ω indexed by the true outcome ω.

It then follows that

(16) The bet offered by the bookie has net gain $\leq a$ iff the true outcome belongs to $\{\omega : Y_\omega(\omega) \leq a\}$.

However, the bookie may not know (15) even if the agent knows it. If the bookie does not know (15), then he is not positioned to know (16) either. But then the bookie, in learning that the true outcome belongs to $\{\omega : Y_\omega(\omega) \leq a\}$, will not thereby learn that his bet has net gain $\leq a$. Furthermore, even if the agent and the bookie both know (15),

https://doi.org/10.1017/S1755020323000345 Published online by Cambridge University Press

the agent may not know that the bookie knows (15). But then the agent does not know that the bookie, by learning that the true outcome belongs to $\{\omega : Y_\omega(\omega) \leq a\}$, thereby learns that his bet has net gain $\leq a$. For that reason, the agent does not know that the bookie can assess the outcome of his own bet. For all the agent knows, the bookie cannot detect whether he won the bet!

Dutch books are supposed to inflict a sure loss. A sure loss means that, from the agent's viewpoint, the agent is guaranteed to lose [39]. If the bookie does not know that the agent lost the bet, then the agent will not actually lose money because the bookie will never collect his winnings. If the agent does not expect the bookie to know that the agent lost money, then the agent does not regard a loss as guaranteed. A loss that is undetectable by the bookie or that the agent thinks is undetectable by the bookie does not constitute a "sure loss."

(14) allows a "loss" that the bookie never detects. It also allows a "loss" that the bookie detects but that the agent does not expect the bookie to detect. In neither case does the book inflict a sure loss. So (14) does not secure genuine Dutch bookability. In contrast, suppose we impose the $\mathcal{G} \otimes \mathcal{F}$-measurability requirement. Then, as argued in Section 4.1, the bookie can detect whether he won the bet provided that he learns which new information was acquired *and* learns which events in $\mathcal{F}$ occurred. I conclude that we should impose the $\mathcal{G} \otimes \mathcal{F}$-measurability requirement on bookie strategies even for the special case of factive Dutch books.[6]

However things stand with factive Dutch books, $\mathcal{G} \otimes \mathcal{F}$-measurability is surely needed for a suitable conception of non-factive Dutch books. Once we lift all factivity assumptions, the bookie may acquire a certainty profile δ_ω such that

$$\delta_\omega \neq \delta_\nu,$$

where ν is the true outcome. In that context, $Y^{-1}(-\infty, a]$ is clearly a much better formal proxy than $\{\omega : Y_\omega(\omega) \leq a\}$ for the proposition *The bet offered by the bookie has net gain* $\leq a$. Even if we agree with Meehan and Zhang that $\mathcal{G} \otimes \mathcal{F}$-measurability is a misguided constraint on bookie strategies in the factive case, we should embrace it in the non-factive case that I am addressing in the present paper.

4.3. Non-factive Dutch books. I now prove a Dutch book theorem and converse Dutch book theorem tailored to Kolmogorov certainty acquisition scenarios.

Non-Factive Dutch Book Theorem for Kolmogorov Conditionalization: *Let* $(\Omega, \mathcal{F}, P)$ *be a probability space, and let* C *be a proper update rule for* $(\Omega, \mathcal{F})$ *and* $\mathcal{G}$. *If* C *is not an rcd for* P *given* $\mathcal{G}$, *then there exists a non-factive Dutch book for* $(\Omega, \mathcal{F}, P)$, $\mathcal{G}$, *and* C.

