A Dutch Book Theorem and Converse Dutch Book Theorem for

Kolmogorov Conditionalization

Michael Rescorla

Abstract: This paper discusses how to update one's credences based on evidence that has initial probability 0. I advance a diachronic norm, *Kolmogorov Conditionalization*, that governs credal reallocation in many such learning scenarios. The norm is based upon Kolmogorov's theory of conditional probability. I prove a Dutch book theorem and converse Dutch book theorem for Kolmogorov Conditionalization. The two theorems establish Kolmogorov Conditionalization as the unique credal reallocation rule that avoids a sure loss in the relevant learning scenarios.

§1. Dutch book arguments for conditionalization

How should you update your credences in light of new evidence? The most widely discussed norm is **Conditionalization**, which requires that:

> If you assign credence $P(H)$ to a proposition H, and you gain new evidence that is exhausted by knowledge of E, then you respond to your new evidence by assigning credence $P(H \mid E)$ to H.

Here $P(H \mid E)$ is the *conditional probability* of H given E. Conditionalization traces back to Bayes's seminal discussion (Bayes and Price, 1763). It is a linchpin of Bayesian decision theory.

Philosophers have pursued various strategies for justifying Conditionalization. One prominent strategy builds upon the classic *Dutch book arguments* advanced by Ramsey (1931)

and de Finetti (1937/1980). A *Dutch book* is a collection of acceptable bets that inflict a sure loss. An agent is *Dutch bookable* when it is possible to rig a Dutch book against her. Dutch bookability is a very undesirable property, because a sufficiently devious bookie can pump a Dutch bookable agent for money by offering her bets that she gladly accepts. Ramsey and de Finetti independently noted that one can rig a Dutch book against an agent whose credences violate the probability calculus axioms. They argued that such an agent's credal allocation is rationally defective. They concluded that credences should conform to the probability calculus axioms. This is a *synchronic* Dutch book argument, because it addresses credences at a moment rather than credal evolution over time. Subsequent authors have pursued *diachronic* Dutch book arguments concerning how to reallocate credence in light of new evidence. In particular, Lewis proved a Dutch book theorem for Conditionalization. Teller (1973) publicized the theorem (with full credit to Lewis), and Lewis (1999) eventually published his own treatment.

Lewis's theorem concerns an idealized agent who learns a proposition drawn from some mutually exclusive, jointly exhaustive set of propositions $E_1, \ldots, E_n$, with $P(E_i) > 0$ for each i. When E_i has non-zero probability, one defines conditional probability through the *ratio formula*:

$$P(H \mid E_i) = \frac{P(H \& E_i)}{P(E_i)} \ .$$

For the proposed learning scenario, we may reformulate Conditionalization as a norm that I will call **Ratio Conditionalization**. This norm requires that:

If you assign credence $P(H)$ to each proposition H, and $E_1, \ldots, E_n$ are mutually exclusive, jointly exhaustive propositions with $P(E_i) > 0$ for all i, and you gain new evidence that is exhausted by knowledge of one particular proposition E_i, then you respond to your new evidence by assigning credence $\frac{P(H \& E_i)}{P(E_i)}$ to H.

Lewis proved that one can rig a *diachronic Dutch book* (a Dutch book involving bets offered at different times) against an agent who violates Ratio Conditionalization. Skyrms (1987) proved a converse theorem: one cannot rig a diachronic Dutch book against an agent who conforms to the probability calculus axioms and who obeys Ratio Conditionalization. These two theorems establish Ratio Conditionalization as the unique update rule that avoids diachronic Dutch books in the specified learning scenario. Lewis and Skyrms argue on that basis that one should conform to Ratio Conditionalization.

When $P(E_i) = 0$, the ratio formula is not well-defined and Ratio Conditionalization does not apply. Ratio Conditionalization does not say how an agent should update her credences based upon evidence that has probability 0. I call such cases *null updating scenarios*. A simple example is conditioning on the value of a *continuous random variable X*, i.e. a random variable with continuum many values. Orthodox probability theory demands that $P(X = x) = 0$ for all but countably many values x. When $P(X = x) = 0$, the ratio formula does not yield well-defined conditional probabilities $P(H \mid X = x)$. Thus, Ratio Conditionalization does not specify how to reallocate credence if you learn that $X = x$. How should you reallocate credence under these circumstances? More generally, what diachronic norms govern null updating?

Kolmogorov (1933/1956) offered an influential mathematical framework that bears directly upon these questions. He advanced a theory of conditional probability general enough to cover scenarios where the conditioning evidence has probability 0. Kolmogorov's treatment plays a foundational role within probability theory. It informs every standard graduate-level textbook. A few philosophers have explored its potential to illuminate credal reallocation (e.g. Easwaran, 2008; Gyenis and Rédei, 2017; Huttegger, 2015). More commonly, though,

philosophers either reject Kolmogorov's approach (e.g. Hájek, 2003, 2011; Howson, 2014; Myrvold, 2015; Seidenfeld, 2001) or else ignore it altogether.

I think that the philosophical community has paid Kolmogorov's insights into conditional probability far less attention than they merit. One can extract from Kolmogorov's discussion a plausible diachronic norm that governs many important null updating scenarios. I call this norm *Kolmogorov Conditionalization*. I will prove a Dutch book theorem and converse Dutch book theorem for Kolmogorov Conditionalization. The theorems establish Kolmogorov Conditionalization as the unique update rule that avoids diachronic Dutch books in relevant learning scenarios.[1]

Considerable stage-setting is needed before I state and prove the two main theorems. §§2-3 offer background remarks on null updating and Dutch book arguments. §4 presents Kolmogorov's theory of conditional probability. §5 introduces Kolmogorov Conditionalization and constructs a diachronic Dutch book for agents who violate it. §6 more rigorously states and proves a strengthened Dutch book theorem. §7 proves a converse Dutch book theorem. §8 discusses the significance of the two theorems. I assume throughout that credences should be countably additive. This assumption is controversial (de Finetti, 1972), (Howson, 2014), (Kadane, Schervish, and Seidenfeld, 1986), (Savage, 1954). However, exploring the two main theorems will keep us busy enough without wading into controversies over countable additivity.

§2. Null updating

---

[1] McGee (1994) generalizes Lewis's setup so as to analyze a special case of null updating. He considers an agent who learns a proposition drawn from some mutually exclusive, jointly exhaustive set $E_1, \ldots, E_n$, where $P(E_i)$ may be 0. For this special case, McGee proves a Dutch book theorem and converse Dutch book theorem involving Popper's (1959) theory of conditional probability. McGee's setup is not general enough to cover even our motivating example of conditioning on a continuous random variable, since a continuous random variable induces an infinite partition of the probability space. The theorems I prove below cover conditioning on a continuous random variable and many other learning scenarios unaddressed by previous Dutch book results.

Why study null updating? One reason is that it figures prominently in scientific practice.

Our world is populated by continuously varying quantities: shapes, sizes, colors, locations, and so on. Bayesian agents must frequently estimate these quantities. Accordingly, continuous random variables are fixtures within scientific applications of the Bayesian paradigm, including *Bayesian statistics* (Lindley, 1982), *probabilistic robotics* (Thrun, Burgard, and Fox, 2006), and *game theory* (Fudenberg and Tirole, 1991). They also play a large role within *Bayesian cognitive science* (Chater and Oaksford, 2008), which provides Bayesian models of core mental processes such as *perception* (Knill and Richards, 1996), *motor control* (Wolpert, 2007), and *navigation* (Madl, et al. 2014). All these fields routinely adduce credences over a continuous hypothesis space. For example, probabilistic robotics aims to construct a robot that estimates its own location. Similarly, Bayesian perceptual psychology hypothesizes that the perceptual system uses Bayesian inference to estimate the shapes, sizes, colors, and locations of observable objects. Scientific applications of Bayesianism cannot get far without a continuous probability space.

Standard probability theory requires that all but countably many values of a continuous random variable receive probability 0. Scientific applications of Bayesianism therefore assign a central role to null updating. For example, one might estimate the mass of some star by measuring perturbations in the orbit of a nearby planet; or one might estimate the location of an underwater missile by taking sonar readings; or one might estimate the color of a distal surface by measuring the light spectrum emitted by the surface. All these learning scenarios, and numerous others that arise in scientific applications, require conditioning on the value of a continuous random variable (or random vector).

Nevertheless, null updating receives surprisingly little philosophical discussion. One reason is that many philosophers endorse *regularity*: the doctrine that one should assign positive probability to metaphysically possible propositions (Jeffrey, 1992), (Kemeny, 1955), (Shimony, 1955), (Skyrms, 1980), (Stalnaker, 1970). Assuming regularity, a rational agent cannot learn a proposition to which she formerly assigned probability 0. However, regularity is implausible (Hájek, 2003, 2012). It conflicts with the aforementioned scientific applications of Bayesianism. It also conflicts with Ratio Conditionalization: conditionalizing on E_i leads you to assign credence 1 to E_i and credence 0 to conflicting propositions, even when those propositions are metaphysically possible. Anyone sympathetic to Ratio Conditionalization should reject regularity.

Jeffrey (1983) argues that Conditionalization has limited applicability to real world agents, since experience rarely authorizes credence 1 for an empirical proposition E_i. More typically, experience authorizes you to reallocate credence across certain select propositions E_1, …, E_n, with no single proposition receiving credence 1. On that basis, you must assign new credences to all remaining propositions H. Jeffrey formulates an update rule (*Jeffrey Conditionalization*) tailored to this learning scenario.

Although Jeffrey focuses his critique on Ratio Conditionalization, a similar worry arises even more forcefully for null updating. We are finite beings with limited representational and discriminative capacities. Our perceptual organs and measuring instruments rarely if ever specify the value of a random variable X with infinite precision (Borel, 1909/1956), (Myrvold, 2015). At best, we learn that X has value *x plus or minus some margin of error*. At best, we learn that X's value falls in some interval, where the interval has positive probability. Thus, one might argue that we rarely if ever learn null propositions with anything approaching complete confidence.

From this viewpoint, null updating looks like an overly fanciful learning scenario that idealizes away crucial features of real-world statistical inference.

To a certain extent, I sympathize with such worries. Null updating scenarios *are* heavily idealized. There is value in studying less idealized learning scenarios. But I insist that there is *also* value in studying null updating. In field after field --- from statistics to economics to robotics to cognitive science --- researchers have found it explanatorily or pragmatically fruitful to consider null updating scenarios. These scenarios form an essential starting point for inquiry into less idealized scenarios. Normative models of null updating persist as benchmarks against which we can compare less idealized models. Thus, current scientific practice establishes null updating as a theoretically important phenomenon that merits intensive foundational investigation. For further argument that we should study null updating, see (McGee, 1994).

§3. Dutch book *theorems* versus Dutch book *arguments*

Even philosophers who agree that null updating merits investigation may doubt that Dutch books shed much light upon it. Although Dutch book arguments are still fairly popular, they have attracted increasingly harsh criticism over the past few decades.

We must carefully distinguish here between Dutch book *theorems* and Dutch book *arguments*. Dutch book *theorems* are mathematical results that admit definitive proof. For example, Lewis proves that one can rig a diachronic Dutch book against anyone who violates Ratio Conditionalization. Dutch book *arguments* use Dutch book theorems to defend a philosophical conclusion: namely, that one should conform to certain norms governing credal allocation. Over the past few decades, philosophers have grown increasingly skeptical that one can convert Dutch book theorems into compelling Dutch book arguments (Hájek, 2009). Most

fundamentally, Dutch book arguments seem to elide *pragmatic* and *epistemic* factors. If you are Dutch bookable, then you are vulnerable to a guaranteed net loss. This vulnerability is a pragmatic defect, so it indicates that your credences are defective *from a pragmatic viewpoint*. But why conclude that your credences are defective *from an epistemic viewpoint*? How can Dutch book considerations establish a failure of *epistemic* rationality?[2]

The present paper focuses on Dutch book *theorems* rather than Dutch book *arguments*. I will prove a Dutch book theorem and converse Dutch book theorem for Kolmogorov Conditionalization. I will *not* use the theorems to argue that Kolmogorov Conditionalization is epistemically privileged. Even without an accompanying Dutch book argument, the theorems are useful and informative. Specifically, they show that Kolmogorov's theory of conditional probability captures fundamental links between conditional probability, credal reallocation, and decision-making.

§4. Kolmogorov's theory of conditional probability

I now present the basic elements of Kolmogorov's theory. Billingsley's (1995, pp. 427-440) more detailed exposition serves as a partial basis for my own exposition. In §4.1, I introduce the learning scenarios modeled by Kolmogorov. In §4.2, I describe how Kolmogorov delineates conditional probabilities tailored to these learning scenarios.

§4.1 Kolmogorov learning scenarios

---

[2] Another prominent worry focuses more specifically on diachronic Dutch book arguments. van Fraassen (1984) proves a diachronic Dutch book theorem for the Principle of Reflection, which is widely regarded as implausible. Some authors conclude that diachronic Dutch book arguments are suspect (Christensen, 1991). van Fraassen himself concludes that we should accept both Conditionalization and Reflection. Others try to isolate a principled difference between Lewis's diachronic Dutch book theorem for Conditionalization and van Fraassen's diachronic Dutch book theorem for Reflection, so that the former may yield a compelling Dutch book argument even though the latter does not (Briggs, 2009), (Mahtani, 2015).

