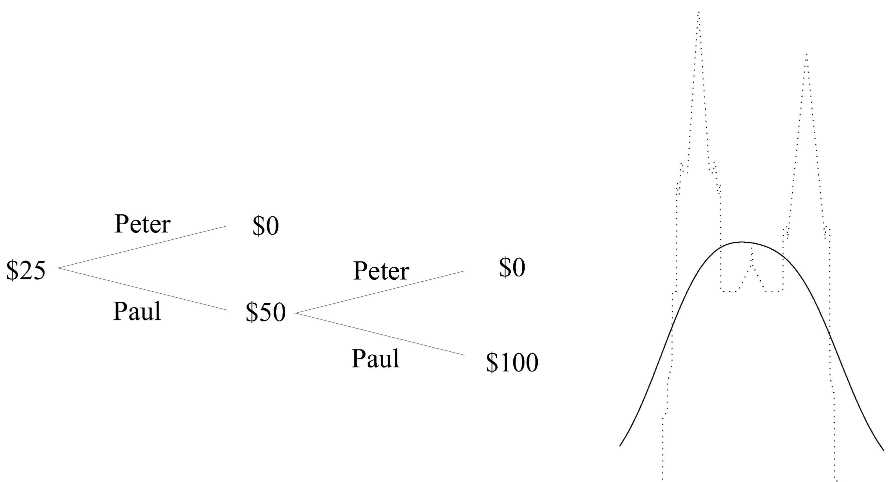
Convergence of opinions
(work in progress)

Vladimir Vovk



The Game-Theoretic Probability and Finance Project

Working Paper #66

First posted December 4, 2023. Last revised December 22, 2023.

Project web site:
http://www.probabilityandfinance.com

Abstract

This paper establishes a game-theoretic version of the classical Blackwell–Dubins result. We consider two forecasters who at each step issue probability forecasts for the infinite future. Our result says that either at least one of the two forecasters will be discredited or their forecasts will converge in total variation.

This paper has also been published as an arXiv report.

Contents

. . . scientific disagreements tend to disappear. . . when new data accumulate. . . .

Harold Jeffreys, 1938 [9, p. 673]; also in [10, Sect. 1.9]

1 Introduction

Blackwell and Dubins [1] prove a classical result about convergence of opinions. However, the standard measure-theoretic form of the Blackwell–Dubins result is restrictive, e.g., because it considers (in its symmetric form) two forecasters who agree on the events of probability zero. What if they don't? For example, suppose that they agree perfectly everywhere apart from two events, E^{I} and E^{II}, for which

$$P^{\mathrm{I}}(E^{\mathrm{I}}) = P^{\mathrm{II}}(E^{\mathrm{II}}) = 10^{-6}, \quad P^{\mathrm{II}}(E^{\mathrm{I}}) = P^{\mathrm{I}}(E^{\mathrm{II}}) = 0.$$

For this case the Blackwell–Dubins result does not say anything. To prevent the possibility of waiting until one or both events become settled, suppose further that, for all n,

$$P^{\mathrm{I}}(E^{\mathrm{I}} \mid \mathcal{F}_n), P^{\mathrm{II}}(E^{\mathrm{I}} \mid \mathcal{F}_n), P^{\mathrm{I}}(E^{\mathrm{II}} \mid \mathcal{F}_n), P^{\mathrm{II}}(E^{\mathrm{II}} \mid \mathcal{F}_n) \in (0,1).$$

Another respect in which the Blackwell–Dubins result is restrictive is that it depends on Bayesian conditioning: the two forecasters revise their beliefs by conditioning on the observed data. This paper uses, instead, a more general game-theoretic picture proposed in [18]. In the simple finite case considered in [18, Sect. 4] the two pictures are equivalent, and we can replace the game-theoretic picture by Bayesian "superconditioning" (Jeffrey, 1988, [5, Sect. 7]), but it is likely that in more general cases the equivalence is lost.

The game-theoretic version proved in Sect. 2 of this paper removes some of the unnecessary restrictions, and it is also more constructive. It shows that at least one of the forecasters can be discredited by successful betting against his forecasts unless the predictions that the forecasters output converge in total variation.

A third respect in which the Blackwell–Dubins result is restrictive is that it assumes that we observe some data with certainty. Section 4 extends the result of Sect. 2 to Jeffrey's picture of radical probabilism, in which no evidence is certain.