Proof: Since C is not an rcd for P given $\mathcal{G}$, $C(A|.)$ must violate the integral formula for some $A \in \mathcal{F}$. It follows that

$$P\{\omega : C(A|\omega) \neq P_{\mathcal{G}}(A|\omega)\} > 0,$$

[6] $(\Omega \times \Omega, \mathcal{G} \otimes \mathcal{F})$ extends the measurable space $(\Omega, \mathcal{F})$ we are using to model the agent's credal states at t_1 and t_2. We may imagine the agent and the bookie, once all bets have been offered, reflecting on the past gambling interaction and transitioning to knowledge modeled by the larger space $(\Omega \times \Omega, \mathcal{G} \otimes \mathcal{F})$. After the transition, the agent and the bookie both gain at least implicit knowledge regarding the outcome of the bet offered at t_2.

https://doi.org/10.1017/S1755020323000345 Published online by Cambridge University Press

Table 1. *The bet has payoff 1 if A occurs, but it is cancelled if G does not occur. Furthermore, the price of the bet $P_{\mathcal{G}}(A|v)$ depends on the true outcome v. At time t_1, the bookie offers to sell this conditional bet to the agent. The final column gives the agent's net gain if she accepts the bookie's offer*

Outcome	Price	Payoff	Net gain
$v \in A \cap G$	$P_{\mathcal{G}}(A\|v)$	1	$1 - P_{\mathcal{G}}(A\|v)$
$v \in A^c \cap G$	$P_{\mathcal{G}}(A\|v)$	0	$-P_{\mathcal{G}}(A\|v)$
$v \notin G$	0	0	0

Table 2. *At time t_2, the bookie offers to buy this conditional bet from the agent. The final column gives the agent's net gain if she accepts the bookie's offer. Net gain is computed by subtracting the payoff from the price, since the bookie is offering to buy rather than sell the bet*

Outcome	Price	Payoff	Net gain
$v \in A \cap G$	$C(A\|v)$	1	$C(A\|v) - 1$
$v \in A^c \cap G$	$C(A\|v)$	0	$C(A\|v)$
$v \notin G$	0	0	0

Table 3. *Net gain for the overall gambling scenario*

Outcome	Net gain
$v \in A \cap G$	$C(A\|v) - P_{\mathcal{G}}(A\|v)$
$v \in A^c \cap G$	$C(A\|v) - P_{\mathcal{G}}(A\|v)$
$v \notin G$	0

where $P_{\mathcal{G}}(A|.)$ is some conditional probability for A given $\mathcal{G}$. Note that

$$\{\omega : C(A|\omega) \neq P_{\mathcal{G}}(A|\omega)\} = \{\omega : C(A|\omega) < P_{\mathcal{G}}(A|\omega)\} \cup \{\omega : C(A|\omega) > P_{\mathcal{G}}(A|\omega)\}.$$

Both sets on the right-hand side belong to $\mathcal{G}$. At least one of these two sets must have non-zero P-measure. Without loss of generality, suppose the first does. Call this set G:

$$G =_{df} \{\omega : C(A|\omega) < P_{\mathcal{G}}(A|\omega)\}.$$

We will use G to rig the Dutch book. Informally, we proceed as follows. At t_1, the bookie offers a conditional bet given by Table 1. At t_2, the bookie offers a new conditional bet given by Table 2. Table 3 summarizes net gain for the two bets, given any outcome $v \in \Omega$. Net gain is 0 when $v \notin G$ and negative when $v \in G$. To ensure a net gain for *all* outcomes v, we add a side bet on G at t_1. We thereby achieve a non-factive Dutch book. In the case where $\{\omega : C(A|\omega) > P_{\mathcal{G}}(A|\omega)\}$ has non-zero P-measure, we merely reverse payoffs from all the aforementioned bets.

https://doi.org/10.1017/S1755020323000345 Published online by Cambridge University Press

Let us now formalize these ideas. As always, we codify bets as random variables. First, define random variable X by

$$X(v) = \begin{cases} 1 - P_{\mathcal{G}}(A|v) & \text{if } v \in A \cap G, \\ -P_{\mathcal{G}}(A|v) & \text{if } v \in A^c \cap G, \\ 0 & \text{if } v \notin G. \end{cases}$$

Intuitively, this is the bet offered at t_1. Define bookie strategy $Y(\omega, v)$ by

$$Y(\omega, v) = \begin{cases} C(A|v) - 1 & \text{if } v \in A \cap G, \\ C(A|v) & \text{if } v \in A^c \cap G, \\ 0 & \text{if } v \notin G. \end{cases}$$