Kolmogorov assumes an idealized agent whose credences are modeled by 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 some purposes, we might regard outcomes as *possible worlds* and events as *propositions*. However, Kolmogorov's theory does not presuppose this philosophical gloss.

Kolmogorov considers a learning setup much more general than that considered by Lewis. In both setups, an idealized agent awaits partial information about the true outcome $\omega \in \Omega$. In Lewis's setup, the agent learns where ω falls within some finite partition of Ω. In Kolmogorov's setup, the agent learns whether ω belongs to each $G \in \mathcal{G}$, where $\mathcal{G} \subseteq \mathcal{F}$ is itself σ-field. Intuitively, the sub-σ-field $\mathcal{G}$ serves as an "information filter." The agent does not learn everything about outcome ω, but she learns about ω as filtered through $\mathcal{G}$. I call learning scenarios of this kind *Kolmogorov learning scenarios*. Thus, a Kolmogorov learning scenario is one where the agent gains full membership knowledge for a sub-σ-field $\mathcal{G} \subseteq \mathcal{F}$ regarding the true outcome ω. If we view events as propositions, then we can say that $\mathcal{G}$ contains all the new propositions learned by our agent.

To illustrate Kolmogorov learning scenarios, consider a continuous random variable $X: \Omega \to \mathbb{R}$. When you learn that $X(\omega) = r$, you thereby learn many additional facts about the events to which ω belongs. Assuming that expression "r" is suitably informative, knowledge that $X(\omega) = r$ allows you to affirm or deny each proposition

$$X(\omega) \in (a, b) \qquad\qquad a, b \in \mathbb{Q}.$$

In that sense, you gain membership knowledge for all sets

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

Let σ(X) be the σ-field generated by these sets, i.e. the result of starting with the sets $X^{-1}(a, b)$ and closing under complementation and countable unions. If your knowledge were closed under complementation and countable union, then you could extrapolate from membership knowledge for the sets $X^{-1}(a, b)$ to full membership knowledge for σ(X). Thus, the sub-σ-field σ(X) helps us model the knowledge of an idealized superhuman who learns that $X = r$ and whose knowledge is closed under complementation and countable union.

Of course, the explicit knowledge of an ordinary finite human is not usually closed under complementation and countable union. However, there is a natural sense in which full membership knowledge for σ(X) is *implicit* in an ability to affirm or deny each proposition

$$X(\omega) \in (a, b) \qquad\qquad a, b \in \mathbb{Q}.$$

In that sense, complete membership information for σ(X) models the *implicit knowledge* gained by an agent who learns that $X = r$.

Even if we use σ(X) to model *implicit* rather than *explicit* knowledge, the envisaged learning scenario assumes superhuman mental capacities. If σ(X) models possible implicit knowledge that an agent might acquire, then the agent has uncountably many possible doxastic states corresponding to each value of X. Finite beings such as ourselves do not have uncountably many possible doxastic states. In particular, we do not have the capacity to represent arbitrary real numbers with infinite precision. Thus, Kolmogorov's model *even viewed as a model of implicit knowledge* assumes an idealized superhuman with infinitary mental capacities. As indicated in §2, idealized models of this kind play an important role in current science. They serve as benchmarks against which we can compare less idealized models. Scientific applications of Bayesian modeling have repeatedly demonstrated the explanatory and pragmatic benefits that accrue when we take these idealized benchmarks as a starting point.

Many important cases of rational credal reallocation can be fruitfully modeled (in an idealized way) as Kolmogorov learning scenarios. I do not say that *all* cases of rational credal reallocation can be so modeled. Some learning scenarios do not naturally fit the Kolmogorov template: scenarios of the kind highlighted by Jeffrey, where experience authorizes you to reallocate credence across certain select propositions $E_1, \ldots, E_n$, with no single proposition receiving credence 1; scenarios involving memory loss or the threat of memory loss (Arntzenius, 2003); scenarios involving conceptual discoveries that add new propositions to your cognitive repertoire (Lewis, 1999); and so on. Thus, Kolmogorov learning scenarios are only one type of learning scenario one might study. Still, they are very important. We also understand them relatively well, thanks in large part to Kolmogorov's efforts. I henceforth focus exclusively on Kolmogorov learning scenarios. (Cf. Easwaran, 2013, pp. 122-123.)

§4.2 Regular conditional distributions

We want to delineate probabilities conditional on information about whether ω belongs to each $G \in \mathcal{G}$. Intuitively, these conditional probabilities constitute a plan for updating credences after gaining full membership knowledge for $\mathcal{G}$ regarding ω. Formally, Kolmogorov isolates a function $P_{\mathcal{G}} : \mathcal{F} \times \Omega \to \mathbb{R}$. We write $P_{\mathcal{G}} (A \mid \omega)$ to denote the value that this function assumes on inputs $A \in \mathcal{F}$ and $\omega \in \Omega$. Think of $P_{\mathcal{G}} (A \mid \omega)$ as the probability of A given ω *as filtered through* $\mathcal{G}$. Kolmogorov places three constraints on $P_{\mathcal{G}}$:

Regularity: For each $\omega \in \Omega$, $P_{\mathcal{G}}$ induces a one-place function $P_{\mathcal{G}} (. \mid \omega) : \mathcal{F} \to \mathbb{R}$. Say that $P_{\mathcal{G}}$ is *regular* just in case:

$P_{\mathcal{G}} (. \mid \omega)$ is a probability measure　　　　　　　*for each $\omega \in \Omega$.*

Intuitively: conditioning on new evidence should always carry you to a probability measure. Note that this is a totally different notion of *regularity* than the notion rejected in §2. It is an unfortunate fact that the literature associates these two completely different notions with the same word "regularity."

$\mathcal{G}$-measurability: For each $A \in \mathcal{F}$, $P_\mathcal{G}$ induces a one-place function $P_\mathcal{G}(A \mid .): \Omega \to \mathbb{R}$. This function reflects a policy for updating the credence assigned to A upon receiving partial information about the true outcome. Kolmogorov requires that

$$P_\mathcal{G}(A \mid .) \text{ is } \mathcal{G}\text{-measurable} \qquad \qquad \text{for each } A \in \mathcal{F}.$$

In other words, he requires that

$$P_\mathcal{G}(A \mid .)^{-1}(-\infty, a] \in \mathcal{G} \qquad \qquad \text{for every } A \in \mathcal{F}, a \in \mathbb{R}.$$

$\mathcal{G}$-measurability reflects the assumption that credences are updated based *solely* upon membership information for $\mathcal{G}$. Intuitively: the agent must learn whether the updated credence for A is $\leq a$, so one of the propositions she implicitly learns should be either the proposition that

$$P_\mathcal{G}(A \mid \omega) \leq a$$

or else the proposition that

$$P_\mathcal{G}(A \mid \omega) > a.$$

If $P_\mathcal{G}(A \mid .)^{-1}(-\infty, a] \in \mathcal{G}$, then someone who learns whether $\omega \in G$ for each $G \in \mathcal{G}$ can (at least in principle) determine whether $P_\mathcal{G}(A \mid \omega) \leq a$. She need simply examine whether $\omega \in P_\mathcal{G}(A \mid .)^{-1}(-\infty, a]$. In contrast, suppose that $P_\mathcal{G}(A \mid .)^{-1}(-\infty, a] \notin \mathcal{G}$. Then our agent's newly acquired membership knowledge about $\mathcal{G}$ does not include knowledge whether $\omega \in P_\mathcal{G}(A \mid .)^{-1}(-\infty, a]$. She does not acquire even implicit knowledge whether $P_\mathcal{G}(A \mid \omega) \leq a$. Thus, $\mathcal{G}$-measurability captures

the intuitive idea that someone who gains full membership knowledge for $\mathcal{G}$ regarding ω thereby gains implicit knowledge whether $P_{\mathcal{G}}(A \mid \omega) \leq a$.

The integral formula: Conditional probabilities must be appropriately related to unconditional probabilities. When the conditioning event has non-zero probability, the ratio formula specifies this "appropriate relation." To analyze probabilities conditional on events with probability 0, Kolmogorov offers a constraint called the *integral formula*:

$$P(A \cap G) = \int_G P_{\mathcal{G}}(A \mid \omega) dP(\omega), \qquad\qquad \textit{for any } G \in \mathcal{G}.$$

The integral formula generalizes the law of total probability from elementary probability theory.

A $\mathcal{G}$-measurable function $P_{\mathcal{G}}(A \mid .): \Omega \to \mathbb{R}$ that satisfies the integral formula is a *conditional probability for A given $\mathcal{G}$*. A two-place function $P_{\mathcal{G}}(. \mid .)$ satisfying all three constraints is a *regular conditional distribution (rcd) for P given $\mathcal{G}$*. One can prove that there always exists a conditional probability for *A* given $\mathcal{G}$ (Billingsley, 1995, p. 430). One can also prove that there exists an rcd for *P* given $\mathcal{G}$ in a wide variety of cases (Durrett, 1991, pp. 198-200), (Rao, 2005, pp. 125-182), including all or virtually all cases that arise in empirical applications.

In certain pathological cases, there is no rcd for *P* given $\mathcal{G}$ (Billingsley, 1995, p. 443). Some critics regard these cases as a serious problem for Kolmogorov's theory (Seidenfeld, 2001). However, I think that they should not worry any theorist who already accepts countable additivity. Vitali proved that certain σ-fields, such as the power set of the unit interval, do not admit probability measures. His proof uses the Axiom of Choice to define a nonmeasurable set. Standard examples where rcds do not exist likewise feature a nonmeasurable set (Seidenfeld, Schervish, and Kadane, 2001). Once we accept that countably additive *unconditional*

probabilities do not always exist, we should not feel particularly disturbed that countably

additive *conditional* probabilities do not always exist.

Notably, Kolmogorov *relativizes* conditional probabilities to conditioning sub-σ-fields.

One never conditions upon an isolated probability zero event. Rather, one conditions upon an

outcome *as filtered through a sub-σ-field* $\mathcal{G}$. Suppose random variables *X* and *Y* are such that

$X(\omega) = x$ iff $Y(\omega) = y$ *for all* $\omega \in \Omega$,

where the event $\{X = x\}$ (i.e. the event $\{Y = y\}$) has probability 0. In many such cases, there

exists $A \in \mathcal{F}$ such that

$P_{\sigma(X)}(A \mid \omega) \neq P_{\sigma(Y)}(A \mid \omega)$

for all $\omega \in \{X = x\}$, where $P_{\sigma(X)}(A \mid .)$ is a conditional probability for *A* given $\sigma(X)$ and $P_{\sigma(Y)}(A \mid .)$

is a conditional probability for *A* given $\sigma(Y)$. According to Kolmogorov, there is no determinate

answer as to the conditional probability of *A* given $\{X = x\}$. A determinate conditional

probability arises only once one regards $\{X = x\}$ as embedded in a surrounding sub-σ-field.[3]

Kolmogorov's relativistic approach is controversial among philosophers. Many authors insist

that we should instead strive to isolate unrelativized conditional probabilities (Hill, 1980),

(Howson, 2014), (Kadane, Schervish, and Seidenfeld, 1986), (Myrvold, 2015). See (Gyenis,

Hofer-Szabó, and Rédei, 2017) and (Rescorla, 2015) for defense of Kolmogorov's relativistic

approach.

---

[3] A famous example, nowadays called *the Borel-Kolmogorov paradox*, arises when a point is picked randomly on the earth's surface. How should we condition on news that the point falls on some great circle *C*? One can motivate different intuitively compelling answers, depending on whether one regards *C* as the equator or as two meridians fused together. According to Kolmogorov, there is no single determinate answer. We must first settle upon a conditioning sub-σ-field that contains *C*. We can pick a sub-σ-field that corresponds to regarding *C* as the equator, or we can pick a different sub-σ-field that corresponds to regarding *C* as two meridians fused together. The different conditioning sub-σ-fields engender different rcds. Gyenis, Hofer-Szabó, and Rédei (2017) and Rescorla (2015) discuss the Borel-Kolmogorov paradox at length.

Even when we hold the conditioning sub-σ-field fixed, Kolmogorov's theory determines $P_\mathfrak{s}(A \mid .): \Omega \to \mathbb{R}$ only up to measure 0. One can vary $P_\mathfrak{s}(A \mid \omega)$ arbitrarily on outcomes ω comprising a probability zero event, provided one takes care that for each ω the results assemble into a probability measure $P_\mathfrak{s}(. \mid \omega)$. This indeterminacy arises because the integral formula mentions $P_\mathfrak{s}(A \mid \omega)$ only inside an integral sign. Integration ignores differences of measure 0, so the integral formula does not pin down unique conditional probabilities. In the mathematical literature, candidate conditional probabilities $P_\mathfrak{s}(A \mid .)$ equal up to measure 0 are sometimes called *versions* of the conditional probability, which is then regarded as an equivalence class of its versions. Using this terminology, Kolmogorov's theory does not privilege any specific version of the conditional probability. Many authors find the resulting indeterminacy worrisome. Pfanzagl (1979), Rao (2005), and Tjur (1980) try to mitigate the indeterminacy by supplementing Kolmogorov's theory with additional constraints on conditional probabilities.[4]

Over the past century, researchers such as Popper (1959) and Rényi (1955) have proposed various alternatives to Kolmogorov's theory of conditional probability. These alternative theories tend to receive better philosophical press than Kolmogorov's theory. However, Kolmogorov's approach is orthodox within contemporary mathematical practice. One of its main advantages is that the integral formula tightly constrains the relation between conditional and unconditional probabilities. Popper and Rényi supply no comparably substantive constraints. In practice, Kolmogorov offers far more guidance for computing conditional probabilities than either Popper and Rényi. Of course, one must assess whether Kolmogorov offers *good* guidance. The theorems proved below provide insight into that question.

---

[4] Another worry about Kolmogorov's approach is that, when the conditioning sub-σ-field is sufficiently odd, the resulting rcds have properties that conditional properties seemingly should not have (Hájek, 2009), (Seidenfeld, Schervish, and Kadane, 2001). See Easwaran (2011) for one response to this worry.

§5. A Dutch book theorem for Kolmogorov Conditionalization

Kolmogorov's explicit concern is *conditional probability* rather than *credal reallocation*. Still, Kolmogorov's theory naturally suggests a norm that governs credal reallocation in Kolmogorov learning scenarios. Suppose that $(\Omega, \mathcal{F}, P)$ is a probability space and $\mathcal{G} \subseteq \mathcal{F}$ is a sub-σ-field. ***Kolmogorov Conditionalization*** requires that:

> If you begin with credences P over $\mathcal{F}$, and there exists an rcd for P given $\mathcal{G}$, then, for one such rcd $P_\mathcal{G}$, whenever you gain new evidence that is exhausted by full membership knowledge for $\mathcal{G}$ regarding the true outcome ω, you respond by adopting new credences $P_\mathcal{G}(. \mid \omega)$ over $\mathcal{F}$.

Kolmogorov Conditionalization requires you to update credences whenever possible using an rcd. If $\mathcal{G}$ is generated by a partition $E_1, \ldots, E_n$ of events such that $P(E_i) > 0$, one can show that the ratio formula yields a unique rcd for P given $\mathcal{G}$. In this special case, Kolmogorov Conditionalization reduces to Ratio Conditionalization. If there are infinitely many rcds for P given $\mathcal{G}$, you can satisfy Kolmogorov Conditionalization by using any one of them as your credal reallocation policy. If there exists no rcd for P given $\mathcal{G}$, Kolmogorov Conditionalization does not say how to proceed.[5]

I will now show that anyone who violates Kolmogorov Conditionalization is Dutch bookable. I assume the following setup. At time t_1, $(\Omega, \mathcal{F}, P)$ models your credal allocation. At time t_2, you and I will both gain full membership knowledge for $\mathcal{G} \subseteq \mathcal{F}$. Let $C(A \mid \omega)$ be the credence you would assign to $A \in \mathcal{F}$ upon gaining this knowledge for outcome ω. Given the argumentation of §4.2, we may assume that

---

[5] I do not assume that every probability space $(\Omega, \mathcal{F}, P)$ models a possible credal allocation or that every sub-σ-field $\mathcal{G}$ models a possible learning scenario. In certain cases, $(\Omega, \mathcal{F}, P)$ or $\mathcal{G}$ may be so bizarre that no possible agent satisfies the antecedent of Kolmogorov Conditionalization (even though there exists an rcd for P given $\mathcal{G}$).

$C(\,.\mid\omega)$ is a probability measure *for each $\omega\in\Omega$.*

$C(A\mid.\,)$ is $\mathcal{G}$-measurable *for all $A\in\mathcal{F}$*

These two assumptions echo the first two clauses in the definition of rcd.

If $C(A\mid.\,)$ satisfies the integral formula for each $A\in\mathcal{F}$, then C is an rcd and your credal reallocation policy conforms to Kolmogorov Conditionalization. Let us therefore assume that $C(A\mid.\,)$ violates the integral formula for some $A\in\mathcal{F}$. It follows that

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

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

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

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:

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

I now use G to rig a Dutch book containing sequential bets offered at times t_1 and t_2.

At time t_1, I offer you a conditional bet: if we learn at time t_2 that $\omega\in G$, then I will sell you for price $P_{\mathcal{G}}(A\mid\omega)$ a wager that pays off as follows:

$$\omega\in A\Rightarrow payoff=1$$

$$\omega\notin A\Rightarrow payoff=0.^{[6]}$$

We will be able in principle to determine $P_{\mathcal{G}}(A\mid\omega)$ at time t_2, because membership knowledge for $\mathcal{G}$ fixes $P_{\mathcal{G}}(A\mid\omega)$. Call this *bet 1*. Table 1 summarizes net gain for bet 1 given any outcome $\omega\in\Omega$. Bet 1 has payoff 1 when $\omega\in A\cap G$ and payoff 0 otherwise, so its expected payoff is $P(A\cap G)$. You pay price $P_{\mathcal{G}}(A\mid\omega)$ when $\omega\in G$ and price 0 otherwise, so the expected price is

---

[6] Here and elsewhere "$\Rightarrow$" is the material conditional.

$$\int_G P_{\mathbb{S}}(A \mid \omega)dP(\omega) \,.$$

By the integral formula, the expected price is $P(A \cap G)$. Since the expected payoff and the

expected price are equal, the expected net gain is 0. You therefore accept bet 1 as fair. You

commit to buying the specified wager for the specified price if it turns out that $\omega \in G$. (Cf.

Billingsley, 1995, p. 431.)

INSERT TABLE 1 ABOUT HERE

At time t_2, we both learn whether $\omega \in G$. If $\omega \notin G$, then no money changes hands. If $\omega \in G$,

then we enact the gambling transaction agreed upon at time t_1. Furthermore, I now ask *you* to sell

me for price $C(A \mid \omega)$ a wager that pays off as follows:

$$\omega \in A \Rightarrow payoff = 1$$

$$\omega \notin A \Rightarrow payoff = 0.$$

Call this *bet 2*. Table 2 summarizes net gain for bet 2. Bet 2 has payoff 1 when $\omega \in A$ and payoff

0 otherwise, so its expected payoff is your credence in A at time 2: namely, $C(A \mid \omega)$. This is also

bet 2's price, so expected net gain from bet 2 is 0. You therefore accept bet 2 as fair.

INSERT TABLE 2 ABOUT HERE

Table 3 summarizes your net gain for the overall gambling scenario, given any outcome

$\omega \in \Omega$. Net gain is negative if $\omega \in G$ and 0 if $\omega \notin G$. Since $P(G) > 0$, the overall gambling scenario

offers a positive probability of net loss and no compensating prospect of net gain. To quote

Lewis: "I can inflict on you a risk of loss uncompensated by any chance of gain" (1999, p. 406).

When $C(A \mid \omega) > P_{\S}(A \mid \omega)$ with positive probability, I can simply reverse the two bets: I buy bet 1 and sell bet 2. Either way, you are vulnerable to fair bets that offer a positive probability of net loss and no compensating positive probability of net profit.

Here we may distinguish a weaker and stronger notion of *Dutch bookability*. A *strong Dutch book* is a set of acceptable bets with guaranteed net loss. A *weak Dutch book* is a set of acceptable bets with a positive probability of net loss and no compensating positive probability of net profit. In §1, I used the phrase "Dutch book" to mean *strong Dutch book*, as is typical in the literature. I have just shown that anyone who violates Kolmogorov Conditionalization is *weakly* Dutch bookable. A minor emendation in §6 will show that any such agent is *strongly* Dutch bookable. Clearly, weak Dutch bookability is already very undesirable.

To rig a weak diachronic Dutch book against you, I must know your plan $C(A \mid \omega)$ for updating credences. Only then can I decide whether to buy or sell the relevant bets. A similar situation prevails in Lewis's Dutch book theorem for Ratio Conditionalization. As Lewis puts it, "I still have no safe strategy for exploiting you unless I *know* in advance what you will do instead of conditionalizing" (1999, p. 407). However, I think that we should not overemphasize "strategies" for exploiting non-conditionalizers. The core issue here does not concern *competition* between agents. The core issue concerns internal defects in *a single agent's* credal allocation over time. The main worry here is not that you are vulnerable to exploitation but rather that *your own credal allocation* depicts your credal reallocation policy as promoting pointlessly risky behavior. *By your own lights*, your credal reallocation policy can lead you to incur a

positive probability of net loss with no compensating positive probability of net gain. Thus, your credal reallocation policy has highly undesirable pragmatic properties *by your own lights*.[7]

§6. Dutch books formalized

I have shown that agents who violate Kolmogorov Conditionalization are vulnerable to a weak Dutch book. Might agents who obey Kolmogorov Conditionalization be just as badly off? Say that a *Kolmogorov conditionalizer* is someone who satisfies the antecedent of Kolmogorov Conditionalization and who conforms to Kolmogorov Conditionalization. A Kolmogorov conditionalizer has initial credences that admit a suitable rcd, and she reallocates credences using such an rcd. I will prove that one cannot rig a weak Dutch book against a Kolmogorov conditionalizer. To prove a negative result of this kind, I must first formalize the notions *update rule*, *bet*, *bookie strategy*, and *weak Dutch book*.

I assume an agent who at time t_1 has initial credences P over events from $(\Omega, \mathcal{F})$. At time t_2, she gains membership knowledge for sub-σ-field $\mathcal{G} \subseteq \mathcal{F}$. I assume that she updates her credences in conformity to some *update rule*. Intuitively, an update rule is a policy for reallocating credence given membership information for $\mathcal{G}$. One might therefore consider functions $U: \Omega \rightarrow M$, where M is the space of probability measures over $\mathcal{F}$. However, it will prove more convenient to focus instead on functions $C: \mathcal{F} \times \Omega \rightarrow \mathbb{R}$ satisfying two constraints:

$C(\,.\,, \omega)$ is a probability measure $\qquad\qquad$ *for all $\omega \in \Omega$.*

$C(A, \,.\,)$ is $\mathcal{G}$-measurable $\qquad\qquad$ *for all $A \in \mathcal{F}$.*

I say that any such C is an *update rule for (Ω, $\mathcal{F}$, $\mathcal{G}$)*. I will notate $C(A, \omega)$ as $C(A \mid \omega)$. If $C(A \mid .)$

satisfies the integral formula for each $A \in \mathcal{F}$, then C is an rcd for P given $\mathcal{G}$. Obviously, many

update rules are possible that violate the integral formula.

The agent faces a bookie who can offer bets over $\mathcal{F}$ 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 a random

variable $X: \Omega \to \overline{\mathbb{R}}$. Here $X(\omega)$ is the net gain for outcome ω. Infinite gains $-\infty$ and ∞ arise in

scenarios such as Pascal's Wager, but they have doubtful relevance to any realistic gambling

scenario. I allow them anyway. A random variable X is $\mathcal{F}$-measurable:

$X^{-1}(-\infty, a] \in \mathcal{F}$ *for each $a \in \mathbb{R}$*

$X^{-1}\{-\infty\} \in \mathcal{F}$

$X^{-1}\{\infty\} \in \mathcal{F}$.

$\mathcal{F}$-measurability reflects the idea that bets concern events in $\mathcal{F}$. Learning which events in $\mathcal{F}$

occurred must suffice at least in principle to determine the bet's net gain. This constraint leads

naturally to $\mathcal{F}$-measurability, as in §4.2's discussion of $\mathcal{G}$-measurability.

If X is a random variable and μ is a probability measure, then the *expectation of X with*

respect to μ is written as $E_\mu[X]$ and defined in the usual way:

$$E_\mu[X] =_{df} \int X d\mu .$$

Depending on our choice of X, $E_\mu[X]$ may or may not be well-defined. Standard decision theory

only applies to scenarios where $E_\mu[X]$ is well-defined, since only then does expected utility

maximization offers any guidance. It is common in probability theory and Bayesian decision

theory to restrict attention to such scenarios by demanding that bets have well-defined expectations. I impose no such restriction here. If the bookie wishes to offer bets without well-defined expected values, then I raise no objection.

Potential bets are evaluated for acceptability with respect to the agent's current credences. A bet *X* is *fair relative to* μ iff its expected value with respect to μ is 0:

$$E_\mu[X] = 0,$$

favorable relative to μ iff its expected value with respect to μ is positive:

$$E_\mu[X] > 0,$$

and *acceptable relative to* μ iff it is fair or favorable relative to μ. Thus, a bet is unacceptable if it has a negative expectation *or* it does not have a well-defined expectation. I assume that our agent adopts a policy of accepting all acceptable bets that are offered and rejecting all unacceptable bets. The assumed policy may not be rationally obligatory, but it is rationally permissible. We may legitimately consider an agent who adopts this policy.

I assume that the bookie can offer only finitely many bets at a given time.[8] All such bets are summable into a single random variable, so we may assume that the bookie offers at most a single bet at t_1 and a single bet at t_2.[9] At t_1, the agent evaluates whether the proposed bet is

---

[8] McGee (1999) shows that, if payoffs are unbounded and infinitely many bets are allowed, then an agent satisfying very weak assumptions faces an *infinite Dutch book*: an infinite collection of acceptable (indeed, favorable) bets that inflict a sure loss. McGee concludes that any rational agent must have bounded utilities. Following Arntzenius, Elga, and Hawthorne (2003), I draw a different conclusion: infinite Dutch books do not reveal any pragmatic or epistemic defect. A basic mathematical fact is that the expected value of the sum of infinitely many random variables need not equal the sum of their expected values. In betting terms: a book containing infinitely many bets need not be acceptable even though each individual bet is acceptable (indeed, favorable). That an agent would regard each bet as acceptable when presented individually does not entail that the agent should accept the overall package of infinitely many bets. I therefore ignore books that contain infinitely many bets.