Note that $Y(\omega, v)$ does not depend upon ω. Intuitively: the bookie offers the bet from Table 2 no matter which new certainties over $\mathcal{G}$ the agent gains at t_2. Let $L =_{df} \int_G \big[C(A|v) - P_{\mathcal{G}}(A|v)\big]\, dP(v)$ if this integral is finite. If the integral is infinite, then let L be any finite negative number. Either way, we have

$$\int_G \big[C(A|v) - P_{\mathcal{G}}(A|v)\big]\, dP(v) \leq L.$$

Define random variable Z:

$$Z(v) = \begin{cases} (P(G) - 1)\big[C(A|v) - P_{\mathcal{G}}(A|v)\big] & \text{if } v \in G, \\ L & \text{if } v \notin G. \end{cases}$$

This is the side bet on G advertised above. We show that $(X + Z, Y)$ is a non-factive Dutch book for $(\Omega, \mathcal{F}, P)$, $\mathcal{G}$, and C.

One can easily check that Y is a bookie strategy and that $(X + Z, Y)$ satisfies clause (iv*) in the definition of non-factive Dutch book. To prove that X is acceptable relative to P, we write

$$X(v) = I_G(v)\big[I_A(v) - P_{\mathcal{G}}(A|v)\big],$$

where, for any set S, I_S is the indicator function for S:

$$I_s(\omega) = \begin{cases} 1 & \text{if } \omega \in S, \\ 0 & \text{if } \omega \notin S. \end{cases}$$

Routine computation using the integral formula confirms that

$$E_P[X] = \int_\Omega I_G(v)\big[I_A(v) - P_{\mathcal{G}}(A|v)\big]\, dP(v) = 0.$$

Similarly, we write

$$Z(v) = I_G(v)(P(G) - 1)\big[C(A|v) - P_{\mathcal{G}}(A|v)\big] + I_{G^c}(v)L,$$

and routine computation confirms that

$$E_P[Z] = \int_\Omega \big[I_G(v)(P(G) - 1)\big[C(A|v) - P_{\mathcal{G}}(A|v)\big] + I_{G^c}(v)L\big]\, dP(v) \geq 0.$$

https://doi.org/10.1017/S1755020323000345 Published online by Cambridge University Press

See [36] for the details of these computations. Thus, $X + Z$ is acceptable relative to P. To show that Y satisfies clause (iii), pick any ω. We have assumed that C is proper at ω:

(17) $C_\omega(B) = 1$ if $\omega \in B$,
(18) $C_\omega(B) = 0$ if $\omega \notin B$,

for all $B \in \mathcal{G}$. We will show that Y_ω's expectation is 0 relative to C_ω:

$$E_{C_\omega}[Y_\omega] = \int Y_\omega dC_\omega = 0.$$

Define

$$J =_{df} \{\pi : C(A|\omega) = C(A|\pi)\}.$$

Since $\omega \in J \in \mathcal{G}$, (17) and (18) entail that

$$C_\omega(J) = 1,$$

$$C_\omega(J^c) = 0.$$

We may write

$$Y_\omega(v) = I_G(v)\left[C(A|v) - I_A(v)\right],$$

and then compute

$$\begin{aligned}
E_{C_\omega}[Y_\omega] &= \int_\Omega I_G\left[C(A|v) - I_A\right]dC_\omega = \int_\Omega I_G C(A|v)dC_\omega - \int_\Omega I_G I_A dC_\omega \\
&= \int_G C(A|v)dC_\omega - \int_{G\cap A} dC_\omega = \int_{G\cap J} C(A|v)dC_\omega + \int_{G\cap J^c} C(A|v)dC_\omega - \int_{G\cap A} dC_\omega \\
&= \int_{G\cap J} C(A|v)dC_\omega + 0 - C_\omega(G\cap A) = \int_{G\cap J} C(A|\omega)dC_\omega - C(A\cap G|\omega) \\
&= C(A|\omega)\int_{G\cap J} dC_\omega - C(A\cap G|\omega) = C(A|\omega)C(G\cap J|\omega) - C(A\cap G|\omega),
\end{aligned}$$