[9] The sum of two well-defined bets may not itself be well-defined for certain ω. For example, we may have $X_1(\omega) = \infty$ and $X_2(\omega) = -\infty$. However, we may assume that all bets offered at a given time have a well-defined sum for each ω. If a book violates this assumption, then it is indeterminate what net gain an agent who accepts the book receives in certain outcomes. This book does not seem to me to constitute a well-defined gambling scenario. In any event, such books raise many problems that are orthogonal to our main concerns. We may legitimately ignore these books. (Alternatively, we could allow books where all bets offered at a given time have a well-defined sum *except for*

acceptable relative to her initial credences P. At t_2, she evaluates acceptability using updated credences that incorporate membership knowledge for $\mathcal{G}$. More precisely, she computes expectations relative to her new credences $C(\,.\mid\omega)$, where ω is the true outcome. In this spirit, let $\mu_\omega(\,.\,) = C(\,.\mid\omega)$, and say that bet X *is acceptable given* ω iff

$$E_{\mu_\omega}[X] = \int X d\mu_\omega \geq 0\,.$$

The agent accepts bet X at t_2 just in case X is acceptable given the true outcome.

A diachronic Dutch book has two elements: a bet X offered at t_1; and a strategy used by the bookie when deciding which bet to offer at t_2. The Dutch book from §5 features the strategy:

Offer bet 2 if $\omega \in G$; offer no bet if $\omega \notin G$.

Of course, we can imagine more complex bookie strategies. We would like a framework for modeling all such strategies.

A bookie strategy maps information received at t_2 into a bet offered at t_2. In the general case, a bookie might receive information about events drawn from space $(\Omega_1, \mathcal{F}_1)$ and offer a bet concerning events in a different space $(\Omega_2, \mathcal{F}_2)$. For present purposes, we need not proceed so generally. We are focused exclusively on Kolmogorov learning scenarios, where the agent updates credences for $(\Omega, \mathcal{F})$ based on membership knowledge for $\mathcal{G} \subseteq \mathcal{F}$. Moreover, we only consider scenarios where the bookie and the agent receive the same information at t_2. So the only relevant bookie strategies are those that map membership information about $\mathcal{G}$ into bets over $\mathcal{F}$. We use $(\Omega, \mathcal{G})$ to model the information that the bookie consults when selecting which bet to offer, and we use $(\Omega, \mathcal{F})$ to model the information that an observer consults when evaluating the net gain from whatever bet the bookie selects.

---

outcomes ω contained in a probability zero event. This generalization would not appreciably impact the definitions or proofs offered below.)

In this setting, it is natural to define a bookie strategy as a mapping from Ω to random variables over $(\Omega, \mathcal{F})$. However, it will prove convenient to proceed more circuitously. Let

$$\mathcal{G} \otimes \mathcal{F} =_{df} \sigma(\mathcal{G} \times \mathcal{F})$$

be the σ-field generated by the rectangles

$$G \times F \qquad\qquad\qquad\qquad\qquad\qquad G \in \mathcal{G}, F \in \mathcal{F}.$$

Consider the measurable space $(\Omega \times \Omega, \mathcal{G} \otimes \mathcal{F})$. A *bookie strategy* is a $\mathcal{G} \otimes \mathcal{F}$–measurable function $Y \colon \Omega \times \Omega \to \overline{\mathbb{R}}$. For fixed ω, $Y(\omega, \,. \,) \colon \Omega \to \overline{\mathbb{R}}$ is the bet that the bookie offers upon learning whether $\omega \in G$ for each $G \in \mathcal{G}$. Since Y is $\mathcal{G} \otimes \mathcal{F}$–measurable, one can easily show that

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

so that $Y(\omega, \,. \,)$ is indeed a bet according to our official definition. I will commonly abbreviate $Y(\omega, \,. \,)$ as Y_ω. To model scenarios where the bookie offers no bet upon learning full membership information for $\mathcal{G}$ regarding ω, I set $Y_\omega(\nu) = 0$ for all inputs ν.

The $\mathcal{G} \otimes \mathcal{F}$-measurability requirement may look a bit mysterious, so let me elucidate it. Suppose for purposes of this paragraph that "information" received by the bookie may be non-veridical. We use $\mathcal{G} \otimes \mathcal{F}$ to model the implicit knowledge of an observer who learns what information was transmitted to the bookie *and* learns which events in $\mathcal{F}$ occurred. Any such observer should be able in principle to determine whether the bookie's selected bet has gain $\leq a$. She should acquire implicit knowledge whether the proposition *The bet selected by the bookie has net gain $\leq a$* is true. This proposition corresponds to the event

$$\{(\omega, \nu) \colon Y_\omega(\nu) \leq a\} = Y^{1}(-\infty, a].$$

$\mathcal{G} \otimes \mathcal{F}$-measurability requires that each such event belong to $\mathcal{G} \otimes \mathcal{F}$. Thus, $\mathcal{G} \otimes \mathcal{F}$–measurability requires that an observer who knows what information the bookie received *and* which events in $\mathcal{F}$ occurred is able to decide whether the bet selected by the bookie has net gain $\leq$ or $> a$.

For any bookie strategy Y, there is a very important "diagonal function" $Y^*: \Omega \to \overline{\mathbb{R}}$ defined by

$$Y^*(\omega) =_{df} Y_\omega(\omega).$$

See Figure 1. $Y^*(\omega)$ is net gain in outcome ω for the bet dictated by bookie strategy Y in outcome ω. So Y^* specifies the agent's net gain if she accepts whatever bet the bookie offers at t_2.

INSERT FIGURE 1 ABOUT HERE

We can now formalize Dutch bookability. A **strong Dutch book for probability space** $(\mathbf{\Omega}, \mathcal{F}, P)$, **sub-$\sigma$-field** $\mathcal{G} \subseteq \mathcal{F}$, **and update rule C** is a pair (X, Y) such that

    (a) X is a bet that is acceptable relative to P.

    (b) Y is a bookie strategy.

    (c) Y_ω is acceptable given $\omega \Rightarrow X(\omega) + Y_\omega(\omega) < 0$.

    (d) Y_ω is not acceptable given $\omega \Rightarrow X(\omega) < 0$.

(a) requires that X have a nonnegative expectation at t_1. This condition ensures that the agent will accept bet X at t_1. Collectively, (a)-(d) ensure that a bookie who offers bet X and pursues bookie strategy Y will inflict a net loss in all outcomes. Note that (d) constrains net gain from X rather than net gain from $X + Y_\omega$. Our betting agent will reject bet Y_ω in outcomes ω where Y_ω is unacceptable, so only X matters for computing net gain in such outcomes. See Figure 2.

A *weak Dutch book for probability space* $(\Omega, \mathcal{F}, P)$, *sub-σ-field* $\mathcal{G} \subseteq \mathcal{F}$, *and update rule* C is a pair (X, Y) such that

    (a) X is a bet that is acceptable relative to P.

    (b) Y is a bookie strategy.

    (c) $P\{\omega\colon Y_\omega$ is acceptable given ω & $X(\omega) + Y_\omega(\omega) < 0\} > 0$.

    (d) $P\{\omega\colon Y_\omega$ is acceptable given ω & $X(\omega) + Y_\omega(\omega) > 0\} = 0$.

    (e) $P\{\omega\colon Y_\omega$ is not acceptable given ω & $X(\omega) > 0\} = 0$.

(a) ensures that the agent will happily accept bet X at t_1. (c) ensures a positive probability that the agent accepts bet Y_ω at t_2 and thereby incurs a net loss. (d) ensures that there is no compensating positive probability of net profit in outcomes ω where the agent accepts bet Y_ω. (e) ensures that there is no compensating positive probability of net profit in outcomes ω where the agent rejects bet Y_ω as unacceptable. (e) is very important, because without it (X, Y) need not comprise anything like a strategy for exploiting the agent. A strategy for exploiting an agent must have zero probability of rewarding the agent with net profit.

§10 will show that a strong Dutch book is a weak Dutch book. When there exists a strong (weak) Dutch book for $(\Omega, \mathcal{F}, P)$, $\mathcal{G}$, and C, say that C is *strongly (weakly) Dutch bookable*.

A weaker notion that sometimes figures in the literature is *semi-Dutch book*. A semi-Dutch book is a set of acceptable bets with a possibility of net loss and no possibility of net gain. This is less demanding than the notion *weak Dutch book*, because the probability of net loss from a semi-Dutch book may be 0. A weak Dutch book offers a *positive probability* of net loss, not just a *possibility* of net loss. To illustrate, suppose that your credences violate regularity *in the sense of §2*, i.e. you assign credence 0 to a metaphysically possible proposition. Then you should

happily pay price 1 for a wager that returns payoff 1 just in case the proposition is false. The resulting bet is a semi-Dutch book but not a weak Dutch book.

Shimony deploys semi-Dutch bookability to defend regularity. I agree with Hájek (2009, 2012) that Shimony's argument is not compelling. As Hájek (2009, pp. 188-189) urges, the mere possibility of net loss need not be worrisome if you are 100% confident that the possibility will not materialize. *From your viewpoint*, a semi-Dutch book that is not a weak Dutch book carries no risk of net loss. It seems rationally permissible for you to regard such a book as perfectly agreeable. You *may* be worried by a probability 0 possibility of net loss, but rationality does not *require* you to be worried. Thus, semi-Dutch bookability does not in itself suggest that any serious pragmatic or epistemic defect afflicts your credal reallocations.

Kolmogorov conditionalizers are often semi-Dutch bookable. Suppose you plan to update credal assignment $P(A)$ using $C(A \mid .)$, a conditional probability for A given $\mathcal{G}$. Suppose that there exists $P_{\mathcal{G}}(A \mid .)$, a conditional probability for A given $\mathcal{G}$, such that $C(A \mid .)$ and $P_{\mathcal{G}}(A \mid .)$ disagree on outcomes belonging to some set G. Then I can employ the strategy from §5. I can construct sequential fair bets that inflict upon you a net loss for outcomes in G and net gain 0 on outcomes outside G. The catch is that G itself must have probability 0, so that your net loss only occurs with probability 0. In that case, my strategy is a semi-Dutch book but not a weak Dutch book. As argued in the previous paragraph, vulnerability to semi-Dutch books does not suggest that Kolmogorov Conditionalization is pragmatically or epistemically problematic.

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

Proof: $C(A \mid .)$ must violate the integral formula for some $A \in \mathcal{F}$. Let $P_\mathcal{G}(A \mid .)$ be a conditional

probability for A given $\mathcal{G}$. As in §5, we may assume without loss of generality that $G =_{df} \{ \omega: C(A$

$\mid \omega) < P_\mathcal{G}(A \mid \omega) \}$ has non-zero P-measure. We formalize the procedure from §5, supplemented

with a sidebet on G at t_1 to ensure a net loss when $\omega \notin G$. Define random variable X by

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

For $\omega \in G$, define random variable Y_ω by

$$Y_\omega(v) = \begin{cases} C(A \mid \omega) - 1 & \text{if } v \in A \\ C(A \mid \omega) & \text{if } v \notin A \end{cases}$$

and for $\omega \notin G$ define random variable Y_ω by

$$Y_\omega(v) = 0.$$

These definitions determine a bookie strategy $Y(\omega, v) =_{df} Y_\omega(v)$. X formalizes bet 1 from §5. Y

formalizes the strategy: *offer bet 2 if $\omega \in G$; offer no bet if $\omega \notin G$*. Let

$L =_{df} \int_G \left[C(A \mid \omega) - P_\mathcal{G}(A \mid \omega) \right] dP(\omega)$ if this integral is finite. If the integral is infinite, then let L be

any finite negative number. Either way, we have

$$\int_G \left[C(A \mid \omega) - P_\mathcal{G}(A \mid \omega) \right] dP(\omega) \leq L.$$

Define random variable Z:

$$Z(\omega) = \begin{cases} (P(G) - 1) \left[C(A \mid \omega) - P_\mathcal{G}(A \mid \omega) \right] & \text{if } \omega \in G \\ L & \text{if } \omega \notin G \end{cases}$$

We show that $(X + Z, Y)$ is a strong Dutch book for $(\Omega, \mathcal{F}, P)$, $\mathcal{G}$, and C.