where ω in these equations is held fixed and is not an integration variable. Now, either $\omega \in G$ or $\omega \notin G$. In the first case, we have $\omega \in G \cap J \in \mathcal{G}$, so (17) ensures that

$$C(G\cap J|\omega) = 1.$$

(17) also ensures that $C(G|\omega) = 1$, so that

$$C(A\cap G|\omega) = C(A|\omega).$$

Thus, our expression for Y_ω's expectation reduces to

$$E_{C_\omega}[Y_\omega] = C(A|\omega)C(G\cap J|\omega) - C(A\cap G|\omega) = C(A|\omega) - C(A|\omega) = 0.$$

If on the other hand $\omega \notin G$, then (18) ensures that

$$C(G\cap J|\omega) = 0$$

https://doi.org/10.1017/S1755020323000345 Published online by Cambridge University Press

and

$$C(A \cap G|\omega) = 0,$$

so our expression for Y_ω's expectation again reduces to 0. Either way, Y_ω's expectation is 0. Hence, $(X + Z, Y)$ is a non-factive Dutch book for $(\Omega, \mathcal{F}, P)$, $\mathcal{G}$, and C. □

We can strengthen the theorem so that all bets have *positive* expected net payoff. To accomplish this, we add a small "sweetener" that is large enough to boost expected net payoffs above zero but small enough not to cancel out the net loss inflicted by the original book. See [36] for details regarding the factive case; those details require only slight modification to fit the non-factive case.

Non-Factive Converse Dutch Book Theorem for Kolmogorov Conditionalization: *Let $(\Omega, \mathcal{F}, P)$ be a probability space, and let C be an rcd for P given $\mathcal{G}$. Then there does not exist a non-factive Dutch book for $(\Omega, \mathcal{F}, P)$, $\mathcal{G}$, and C.*

Proof: Clause (iv*) in the definition of non-factive Dutch book entails clause (iv) in the definition of factive Dutch book. So a non-factive Dutch book for $(\Omega, \mathcal{F}, P)$, $\mathcal{G}$, and C is a factive Dutch book for $(\Omega, \mathcal{F}, P)$, $\mathcal{G}$, and C. The Factive Converse Dutch Book Theorem entails that a factive Dutch book does not exist. It follows that a non-factive Dutch book does not exist. □

The non-factive Dutch book theorem confines attention to proper update rules. It states that, if a proper update rule violates the integral formula, then a non-factive Dutch book exists. Note the sharp contrast with the factive Dutch book theorem proved in [36], which does not assume propriety. The contrast reflects a crucial difference between Kolmogorov certainty acquisition scenarios and Kolmogorov knowledge acquisition scenarios. In a Kolmogorov certainty acquisition scenario, the agent gains new certainties regarding $\mathcal{G}$ and updates her credences on that basis. The only update rules that can deliver such an update are proper update rules—hence the restriction to proper updates rules in the non-factive Dutch book theorem. In a Kolmogorov knowledge acquisition scenario, the agent gains *knowledge* regarding $\mathcal{G}$ and updates her credences on that basis. As emphasized in Section 3.4, knowledge is not certainty. An agent who comes to know E does not necessarily become certain that E. It is possible in principle for the agent to use an improper update rule—hence the attention to all update rules, and not just proper update rules, in the factive Dutch book theorem.

There is also a sharp contrast between the role that propriety plays in my discussion and the role that it plays in Meehan and Zhang's discussion. Meehan and Zhang attempt to show that impropriety induces Dutch bookability. Section 4.3 raised doubts about whether the attempt succeeds. Be that as it may, the non-factive Dutch book theorem does not even purport to show that impropriety induces Dutch bookability. Propriety figures in Meehan and Zhang's discussion as a putative rational norm that an update policy may or may not violate. It figures in my discussion as a usability condition: only proper update rules can be used to update credences in light of new certainties, because impropriety induces a conflict (10) between the update rule and the new certainties. My setup treats propriety as a necessary condition on usable update rules rather than a putative rational norm that requires justification.