§5 already showed that X is fair relative to P, but we now offer a somewhat more formal proof. For any set S, let I_S be the indicator function for S:

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

For $\omega \in G$, we have

$$X(\omega) = \left[1 - P_{\mathfrak{S}}(A \mid \omega)\right] I_A(\omega) - P_{\mathfrak{S}}(A \mid \omega) I_{A^c}(\omega) = I_A(\omega) - P_{\mathfrak{S}}(A \mid \omega).$$

For any $\omega \in \Omega$, we have

$$X(\omega) = I_G(\omega)\left[I_A(\omega) - P_{\mathfrak{S}}(A \mid \omega)\right],$$

so that

$$E_P[X] = \int_\Omega I_G(\omega)\left[I_A(\omega) - P_{\mathfrak{S}}(A \mid \omega)\right] dP(\omega) = \int_\Omega I_G I_A dP - \int_\Omega I_G(\omega) P_{\mathfrak{S}}(A \mid \omega) dP(\omega)$$

$$= \int_\Omega I_{A \cap G} dP - \int_G P_{\mathfrak{S}}(A \mid \omega) dP(\omega) = P(A \cap G) - \int_G P_{\mathfrak{S}}(A \mid \omega) dP(\omega) = P(A \cap G) - P(A \cap G) = 0,$$

where the penultimate identity follows by the integral formula. To confirm that Z is acceptable relative to P, note that

$$Z(\omega) = I_G(\omega)(P(G) - 1)\left[C(A \mid \omega) - P_{\mathfrak{S}}(A \mid \omega)\right] + I_{G^c}(\omega) L,$$

so that

$$E_P[Z] = \int_\Omega \left[I_G(\omega)(P(G) - 1)\left[C(A \mid \omega) - P_{\mathfrak{S}}(A \mid \omega)\right] + I_{G^c} L\right] dP(\omega)$$

$$= \int_\Omega I_G(\omega)(P(G) - 1)\left[C(A \mid \omega) - P_{\mathfrak{S}}(A \mid \omega)\right] dP(\omega) + \int_\Omega I_{G^c} L dP(\omega)$$

$$= \int_G (P(G) - 1)\left[C(A \mid \omega) - P_{\mathfrak{S}}(A \mid \omega)\right] dP(\omega) + L P(G^c)$$

$$\geq (P(G) - 1) L + L P(G^c) = L(P(G) + P(G^c) - 1) = L(1 - 1) = 0.$$

As for Y, we may write

$$Y_\omega = C(A \mid \omega) - I_A$$

for any $\omega \in G$. Formalizing the reasoning from §5, we check that Y_ω is fair given ω for any $\omega \in G$:

$$E_{\mu_\omega}[Y_\omega] = \int_\Omega [C(A \mid \omega) - I_A] d\mu_\omega = \int_\Omega C(A \mid \omega) d\mu_\omega - \int_\Omega I_A d\mu_\omega$$

$$= C(A \mid \omega) \int_\Omega d\mu_\omega - \mu_\omega(A) = C(A \mid \omega) - C(A \mid \omega) = 0,$$

where ω in these equations is held fixed and is not an integration variable. Clearly, Y_ω is also fair given ω for any $\omega \notin G$.

To complete the proof, we must show that $X(\omega) + Z(\omega) + Y_\omega(\omega) < 0$ for all $\omega \in \Omega$. If $\omega \in G$, then routine calculation confirms that

$$X(\omega) + Z(\omega) + Y_\omega(\omega) = P(G) [C(A \mid \omega) - P_{\mathcal{G}}(A \mid \omega)] < 0.$$

If $\omega \notin G$, then

$$X(\omega) + Z(\omega) + Y_\omega(\omega) = L < 0.$$

Hence, $(X + Z, Y)$ is a strong Dutch book for $(\Omega, \mathcal{F}, P)$, $\mathcal{G}$, and C. □

Skyrms (1992) suggests that a genuine Dutch book should contain bets that are favorable, not just acceptable. We can strengthen the foregoing theorem to accommodate Skyrms's viewpoint. In particular, we can supplement all bets with a "sweetener" as follows. Define

$$S(\omega) = \frac{P(G)[P_{\mathcal{G}}(A \mid \omega) - C(A \mid \omega)]}{4} .$$

Define a bookie strategy $W(\omega, v)$:

$$W(\omega, v) = Y(\omega, v) + I_G(\omega)\big[I_G(v)S(v) + I_{G^c}(v)\big].$$

Consider the sweetened pair $((X + I_G S) + (Z + I_G S), W)$. One can show that $X + I_G S$ and $Z + I_G S$ are each favorable at t_1 and that W_ω is favorable given ω for every $\omega \in G$. One can also check that

$W_\omega = 0$ for any $\omega \notin G$, corresponding to a situation where no bet is offered at t_2. So $((X + I_G S) +$

$(Z + I_G S), W)$ models a gambling scenario where all proffered bets are favorable. Nevertheless,

net payoff from the overall scenario is always < 0. Hence, agents who violate Kolmogorov

Conditionalization are vulnerable to a set of *favorable* bets that inflict a sure loss.

§7. A converse Dutch book theorem for Kolmogorov Conditionalization

This section proves that Kolmogorov conditionalizers are not Dutch bookable. The proof

resembles Skyrms's proof of the converse Dutch book theorem for Ratio Conditionalization. In

both proofs, the basic idea is that a diachronic Dutch book for a conditionalizer could be

converted into a synchronic book with impossible properties. Developing this idea requires much

more mathematical machinery for the general case of Kolmogorov Conditionalization than for

the special case of Ratio Conditionalization. I will first offer some heuristic remarks and then

present a rigorous proof.

Suppose for reductio that you are a Kolmogorov conditionalizer and that there exists a

weak diachronic Dutch book (X, Y) for your update rule. Let

$\Delta =_{df} \{ \omega : Y_\omega \text{ is acceptable given } \omega \}$.

Now consider a bet Z^* defined as follows:

$\omega \in \Delta \Rightarrow Z^*(\omega) = Y_\omega(\omega)$

$\omega \notin \Delta \Rightarrow Z^*(\omega) = 0$.

Think of Z^* as a conditional bet offered at t_1:

You accept the bet Y_ω offered at t_2 if that bet is acceptable; otherwise you decline.

See Figure 3. This conditional bet should be acceptable at t_1, since it only commits you to betting in situations where you find the bet offered at t_2 acceptable. Given that X and Z^* are individually acceptable at t_1, the combined bet $X + Z^*$ is also acceptable at t_1. In other words:

(1) $\quad E_P[X + Z^*] \geq 0.$

Since (X, Y) is a weak Dutch book, $X(\omega) + Y_\omega(\omega) < 0$ with positive probability inside Δ. Thus:

(2) $\quad X + Z^* < 0$ with positive probability inside Δ.

Since (X, Y) is a weak Dutch book, $X(\omega) + Y_\omega(\omega) > 0$ with probability 0 inside Δ. Thus:

(3) $\quad X + Z^* > 0$ with probability 0 inside Δ.

Since (X, Y) is a weak Dutch book, $X > 0$ with probability 0 outside Δ. Thus:

(4) $\quad X + Z^* > 0$ with probability 0 outside Δ.

(1)-(4) are mutually inconsistent: the negative values ensured by (2) find no counterbalancing positive values to generate the nonnegative expected value promised by (1). By contradiction, Kolmogorov conditionalizers are not weakly Dutch bookable.

This reasoning hinges upon the presupposition that Z^* is acceptable at t_1. The presupposition is plausible, but why should we believe it? In fact, the presupposition is *not* true for agents who violate Kolmogorov Conditionalization. Take the bookie strategy Y from the Dutch book theorem:

Offer bet 2 if $\omega \in G$; offer no bet if $\omega \notin G$.

Y_ω is acceptable given ω for every $\omega \in \Omega$, yet one can easily check that the corresponding bet Z^*:

You accept bet 2 if $\omega \in G$; you do not bet if $\omega \notin G$.

is unacceptable at t_1. However, I prove below that the presupposition is true for Kolmogorov conditionalizers.

The proof requires a crucial lemma. Suppose that $P_\mathcal{G}$ is an rcd for P given $\mathcal{G}$. Let $\mu_\omega(.) = P_\mathcal{G} (. , \omega)$. For any bookie strategy Z, consider the function $\mathbb{E}_Z: \Omega \to \overline{\mathbb{R}}$ defined by

$$\mathbb{E}_Z(\omega) =_{df} E_{\mu_\omega}[Z_\omega] = \int Z_\omega d\mu_\omega = \int_\Omega Z(\omega, v) \mu_\omega(dv).$$

$\mathbb{E}_Z(\omega)$ is the expected value a Kolmogorov conditionalizer computes at t_2 for the bet dictated by bookie strategy Z. Thus, $\mathbb{E}_Z$ maps each outcome to the net gain that a Kolmogorov conditionalizer would expect at t_2 if she accepted the bet offered at t_2. The lemma basically says that averaging together these expected net gains yields the same result as computing the net gain our agent would expect at t_1 if she resolved to accept the bet offered at t_2. More carefully: if the diagonal function Z^* has an expectation with respect to initial credences P, then that expectation equals the expectation of $\mathbb{E}_Z$ with respect to P.

Lemma: *Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\mathcal{G} \subseteq \mathcal{F}$ a sub-σ-field. Suppose there exists $P_\mathcal{G}$, an rcd for P given $\mathcal{G}$, and let $\mu_\omega(.) = P_\mathcal{G} (. , \omega)$. Let $Z: \Omega \times \Omega \to \overline{\mathbb{R}}$ be $\mathcal{G} \otimes \mathcal{F}$-measurable. Define $\mathbb{E}_z(\omega) =_{df} E_{\mu_\omega}[Z_\omega]$, which may be infinite or undefined for certain ω. Then $\mathbb{E}_Z$ is $\mathcal{G}$-measurable. Let Z^* be the diagonal function defined by $Z^*(\omega) = Z_\omega(\omega)$. If $E_P[Z^*]$ exists, then $\mathbb{E}_Z$ is defined for P-almost all values and $E_P[Z^*] = E_P[\mathbb{E}_Z]$.*

I prove the lemma in a mathematical appendix (§10).

The Dutch book theorem generates a conflict between present and future computations of expected value. The bookie strategy Y:

Offer bet 2 if $\omega \in G$; offer no bet if $\omega \notin G$.

yields a bet with nonnegative expectation as computed at t_2 or else yields a non-bet with net gain

0 for all outcomes. So $E_{\mu_\omega}[Y_\omega] \geq 0$ for every ω, where we have set $\mu_\omega(\,.\,) = C(\,.\mid\omega)$. The

expected value at t_1 of those nonnegative expectations --- $E_P\big[E_{\mu_\omega}[Y_\omega]\big]$ --- is nonnegative as well.

Nevertheless, the conditional bet Z^*:

> *You accept bet 2 if $\omega \in G$; you do not bet if $\omega \notin G$.*

has negative expectation when computed at t_1. Our lemma shows that no such conflict can arise

for Kolmogorov conditionalizers. If $E_P[Z^*]$ exists, then it must equal $E_P[\mathbf{E}_Z]$. If Z is a bookie

strategy that always yields acceptable bets at t_2, and if $E_P[Z^*]$ exists, then Z^* must already be

acceptable at t_1. As I now show, this harmony between present and future expectations

immunizes Kolmogorov conditionalizers from diachronic Dutch books.

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 update rule for $(\Omega,\ \mathcal{F},\ \mathcal{G})$. If C is an

rcd for P given $\mathcal{G}$, then there does not exist a weak Dutch book for $(\Omega,\ \mathcal{F},\ P)$, $\mathcal{G}$, and C.

Proof: Suppose for reductio that there is a weak Dutch book $(X,\ Y)$ for $(\Omega,\ \mathcal{F},\ P)$, $\mathcal{G}$, and C. Let

$\mu_\omega(\,.\,) =_{df} C(\,.\mid\omega)$, and let

$$\Delta =_{df} \{\,\omega\colon Y_\omega \text{ is acceptable given } \omega\,\} = \{\,\omega\colon \mathbf{E}_Y(\omega) \geq 0\,\}.$$

The lemma shows that $\mathbf{E}_Y$ is $\mathcal{G}$-measurable, from which it follows that $\Delta \in \mathcal{G}$. Define bookie

strategy Z by

$$Z(\omega,.) =_{df} \begin{cases} Y(\omega,.) & \text{if } \omega \in \Delta \\ 0 & \text{if } \omega \notin \Delta \end{cases}$$

Since $\Delta \in \mathcal{G}$, Z is $\mathcal{G} \otimes \mathcal{F}$–measurable. Partition Δ as follows:

$$\Delta_- =_{df} \Delta \cap \{ \omega: X(\omega) + Z^*(\omega) < 0\} = \Delta \cap \{ \omega: X(\omega) + Y_\omega(\omega) < 0\}$$

$$\Delta_0 =_{df} \Delta \cap \{ \omega: X(\omega) + Z^*(\omega) = 0\} = \Delta \cap \{ \omega: X(\omega) + Y_\omega(\omega) = 0\}$$

$$\Delta_+ =_{df} \Delta \cap \{ \omega: X(\omega) + Z^*(\omega) > 0\} = \Delta \cap \{ \omega: X(\omega) + Y_\omega(\omega) > 0\}.$$

Since (X, Y) is a weak Dutch book, $P(\Delta_-) > 0$ and $P(\Delta_+) = 0$. We therefore have

$$\int_\Delta (X + Z^*)dP = \int_{\Delta_-}(X + Z^*)dP + \int_{\Delta_0}(X + Z^*)dP + \int_{\Delta_+}(X + Z^*)dP = \int_{\Delta_-}(X + Z^*)dP + 0 + 0 < 0$$

By similar reasoning,

$$\int_{\Delta^c}(X + Z^*)dP \le 0.$$

Thus,

$$\int_\Omega (X + Z^*)dP = \int_\Delta (X + Z^*)dP + \int_{\Delta^c}(X + Z^*)dP < 0,$$

so that $E_P[X + Z^*] < 0$.