https://doi.org/10.1017/S1755020323000345 Published online by Cambridge University Press

§5. Significance of the theorems. Taken together, the non-factive Dutch book theorem and non-factive converse Dutch book theorem show that Kolmogorov conditionalization has highly desirable pragmatic properties. Any alternative credal reallocation policy leaves you vulnerable to a guaranteed net loss in Kolmogorov certainty acquisition scenarios. If you are a Kolmogorov conditionalizer, then you are immunized from that unsavory prospect. Kolmogorov conditionalization is the unique credal reallocation strategy that guards against Dutch bookability in Kolmogorov certainty acquisition scenarios.

The theorems also clarify the status of impropriety. Upon learning that rcds are not always proper, one naturally asks whether some alternative theory of conditional probability might prove superior to Kolmogorov's. Section 4's theorems show that, whatever benefits an alternative theory might offer, it would carry serious costs. Any proper update rule besides an rcd will run afoul of Section 4's Dutch book theorem. If the specter of impropriety impels you to replace Kolmogorov conditionalization with some alternative update strategy, then you face a sure loss *no matter the true outcome and no matter your new certainty profile*.

We must distinguish between Dutch book *theorems* and Dutch book *arguments*. Dutch book *theorems* are uncontroversial mathematical results. Dutch book *arguments* use Dutch book theorems to argue that one should conform to certain credal norms. For example, Lewis [26] and Skyrms [43] cite diachronic Dutch book theorems to argue that epistemic rationality favors conditionalization over rival credal reallocation policies. Critics have raised various objections to Dutch book arguments [17], especially diachronic Dutch book arguments [5]. In this paper, I have not mounted a Dutch book argument. I have *not* argued that epistemic rationality requires you to be a Kolmogorov conditionalizer, or anything along those lines.

To my mind, Section 4's theorems are important primarily for elucidatory and diagnostic purposes, not for their potential to sustain a Dutch book argument. The theorems help us understand the benefits that accrue when your update rule abides by the integral formula. Someone who flouts the integral formula leaves herself vulnerable to a guaranteed net loss in a way that someone who conforms to it does not. Whether or not the theorems show that Kolmogorov conditionalization is an epistemically rational credal reallocation policy, they show that it has uniquely attractive pragmatic virtues. In that respect, the theorems depict Kolmogorov conditionalization as a natural extension of ratio conditionalization to null updating scenarios.

Admittedly, the extension does not go quite as smoothly as one might hope. We have encountered several concerning features of rcds:

(a) Rcds may not exist, in which case Kolmogorov conditionalization is impossible.
(b) Outside the special case where the ratio formula prevails, there are infinitely many rcds for P given $\mathcal{G}$.
(c) Rcds may be improper, in which case Kolmogorov conditionalization is impossible in Kolmogorov certainty acquisition scenarios.

(a) and (c) show that Kolmogorov conditionalization is not always possible. (b) shows that there are sometimes infinitely many distinct ways to be a Kolmogorov conditionalizer; in contrast, ratio conditionalization dictates a single unique credal update.

Despite challenges regarding existence, uniqueness, and impropriety, Kolmogorov conditionalization is a highly general credal reallocation strategy that applies across

https://doi.org/10.1017/S1755020323000345 Published online by Cambridge University Press

a huge range of null updating scenarios. Most likely, it applies to all null updating scenarios that might arise in scientific and philosophical applications of Bayesian decision theory. Section 4's theorems establish that it offers distinctive benefits. All the more fitting, then, that it figures so centrally in contemporary scientific applications of Bayesian decision theory. The time is long past for philosophers to join the broader scientific community by assigning rcds a central role within the foundations of Bayesian inference.

Acknowledgments I thank Alexander Meehan, Stephen Ge, Snow Zhang, and two anonymous referees for this journal for discussion and feedback that helped me improve the paper.