We now apply the lemma to Z^* so as to derive a conflicting value for $E_P[X + Z^*]$. We must first check that $E_P[Z^*]$ exists. For any random variables V and W, define

$$V^+(\omega) =_{df} \begin{cases} V(\omega) & \text{if } V(\omega) \ge 0 \\ 0 & \text{if } V(\omega) < 0 \end{cases}$$

$$W^-(\omega) =_{df} \begin{cases} -W(\omega) & \text{if } W(\omega) \le 0 \\ 0 & \text{if } W(\omega) > 0 \end{cases}$$

and note that

$$W^+ \le (V + W)^+ + V^-.$$

All these functions are nonnegative, so

$$0 \le E_P[W^+] \le E_P[(V + W)^+] + E_P[V^-].$$

In particular,

$$0 \le E_P\left[(Z*)^+\right] \le E_P\left[(X + Z*)^+\right] + E_P\left[X^-\right].$$

We have assumed that X is acceptable, so that $E_P[X] \ge 0$. Thus, $E_P\left[X^-\right]$ must certainly be finite.

We have also just shown that $E_P[X + Z*] < 0$, so that $E_P\left[(X + Z*)^+\right]$ must likewise be finite. It

follows that $E_P\left[(Z*)^+\right]$ is finite. Hence, $E_P[Z*]$ exists. Applying the lemma,

$$E_P[Z*] = E_P\left[\mathbb{E}_Z\right].$$

We have chosen Z so that

$$\mathbb{E}_Z(\omega) = \int Z_\omega d\mu_\omega \ge 0 \qquad\qquad \textit{for all } \omega \in \Omega.$$

The integral of a nonnegative function is nonnegative, so

$$E_P[Z*] = E_P\left[\mathbb{E}_Z\right] \ge 0.$$

Since $E_P[X] \ge 0$, we conclude that

$$E_P[X + Z*] = E_P[X] + E_P[Z*] \ge 0$$

which contradicts our earlier finding that $E_P[X + Z*] < 0$. By reductio, there is no weak Dutch

book for $(\Omega, \mathcal{F}, P)$, $\mathcal{G}$, and C. □

§8. Significance of the two theorems

The Dutch book theorem and converse Dutch book theorem show that Kolmogorov's

theory delineates conditional probabilities with uniquely desirable pragmatic properties. It is

good to avoid Dutch books. Thus, it is good when credal reallocation is invulnerable to Dutch

books. The theorems establish Kolmogorov Conditionalization as the sole credal reallocation

policy that achieves the desired invulnerability in Kolmogorov learning scenarios where rcds

exist. The forbidding mathematics of Kolmogorov's theory should not distract us from the fact

that it codifies fundamental ties between conditional probability, credal reallocation, and decision-making under uncertainty.

To explore the significance of the theorems, let us revisit two worries about Kolmogorov's approach raised in §4: *non-existence of rcds* and *residual indeterminacy after specifying a conditioning sub-σ-field.*

§8.1 Non-existence of rcds

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\mathcal{G} \subseteq \mathcal{F}$ a sub-σ-field such that there exists no rcd for P given $\mathcal{G}$. Suppose that you have initial credences P over $\mathcal{F}$ and then gain new evidence exhausted by full membership knowledge for $\mathcal{G}$.[10] How should you update your credences over events in $\mathcal{F}$? Kolmogorov Conditionalization does not say. It remains silent about learning scenarios where rcds do not exist. We can now identify a good reason for this silence: the Dutch book theorem shows that all options are problematic.

Suppose that $C(A \mid \omega)$ is the credence you would assign to $A \in \mathcal{F}$ upon gaining full membership knowledge for $\mathcal{G}$ regarding outcome ω. As in §5, we may assume that

$C(A \mid .)$ is $\mathcal{G}$-measurable   *for all $A \in \mathcal{F}$.*

Suppose that $C(A \mid .)$ violates the integral formula for some $A \in \mathcal{F}$. Then the Dutch book theorem shows that I can rig a strong Dutch book against you. Thus, you are strongly Dutch bookable if you do not employ some conditional probability $P_{\mathcal{G}}(A \mid .)$ as your policy for reassigning credence to A. The problem is that, if you *do* adopt such a policy for each $A \in \mathcal{F}$, then your

---

[10] I assume for the sake of argument that $(\Omega, \mathcal{F}, P)$ models a possible credal allocation. It is not obvious that this assumption is correct, because the usual examples where rcds do not exist involve a σ-field $\mathcal{F}$ defined non-constructively through the Axiom of Choice.

credences at t_2 will not always constitute a probability measure. You will violate countable

additivity for certain outcomes ω.

It is controversial whether credences should be countably additive. As mentioned in §1, I

wish to side-step these controversies. I assume that credences should be countably additive, at

least for idealized agents who figure in Kolmogorov learning scenarios. Under that assumption,

the Dutch book theorem shows that a learning scenario modeled by $(\Omega, \mathcal{F}, P)$ and $\mathcal{G}$ is

problematic. Anyone in such a learning scenario must either succumb to a strong Dutch book or

else violate countable additivity.

In my opinion, the most important moral here is that you should avoid these problematic

learning scenarios. By choosing a sufficiently inauspicious probability space as your starting

point, you set yourself a rational dilemma between strong Dutch bookability and countable

additivity violations. To avoid the dilemma, you should refrain from having credences modeled

by a pathological probability space that does not admit rcds. You should adopt a policy of

maintaining a well-behaved credal allocation that admits rcds. This policy implicitly guides all

serious empirical applications of Bayesian modeling within statistics, economics, robotics,

cognitive science, and so on. The policy is easy to implement, because naturally arising

probability spaces always admit rcds.

§8.2 Residual indeterminacy

Kolmogorov does not delineate unique conditional probabilities *even after one fixes a*

conditioning sub-σ-field. When $\mathcal{G}$ contains a non-empty probability zero event, $P_{\mathcal{G}}(A \mid .)$ is

uniquely determined only up to measure 0. Nor can Dutch books help pin down conditional

probabilities more determinately: the converse Dutch book theorem shows that updating in

accord with *any* rcd $P_\mathcal{G}$ avoids a diachronic Dutch book. That is why I formulated Kolmogorov Conditionalization as an *indeterministic constraint* on credal reallocation rather than a *deterministic instruction* that yields unique reallocated credences.

We see here a foundational basis for the indeterminacy in Kolmogorov's theory. Kolmogorov specifies conditional probabilities as uniquely as Dutch book considerations allow. *If the only constraint on rational credal reallocation is that one avoid Dutch books*, then Kolmogorov Conditionalization is the most determinate norm we can expect. *If the only constraint on conditional probability is that it subserve credal reallocations that avoid Dutch books*, then Kolmogorov pins down conditional probabilities as determinately as possible.

Would you like more determinate conditional probabilities than Kolmogorov provides? Then you must look beyond Dutch books. You must examine the broader role that conditional probability plays within our cognitive lives. In principle, one might try to motivate more determinate conditional probabilities through either pragmatic or epistemic considerations. However, I doubt that pragmatic factors can pin down conditional probabilities more determinately than Kolmogorov's theory. Rational decision-making compares expected values, and expected values obliterate differences among alternative rcds for P given $\mathcal{G}$. Thus, I doubt that we can render Kolmogorov's theory any more determinate by examining how conditional probability figures in rational decision-making. Whether *epistemic* factors can generate more determinacy is a question worth further exploration.

§9. Conclusion

Kolmogorov's theory of conditional probability is acclaimed by mathematicians and neglected by philosophers. The mathematicians are right. Kolmogorov's theory brilliantly

exemplifies how formal mathematics can elucidate core philosophical concepts. It deserves mention in the same breath with Turing's analysis of computability and Tarski's analyses of truth and logical consequence. At the very least, Kolmogorov articulates a systematic framework that codifies how conditional probability, credal allocation, and decision-making interact in numerous important learning scenarios. I hope that the theorems proved above will promote wider appreciation of rcds as invaluable analytical tools.

§10. Mathematical appendix

This appendix proves the lemma from §7. The lemma follows from the *conditional Fubini theorem* (Fristedt and Gray, 1997, p. 431). I have thought it best to prove the lemma directly, partly because doing so takes only a little more space than proving that the conditional Fubini theorem entails the lemma, partly because a self-contained proof of the conditional Fubini theorem does not seem to be readily accessible anywhere in the literature. Throughout my discussion, I employ the conventions of (Fristedt and Gray, 1997, p. 48, p. 445) regarding *partially defined functions* and *almost surely defined random variables*.

Proof of the lemma: We will first prove the lemma for the special case where Z is everywhere nonnegative, then prove it for general Z. Let us begin by transforming $\int Z * dP$ into a more useful form. Let $T: \Omega \rightarrow \Omega \times \Omega$ be the "diagonal embedding"

$$T(\omega) = (\omega, \omega),$$

which is a measurable function from $(\Omega, \mathscr{F})$ to $(\Omega \times \Omega, \mathscr{G} \otimes \mathscr{F})$. T induces a measure $P*$ on $\mathscr{G} \otimes \mathscr{F}$:

$$P* =_{df} PT^{-1}.$$

For any $E \in \mathscr{G} \otimes \mathscr{F}$,

$$P*(E) = PT^{-1}(E) = P\{\omega\colon (\omega,\omega) \in E\}.$$

For any nonnegative Z,

$$\int_{\Omega} Z * dP = \int_{\Omega} ZTdP = \int_{\Omega \times \Omega} ZdPT^{-1} = \int_{\Omega \times \Omega} ZdP*,$$

where the second identity follows by change of variable (Billingsley, 1995, p. 216). We will

prove that

(5) $$\int_{\Omega \times \Omega} ZdP* = \int \mathbb{E}_Z dP = \int_{\Omega}\left[\int Z_{\omega} du_{\omega}\right] dP(\omega)$$

for nonnegative Z, which entails that for all such Z

$$\int_{\Omega} Z * dP = \int_{\Omega \times \Omega} ZdP* = \int \mathbb{E}_Z dP = \int_{\Omega}\left[\int Z_{\omega} du_{\omega}\right] dP(\omega) \ .$$

Following a common strategy from probability theory, we first prove (5) for indicator functions

and then build our way up to arbitrary nonnegative Z.

Take any measurable rectangle $G{\times}F$ with $G\in\mathcal{G}$ and $F\in\mathcal{F}$, and let $I_{G{\times}F}$ be the

corresponding indicator function. For this special case, (5) becomes

$$\int_{\Omega \times \Omega} I_{G{\times}F} dP* = \int_{\Omega}\left[\int (I_{G{\times}F})_{\omega} du_{\omega}\right] dP(\omega) \ .$$

For the left-hand side, note that by change of variable

$$\int_{\Omega \times \Omega} I_{G{\times}F} dP* = \int_{\Omega \times \Omega} I_{G{\times}F} dPT^{-1} = \int_{\Omega} I_{G{\times}F} TdP = \int_{\Omega} I_G I_F dP = P(G\cap F) \ .$$

For the right-hand side, note that $(I_{G{\times}F})_{\omega}$ is I_F for all $\omega\in G$ and is 0 otherwise. Thus, the right-

hand side reduces to

$$\int_{\Omega}\left[\int (I_{G{\times}F})_{\omega} du_{\omega}\right] dP(\omega) = \int_{G}\left[\int I_F du_{\omega}\right] dP(\omega) = \int_{G} \mu_{\omega}(F) dP(\omega) = P(G\cap F),$$

where the last identity follows by the integral formula. Our analysis also shows that

$$
\mathbb{E}_{I_{G\times F}}(\omega) = \begin{cases} \mu_\omega(F) & \text{if } \omega \in G \\ 0 & \text{if } \omega \notin G \end{cases},
$$

which is a well-defined $\mathcal{G}$-measurable function of ω. Thus, the lemma is established for the special case of $I_{G\times F}$.

Now consider any arbitrary indicator variable I_E, with $E \in \mathcal{G} \otimes \mathcal{F}$. Take the class M containing all $E \in \mathcal{G} \otimes \mathcal{F}$ such that

$$
\int (I_E)_\omega \, d\mu_\omega \text{ is a } \mathcal{G}\text{-measurable function of } \omega.
$$

$$
\int_{\Omega\times\Omega} I_E \, dP^* = \int_\Omega \left[\int (I_E)_\omega \, du_\omega \right] dP(\omega) .
$$

One can show that M is closed under complementation and countable disjoint union. I address countable disjoint union, leaving complementation to the reader. Suppose that E is the union of countably many sets E_n, where these sets are pairwise disjoint and where (5) holds of each indicator function I_{E_n}. Then

$$
\int_{\Omega\times\Omega} I_E \, dP^* = \int_{\Omega\times\Omega} \sum_{n=1}^\infty I_{E_n} \, dP^* = \sum_{n=1}^\infty \int_{\Omega\times\Omega} I_{E_n} \, dP^* = \sum_{n=1}^\infty \int_\Omega \left[\int (I_{E_n})_\omega \, du_\omega \right] dP(\omega)
$$

$$
= \int_\Omega \left[\int \sum_{n=1}^\infty (I_{E_n})_\omega \, du_\omega \right] dP(\omega) = \int_\Omega \left[\int (I_E)_\omega \, du_\omega \right] dP(\omega) ,
$$

where we have repeatedly used an infinite series version of the monotone convergence theorem (Billingsley, 1995, p. 211). Thus, (5) holds of I_E. To establish $\mathcal{G}$-measurability, one may write

$$
\int (I_E)_\omega \, du_\omega = \int \sum_{n=1}^\infty (I_{E_n})_\omega \, du_\omega = \sum_{n=1}^\infty \int (I_{E_n})_\omega \, du_\omega
$$

and use that the limit of $\mathcal{G}$-measurable functions is itself $\mathcal{G}$-measurable if that limit exists everywhere (Billingsley, 1995, p. 184). Since M contains all measurable rectangles $G{\times}F$ with $G{\in}\mathcal{G}$ and $F{\in}\mathcal{F}$ and is closed under complementation and countable disjoint union, it follows that M contains all members of $\mathcal{G}{\otimes}\mathcal{F}$ (Billingsley, 1995, pp. 41-42). This proves the lemma for arbitrary indicator variables I_E.

Using the linearity of integration, one can extend the lemma to any nonnegative measurable *simple function*, i.e. any finite linear combination

$$\sum_{i=1}^{n} c_i I_{E_i}$$

such that $E_i{\in}\mathcal{G}{\otimes}\mathcal{F}$, $c_i \geq 0$, and the sets E_i form a partition of $\Omega \times \Omega$. Given a nonnegative $\mathcal{G}{\otimes}\mathcal{F}$-measurable Z, there is a sequence $\{ Z_n \}$ of nonnegative measurable simple functions such that

$$Z(\omega,\nu) = \lim_{n\to\infty} Z_n(\omega,\nu),$$

and such that the sequence $\{Z_n(\omega,\nu)\}$ is non-decreasing, for each ω, ν (Billingsley, 1995, p. 185). By the monotone convergence theorem (Billingsley, 1995, p. 208),

$$\int_{\Omega\times\Omega} Z dP^* = \int_{\Omega\times\Omega} \lim_{n\to\infty} Z_n dP^* = \lim_{n\to\infty} \int_{\Omega\times\Omega} Z_n dP^* = \lim_{n\to\infty} \int_{\Omega}\left[\int_{\Omega} Z_n(\omega,\nu) u_\omega(d\nu) \right] dP(\omega) =$$

$$= \int_{\Omega}\left[\int_{\Omega} \lim_{n\to\infty} Z_n(\omega,\nu) u_\omega(d\nu) \right] dP(\omega) = \int_{\Omega}\left[\int_{\Omega} Z(\omega,\nu) u_\omega(d\nu) \right] dP(\omega),$$

which proves the lemma for arbitrary nonnegative $\mathcal{G}{\otimes}\mathcal{F}$-measurable Z.

Now fix an arbitrary $\mathcal{G}{\otimes}\mathcal{F}$–measurable function Z such that $E_P[Z^*]$ exists. We have already established that $\mathbb{E}_{Z^+}$ and $\mathbb{E}_{Z^-}$ are $\mathcal{G}$-measurable functions. We may write

$$\mathbb{E}_Z(\omega) = \int_{\Omega} Z(\omega,\nu)\mu_\omega(d\nu) = \int_{\Omega} Z(\omega,\nu)^+ \mu_\omega(d\nu) - \int_{\Omega} Z(\omega,\nu)^- \mu_\omega(d\nu) = \mathbb{E}_{Z^+}(\omega) - \mathbb{E}_{Z^-}(\omega).$$

As the difference of $\mathcal{G}$-measurable functions, $\mathbb{E}_Z = \mathbb{E}_{Z^+} - \mathbb{E}_{Z^-}$ is $\mathcal{G}$-measurable. We compute

$$E_P[Z*] = \int_\Omega Z * dP = \int_\Omega (Z*)^+ \, dP - \int_\Omega (Z*)^- \, dP = \int_\Omega (Z^+)* dP - \int_\Omega (Z^-)* dP$$

$$= \int_{\Omega\times\Omega} Z^+ dP* - \int_{\Omega\times\Omega} Z^- dP* = E_P\big[\mathbb{E}_{Z^+}\big] - E_P\big[\mathbb{E}_{Z^-}\big].$$

The final identity uses (5), which is legitimate because Z^+ and Z^- are both nonnegative. The expectation on the left-hand side exists, so at least one of the expectations on the right-hand side must be finite. Suppose without loss of generality that $E_P\big[\mathbb{E}_{Z^+}\big]$ is finite.

Choose any random variables V and W such that $E_P[V]$ is finite and $E_P[W]$ exists (possibly with infinite value). Since $E_P[V]$ is finite, V must have finite value except possibly inside a set A of P-measure 0. Define random variable $\widehat{V}$ by

$$\widehat{V}(\omega) =_{df} \begin{cases} V(\omega) & \text{if } \omega \notin A \\ 0 & \text{if } \omega \in A \end{cases}$$

Random variables that agree almost everywhere have the same expectation, so

$$E_P[V] = E_P\big[\widehat{V}\big].$$

$\widehat{V} - W$ is well-defined everywhere and agrees with $V - W$ except possibly inside A. Since $E_P\big[\widehat{V}\big] - E_P[W]$ exists, the linearity of expectations entails that

$$E_P\big[\widehat{V} - W\big] = E_P\big[\widehat{V}\big] - E_P[W].$$

$V - W$ is well-defined everywhere except possibly inside A, so it is an almost surely defined random variable. Its expectation is

$$E_P[V - W] =_{df} E_P\big[\widehat{V} - W\big] = E_P\big[\widehat{V}\big] - E_P[W] = E_P[V] - E_P[W].$$

Taking $\mathbb{E}_{Z^+}$ for V and $\mathbb{E}_{Z^-}$ for W, we conclude that $\mathbb{E}_Z = \mathbb{E}_{Z^+} - \mathbb{E}_{Z^-}$ is an almost surely defined random variable and that

$$E_P\left[\mathbf{E}_{Z^+} - \mathbf{E}_{Z^-}\right] = E_P\left[\mathbf{E}_{Z^+}\right] - E_P\left[\mathbf{E}_{Z^-}\right].$$

Thus,

$$E_P[Z*] = E_P\left[\mathbf{E}_{Z^+}\right] - E_P\left[\mathbf{E}_{Z^-}\right] = E_P\left[\mathbf{E}_{Z^+} - \mathbf{E}_{Z^-}\right] = E_P\left[\mathbf{E}_Z\right],$$

which completes the proof.   □

Inspecting the proof, we see that the integral formula is used only once and is not used in showing $\mathcal{G}$-measurability. Let C be an update rule and Y a bookie strategy. Define

$$\mu_\omega(\,.\,) =_{df} C(\,.\mid\omega\,)$$

$$\mathbf{E}_Y(\omega) =_{df} E_{\mu_\omega}[Y_\omega].$$

Our proof shows that $\mathbf{E}_Y$ is $\mathcal{G}$-measurable, whether or not C satisfies the integral formula. If we define

$$\Delta =_{df} \{\,\omega\colon Y_\omega \text{ is acceptable given } \omega\} = \{\,\omega\colon \mathbf{E}_Y(\omega) \geq 0\,\},$$

then the $\mathcal{G}$-measurability of $\mathbf{E}_Y$ entails that $\Delta \in \mathcal{G}$.

We can now show that strong Dutch books are weak Dutch books. Let (X, Y) be a strong Dutch book for $(\Omega, \mathcal{F}, P)$, $\mathcal{G}$, and C. Conditions (a) and (b) in the definition of weak Dutch book are immediate. Defining Δ as in the previous paragraph, note that

(6) $\omega \in \Delta \Rightarrow X(\omega) + Y_\omega(\omega) < 0$

(7) $\omega \notin \Delta \Rightarrow X(\omega) < 0$

Condition (d) in the definition of weak Dutch book follows from (6), while condition (e) follows from (7). For condition (c), note that $P(\Delta)$ is well-defined since $\Delta \in \mathcal{G}$. If $P(\Delta) = 0$, then

$$\int_\Omega X dP = \int_{\Delta^c} X dP < 0,$$

contradicting our assumption that X is acceptable. We must therefore have $P(\Delta) > 0$, which

together with (6) entails condition (c).

Acknowledgments

Works Cited

Arntzenius, F. 2003. "Some Problems for Conditionalization and Reflection." *Journal of Philosophy* 100: 356–70.

Arntzenius, F., Elga, A., and Hawthorne, J. 2004. "Bayesianism, Infinite Decisions, and Binding." *Mind* 113:251–283.

Bayes, T, and Price, R. 1763. "An Essay Towards Solving a Problem in the Doctrine of Chances. By the Late Rev. Mr. Bayes, F.R.S. Communicated by Mr. Price, in a Letter to John Canton, A.M.F.R.S." *Philosophical Transactions of the Royal Society of London* 53 : 470-418.

Bertrand, J. 1889. *Calculs des Probabilités*. Paris: Gauthier-Villars.

Billingsley, P. 1995. *Probability and Measure*. 3rd ed. New York: Wiley.

Borel, E. 1909/1956. *Elements of the Theory of Probability*. Trans. J. Freund. Englewood Cliffs: Prentice-Hall.

Briggs, R. 2009. "Distorted Reflection." *Philosophical Review* 118: 59-85.

Chater, N., and Oaksford, C. 2008. *The Probabilistic Mind*. Oxford: Oxford University Press.

Christensen, David. 1991. "Clever Bookies and Coherent Beliefs." *Philosophical Review* 100: 229–47.

de Finetti, B. 1937/1980. "Foresight. Its Logical Laws, Its Subjective Sources." Rpt. in *Studies in Subjective Probability*, eds. H. E. Kyburg, Jr. and H. E. Smokler. Huntington: Robert E. Krieger.

de Finetti, B. 1972. *Probability, Induction, and Statistics*. New York: Wiley.

Davidson, D. 2004. *Problems of Rationality*. Oxford: Oxford University Press.

Durrett, R. 1991. *Probability: Theory and Examples*. Pacific Grove: Wadsworth.

Earman, J. 1992. *Bayes or Bust?*. Cambridge: MIT Press.

Easwaran, K. 2008. *The Foundations of Conditional Probability*. PhD dissertation, University of California, Berkeley. Ann Arbor: ProQuest/UMI. (Publication No. 3331592.)

---. 2011. "Varieties of Conditional Probability." In *Philosophy of Statistics*, eds. P. Bandyopadhyay and M. Forster. Burlington: Elsevier.

---. 2013. "Expected Accuracy Supports Conditionalization --- and Conglomerability and

Reflection." *Philosophy of Science* 80: 119-142.

Fristedt, B., and Gray, L. 1997. *A Modern Approach to Probability Theory*. Boston: Birkhäuser.

Fudenberg, D., and Tirole, J. 1991. *Game Theory*. Cambridge: MIT Press.

Gyenis, Z., Hofer-Szabó, G., and Rédei, M. 2017. "Conditioning Using Conditional Expectations: The Borel-Kolmogorov Paradox." *Synthese* 194: 2595-2630.

Gyenis, Z., and Rédei, M. 2017. "General Properties of Bayesian Learning as Statistical Inference Determined by Conditional Expectations." *The Review of Symbolic Logic*. doi:10.1017/S1755020316000502.

Hájek, A. 2003. "What Conditional Probability Could Not Be." *Synthese* 137: 273-323.

---. 2009. "Dutch Book Arguments." In *The Handbook of Rationality and Social Choice*, eds. P. Anand, P. Pattanaik, and C. Puppe. Oxford: Oxford University Press.

---. 2011. "Conditional Probability." In *Philosophy of Statistics*, eds. P. Bandyopadhyay and M. Forster. Burlington: Elsevier.

---. 2012. "Is Strict Coherence Coherent?". *dialectica* 66: 411-424.

Hill, B. 1980. "On Some Statistical Paradoxes and Non-conglomerability." *Trabajos de Estadistica de Investigacion Operativa* 31: 39-66.

Howson, C. 2014. "Finite Additivity, Another Lottery Paradox, and Conditionalization." *Synthese* 191: 989-1012.

Huttegger, S. 2015. "Merging of Opinions and Probability Kinematics." *The Review of Symbolic Logic* 8: 611–648.

Jeffrey, R. 1983. *The Logic of Decision*. 2nd edition. Chicago: University of Chicago Press.

---. 1992. *Probability and the Art of Judgment*. Cambridge: Cambridge University Press.

Kadane, J., Schervish, M., and Seidenfeld, T. 1986. "Statistical Implications of Finitely Additive Probability." In *Bayesian Inference and Decision Techniques*, eds. P. Goel and A. Zellner. Amsterdam: North-Holland.

Kemeny, J. 1955. "Fair Bets and Inductive Probabilities." *Journal of Symbolic Logic* 20: 263–273.

Knill, D., and Richards, W., eds. 1996. *Perception as Bayesian Inference*. Cambridge: Cambridge University Press.

Kolmogorov, A. N. 1933/1956. *Foundations of the Theory of Probability*. 2nd English ed. Trans. N. Morrison. New York: Chelsea.

Levi, Isaac. 1987. "The Demons of Decision." *Monist* 70: 193–211.

Lewis, D. 1999. "Why Conditionalize?". In *Papers in Metaphysics and Epistemology*. Cambridge: Cambridge University Press.

Lindley, D. 1982. "The Bayesian Approach to Statistics." *Some Recent Advances in Statistics*, eds. J. T. de Oliveria and B. Epstein. New York: Academic Press.

McGee, V. 1994. "Learning the Impossible." In *Probability and Conditionals*, eds. E. Eells and B. Skyrms. Cambridge: Cambridge University Press.

---. 1999. "An Airtight Dutch Book." *Analysis* 59: 257-265.

Madl, T., Franklin, S., Chen, K., Montaldi, D., and Trappl, R. 2014. "Bayesian Integration of Information in Hippocampal Place Cells." *PloS One* 9: e89762.