BIBLIOGRAPHY

[1] Bennett, B., Hoffman, D., & Prakash, C. (1996). Observer theory, Bayes theory, and psychophysics. In Knill, D. & Richards, W., editors. *Perception as Bayesian Inference*. Cambridge: Cambridge University Press, pp. 229–235.

[2] Billingsley, P. (1995). *Probability and Measure* (third edition). New York: Wiley.

[3] Blackwell, D., & Dubins, L. (1975). On existence and non-existence of proper, regular conditional distributions. *The Annals of Probability*, **3**, 741–752.

[4] Blackwell, D., & Ryll-Nardzewski, C. (1963). Non-existence of everywhere proper conditional distributions. *The Annals of Mathematical Statistics*, **34**, 223–225.

[5] Christensen, D. (1991). Clever bookies and coherent beliefs. *Philosophical Review*, **100**, 229–247.

[6] de Finetti, B. (1937/1980). Foresight: Its logical laws, its subjective sources. In Kyburg Jr., H. E. & Smokler, H. E., editors. *Rpt. in Studies in Subjective Probability*. Huntington: Robert E. Krieger, pp. 53–118.

[7] ———. (1972). *Probability, Induction, and Statistics*. New York: Wiley.

[8] Easwaran, K. (2008). *The Foundations of Conditional Probability*. Ph.D. Thesis, University of California, Berkeley; Ann Arbor: ProQuest/UMI (Publication No. 3331592).

[9] ———. (2011). Varieties of conditional probability. In Bandyopadhyay, P. & Forster, M., editors. *Philosophy of Statistics*. Burlington: Elsevier, pp. 137–148.

[10] ———. (2019). Conditional probabilities. In Pettigrew, R. & Weisberg, J., editors. *The Open Handbook of Formal Epistemology*. PhilPapers, pp. 131–198.

[11] Florens, J.-P., Mouchart, M., & Rolin, J.-M. (1990). *Elements of Bayesian Statistics*. New York: Marcel Dekker.

[12] Fristedt, B., & Gray, L. (1997). *A Modern Approach to Probability Theory*. Boston: Birkhäuser.

[13] Gaifman, H. (1988). A theory of higher-order probabilities. In Skyrms, B. & Harper, W., editors. *Causation, Chance, and Credence: Proceedings of the Irvine Conference on Probability and Causation*. Boston: Kluwer.

[14] Ghosal, S., & van der Vaart, A. (2017). *Fundamentals of Nonparametric Bayesian Inference*. Cambridge: Cambridge University Press.

[15] Gyenis, Z., Hofer-Szabó, G., & Rédei, M. (2017). Conditioning using conditional expectations: The Borel–Kolmogorov paradox. *Synthese*, **194**, 2595–2630.

https://doi.org/10.1017/S1755020323000345 Published online by Cambridge University Press

[16] Hájek, A. (2003) What conditional probability could not be. *Synthese*, **137**, 273–323.

[17] ———. (2008). Dutch book arguments. In Anand, P., Pattanaik, P., & Puppe, C., editors, *The Handbook of Rationality and Social Choice*. Oxford: Oxford University Press, pp. 173–195.

[18] ———. (2011). Conditional probability. In Bandyopadhyay, P. & Forster, M., editors. *Philosophy of Statistics*. Burlington: Elsevier.

[19] Howson, C. (2014). Finite additivity, another lottery paradox, and conditionalization. *Synthese*, **191**, 989–1012.

[20] Huttegger, S. (2015). Merging of opinions and probability kinematics. *The Review of Symbolic Logic*, **8**, 611–648.

[21] ———. (2017). *The Probabilistic Foundations of Rational Learning*. Cambridge: Cambridge University Press.

[22] Huttegger, S., & Nielsen, M. (2020). Generalized learning and conditional expectation. *Philosophy of Science*, **87**, 868–883.

[23] Kiefer, N., & Nyarko, Y. (1995). Savage–Bayesian models of economics. In Kirman, A. & Salmon, M., editors. *Learning and Rationality in Economics*. Oxford: Blackwell, pp. 40–62.

[24] Kolmogorov, A. N. (1933/1956). *Foundations of the Theory of Probability* (second English edition). N. Morrison, translator. New York: Chelsea.

[25] Lee, J. J. (2018). Formalization of information: Knowledge and belief. *Economic Theory*, **66**, 1007–1022.

[26] Lewis, D. (1999). Why conditionalize?. In *Papers in Metaphysics and Epistemology*. Cambridge: Cambridge University Press, pp. 403–407.

[27] Meehan, A., & Zhang, S. (2020). Jeffrey meets Kolmogorov: A general theory of conditioning. *Journal of Philosophical Logic*, **49**, 941–979.

[28] ———. (2022). Kolmogorov conditionalizers can be Dutch booked (if and only if they are evidentially uncertain). *The Review of Symbolic Logic*, **15**, 722–757.

[29] Mertens, J.-F., & Zamir, S. (1985). Formulation of Bayesian analysis for games of incomplete information. *International Journal of Game Theory*, **14**, 1–29.

[30] Myrvold, W. (2015). You can't always get what you want: Some considerations regarding conditional probabilities. *Erkenntnis*, **80**, 573–603.

[31] Nielsen, M. (2021). A new argument for Kolmogorov conditionalization. *The Review of Symbolic Logic*, **14**, 930–945.

[32] Ramachandran, D. (1979). Existence of independent complements in regular conditional probability spaces. *Annals of Probability*, **7**, 433–443.

[33] Ramsey, F. P. (1931). Truth and probability. In Braithwaite, R. B., editor. *The Foundations of Mathematics and Other Logical Essays*. London: Routledge and Kegan, pp. 156–199.

[34] Rao, M. M. (2005). *Conditional Measures and Applications* (second edition). Boca Raton: CRC Press.

[35] Rescorla, M. (2015). Some epistemological ramifications of the Borel–Kolmogorov paradox. *Synthese*, **192**, 735–767.

[36] ———. (2018). A Dutch book theorem and converse Dutch book theorem for Kolmogorov Conditionalization. *The Review of Symbolic Logic*, **11**, 705–735.

[37] ———. (2021). On the proper formulation of Conditionalization. *Synthese*, **198**, 1935–1965.

https://doi.org/10.1017/S1755020323000345 Published online by Cambridge University Press

[38] ———. (2022). An improved Dutch book theorem for Conditionalization. *Erkenntnis*, **87**, 1013–1041.

[39] ———. (2023). Reflecting on diachronic Dutch books. *Noûs*, **57**, 511–538.

[40] ———. (2024). Bayesian defeat of certainties. *Synthese*, **203**, 50.

[41] Seidenfeld, T. (2001). Remarks on the theory of conditional probability: Some issues of finite versus countable additivity. In Hendricks, V., Pedersen, S., & Jørgensen, K., editors. *Probability Theory: Philosophy, Recent History, and Relations to Science*. Dordrecht: Kluwer, pp. 167–178.

[42] Seidenfeld, T., Schervish, M., & Kadane, J. (2001). Improper regular conditional distributions. *Annals of Probability*, **29**, 1612–1624.

[43] Skyrms, B. (1987). Dynamic coherence and probability kinematics. *Philosophy of Science*, **54**, 1–20.

[44] Sokal, A. (1981). Existence of compatible families of proper regular conditional probabilities. *Zeitschrift für Wahrscheinliehkeitstheorie und verwandte Gebiete*, **56**, 537–548.

[45] Teller, P. (1973). Conditionalization and observation. *Synthese*, **26**, 218–258.

[46] Thrun, S., Burgard, W., & Fox, D. (2005). *Probabilistic Robotics*. Cambridge: MIT Press.

[47] Trotta, R. (2008). Bayes in the sky: Bayesian inference and model selection in cosmology. *Contemporary Physics*, **49**, 71–104.

DEPARTMENT OF PHILOSOPHY
UNIVERSITY OF CALIFORNIA, LOS ANGELES
390 PORTOLA PLAZA
LOS ANGELES, CA 90095, USA
E-mail: rescorla@ucla.edu

https://doi.org/10.1017/S1755020323000345 Published online by Cambridge University Press