Maher, P. 1992. "Diachronic Rationality." *Philosophy of Science* 59: 120-141.

Mahtani, A. 2015. "Dutch Books, Coherence, and Logical Consistency." *Noûs* 49: 522-537.

Myrvold, W. 2015. "You Can't Always Get What You Want: Some Considerations Regarding Conditional Probabilities." *Erkenntnis* 80: pp. 573-603.

Pfanzagl, J. 1979. "Conditional Distributions as Derivatives." *Annals of Probability* 7:

1046-1050.

Popper, K. 1959. *The Logic of Scientific Discovery*. London: Hutchinson.

Ramsey, F. P. 1931. "Truth and Probability." In *The Foundations of Mathematics and Other Logical Essays*, ed. R. B. Braithwaite. London: Routledge and Kegan.

Rao, M. M. 2005. *Conditional Measures and Applications*. 2nd ed. Boca Raton: CRC Press.

Rényi, A. 1955. "On a New Axiomatic Theory of Probability." *Acta Mathematica Academiae Scientiarum Hungarica* 6: 285-335.

Rescorla, M. 2015. "Some Epistemological Ramifications of the Borel-Kolmogorov Paradox." *Synthese* 192: 735-767.

Savage, L.J. 1954. *The Foundations of Statistics*. New York: Wiley.

Seidenfeld, T. 2001. "Remarks on the Theory of Conditional Probability: Some Issues of Finite versus Countable Additivity." In *Probability Theory: Philosophy, Recent History, and Relations to Science*, eds. V. Hendricks, S. Pedersen, and K. Jørgensen. Dordrecht: Kluwer.

Seidenfeld, T., Schervish, M., and Kadane, J. 2001. "Improper Regular Conditional Distributions." *Annals of Probability* 29: 1612-1624.

Shimony, A. 1955. "Coherence and the Axioms of Confirmation." *Journal of Symbolic Logic* 20: 1-28.

Skyrms, B. 1980. *Causal Necessity*. New Haven: Yale University Press.

---. 1987. "Dynamic Coherence and Probability Kinematics." *Philosophy of Science* 54: 1–20.

---. 1992. "Coherence, Probability and Induction." *Philosophical Issues* 2:215-226.

---. 1993. "A Mistake in Dynamic Coherence Arguments?". *Philosophy of Science* 60: 320-328.

Stalnaker, R. 1970. "Probabilities and Conditionals." *Philosophy of Science* 37: 64–80.

Teller, P. 1973. "Conditionalization and Observation." *Synthese* 26: 218–258.

Thrun, S., Burgard, W., and Fox, D. 2006. *Probabilistic Robotics*. Cambridge: MIT Press.

van Fraassen, B. 1984. "Belief and the Will." *Journal of Philosophy* 81: 235–56.

Outcome	Price	Payoff	Net gain
$\omega \in A \cap G$	$P_{\mathfrak{G}}(A \mid \omega)$	1	$1 - P_{\mathfrak{G}}(A \mid \omega)$
$\omega \in A^c \cap G$	$P_{\mathfrak{G}}(A \mid \omega)$	0	$- P_{\mathfrak{G}}(A \mid \omega)$
$\omega \notin G$	0	0	0

Table 1. Net gain for bet 1.

Outcome	Price	Payoff	Net gain
$\omega \in A$	$C(A \mid \omega)$	1	$C(A \mid \omega) - 1$
$\omega \in A^c$	$C(A \mid \omega)$	0	$C(A \mid \omega)$

*Table 2*. Net gain for bet 2. Note that we subtract the payoff from the price, since you sell rather than buy the bet.

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

Table 3. Net gain for the entire gambling scenario, found by adding net gains for bet 1 and bet 2. Bear in mind that bet 2 is only offered when $\omega \in G$.
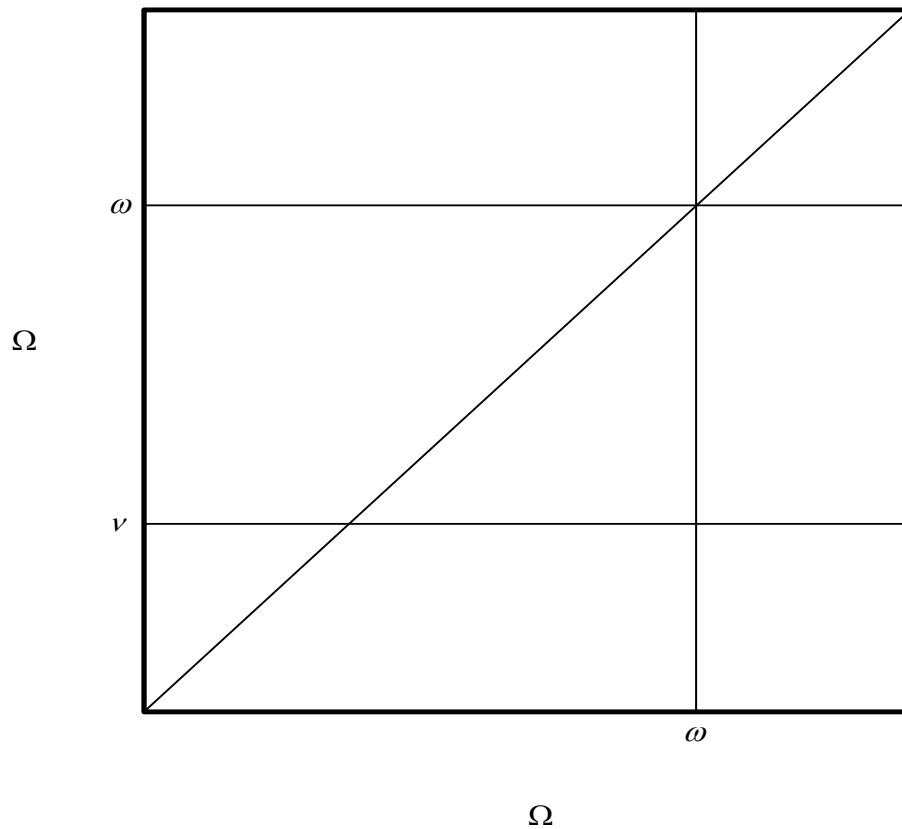
Figure 1. A visualization of $\Omega \times \Omega$. There may not be a natural linear ordering of Ω, but the visualization is still a useful heuristic. The horizontal axis corresponds to $(\Omega, \mathcal{G})$. Points on this axis determine which bet the bookie selects. The vertical axis corresponds to $(\Omega, \mathcal{F})$. Points on this axis determine net gain from whatever bet the bookie selects. When the bookie acquires information about outcome ω, he offers bet Y_ω. For any outcome v, this bet has a well-defined net gain $Y_\omega(v)$, i.e. Y's value on the point where the vertical line intersects the lower horizontal line. In actuality, the outcome ω that determines the bookie's bet is the same outcome ω that determines net gain for that bet. Someone who accepts the bet receives net gain $Y^*(\omega) =_{df} Y_\omega(\omega)$, which is Y's value at the point where the vertical line intersects the diagonal line.
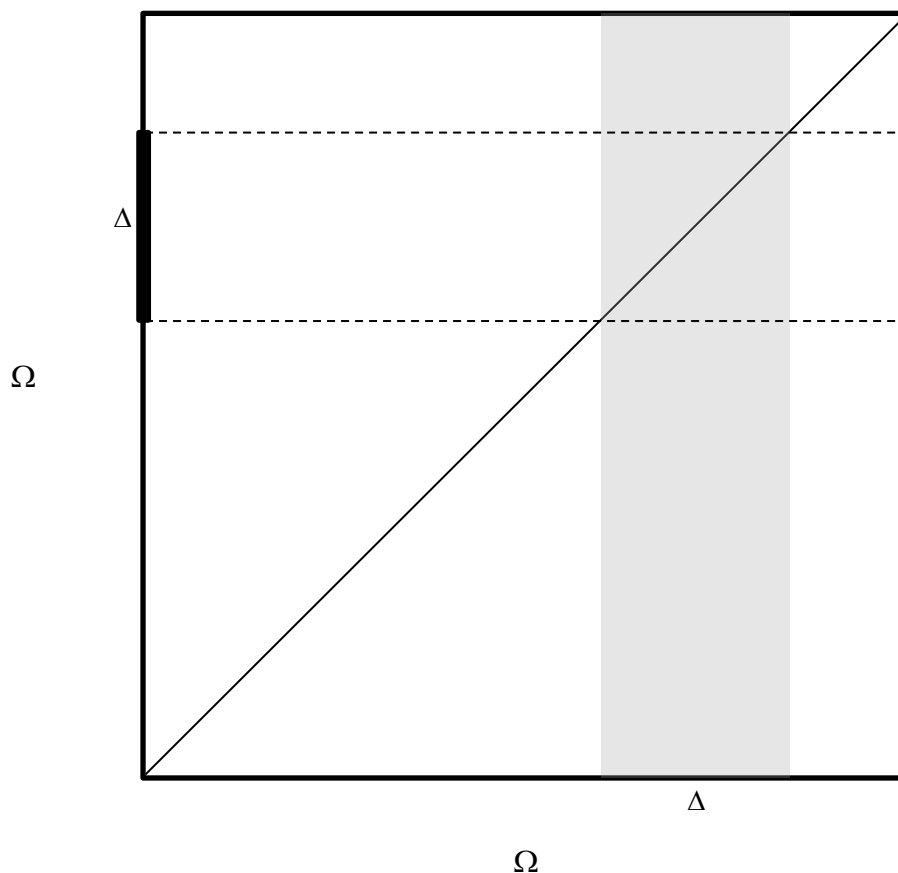
Figure 2. Think of *X* and *Y**** as bets on events over the vertical axis. Let $\Delta =_{df} \{ \omega: Y_\omega$ is acceptable given $\omega\}$. The grey column contains points (ω, *v*) with $\omega \in \Delta$. If (*X*, *Y*) is a strong Dutch book, then $X + Y^* < 0$ inside Δ and $X < 0$ outside Δ. If (*X*, *Y*) is a weak Dutch book, then $X + Y^* < 0$ with positive probability inside Δ, $X + Y^* > 0$ with probability 0 inside Δ, and $X > 0$ with probability 0 outside Δ. Thus, the only values of *Y* that affect whether (*X*, *Y*) is a strong (or weak) Dutch book are its values on points where the diagonal line intersects the grey column. This reflects the fact that our betting agent rejects Y_ω if $\omega \notin \Delta$.
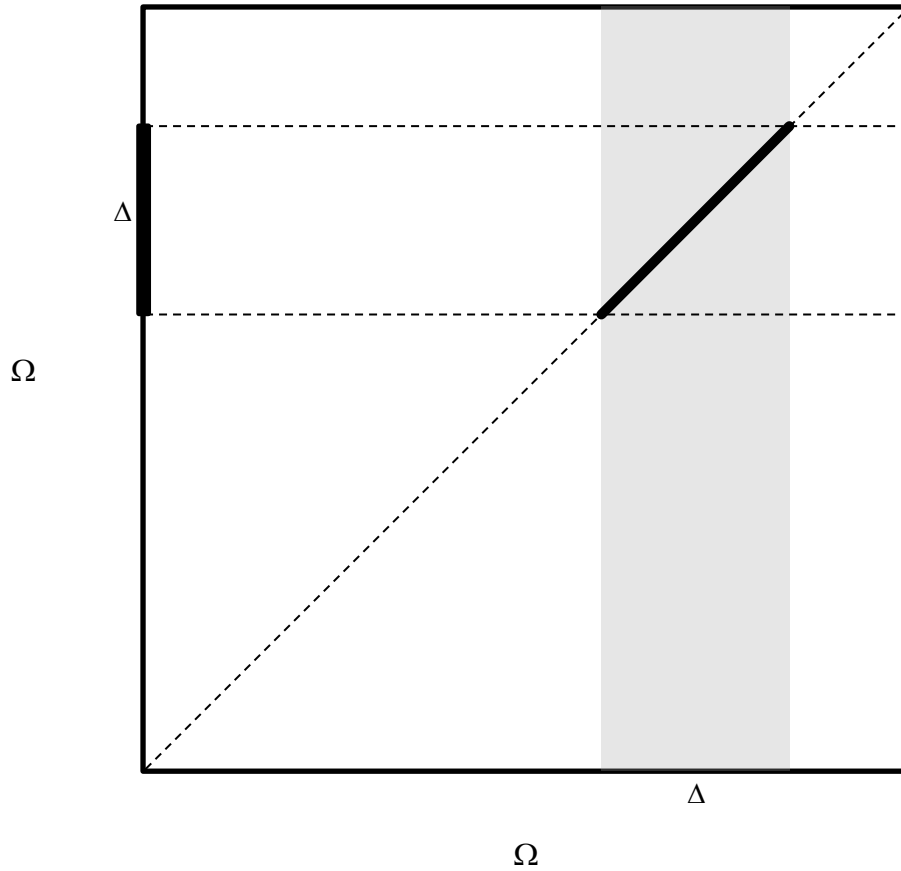
Figure 3. Think of Z^* as a bet on events over the vertical axis. If $\omega \in \Delta$, then $Z^*(\omega)$ is Y's value on the corresponding point of the diagonal line. If $\omega \notin \Delta$, then $Z^*(\omega) = 0$. Informally, Z^* converts bookie strategy Y into a single bet that prunes away all possibilities for expected net loss.