In this paper, the general phenomenon of the convergence of opinions for adequate forecasters will be referred to as *Jeffreys's law*, although originally this expression was used by Dawid [3, Sect. 5.2] for his one-step-ahead result about convergence of opinions:

> I shall call this finding "Jeffreys's Law", after an admittedly distorted interpretation of Jeffreys (1938): "When a law has been applied to a large body of data without any systematic discrepancy

Figure 1: Left panel: Harold Jeffreys (1891–1989). Right panel: Richard Jeffrey (1926–2002).

> being detected... the probability of a further inference from the law approaches certainty whether the law is true or not."

Another quote from the same paper is given as the epigraph to this paper.

This paper and its predecessor [18] were motivated by a conversation with A. Philip Dawid [19]. Among topics of the conversation were one-step-ahead vs multi-step predictions [19, Sect. 7], the Blackwell–Dubins theorem [19, Sect. 7 of the arXiv version], and data generating processes [19, Sect. 7] (cf. Remark 2.4 below).

Two people with similar surnames will play key roles in this paper, Jeffreys and Jeffrey; see Fig. 1.

2 Jeffreys's law

Let us fix a finite *observation space* (equipped with the discrete σ-algebra). Our prediction/testing protocol involves one Sceptic playing separately against two Forecasters, who output probability forecasts $P^{\mathrm{I}}_n, P^{\mathrm{II}}_n \in \mathfrak{P}(\mathbf{Y}^\infty)$ ($\mathfrak{P}(\mathbf{Y}^\infty)$ standing for the family of all probability measures on $\mathbf{Y}^\infty$) at steps $n = 1, 2, \dots$.

Protocol 2.1. Basic additive protocol:

$\mathcal{K}^{\mathrm{I}}_0 = \mathcal{K}^{\mathrm{II}}_0 := 1$
FOR $n = 1, 2, \dots$:
  Forecaster I announces $P^{\mathrm{I}}_n \in \mathfrak{P}(\mathbf{Y}^\infty)$
  Forecaster II announces $P^{\mathrm{II}}_n \in \mathfrak{P}(\mathbf{Y}^\infty)$
  IF $n > 1$:
    $\mathcal{K}^{\mathrm{I}}_{n-1} := \mathcal{K}^{\text{-I}}_{n-1} + \sum_{x\in\mathbf{Y}^*} f^{\mathrm{I}}_{n-1}(y_{n-1}x) P^{\mathrm{I}}_n(x) - \sum_{x\in\mathbf{Y}^*} f^{\mathrm{I}}_{n-1}(x) P^{\mathrm{I}}_{n-1}(x)$
    $\mathcal{K}^{\mathrm{II}}_{n-1} := \mathcal{K}^{\text{-II}}_{n-1} + \sum_{x\in\mathbf{Y}^*} f^{\mathrm{II}}_{n-1}(y_{n-1}x) P^{\mathrm{II}}_n(x) - \sum_{x\in\mathbf{Y}^*} f^{\mathrm{II}}_{n-1}(x) P^{\mathrm{II}}_{n-1}(x)$
  Sceptic announces $f^{\mathrm{I}}_n, f^{\mathrm{II}}_n \in \mathbb{R}^{\mathbf{Y}^*}$ such that
    $f^{\mathrm{I}}_n(x) = f^{\mathrm{II}}_n(x) = 0$ for all but finitely many $x \in \mathbf{Y}^*$
  Reality announces $y_n \in \mathbf{Y}$
  $\mathcal{K}^{\text{-I}}_n := \mathcal{K}^{\mathrm{I}}_{n-1} + f^{\mathrm{I}}_n(y_n) - \sum_{y\in\mathbf{Y}} f^{\mathrm{I}}_n(y) P^{\mathrm{I}}_n(y)$
  $\mathcal{K}^{\text{-II}}_n := \mathcal{K}^{\mathrm{II}}_{n-1} + f^{\mathrm{II}}_n(y_n) - \sum_{y\in\mathbf{Y}} f^{\mathrm{II}}_n(y) P^{\mathrm{II}}_n(y)$

Protocol 2.1 is interpreted in terms of trading in futures contracts, as described in [18, Sect. 3]. At each step Sceptic announces positions in the futures contracts that are implicit in P^{I}_n and P^{II}_n. The positions are allowed to be different from zero only for finitely many futures contracts, and so the sums $\sum_{x\in\mathbf{Y}^*}$ are uncontroversial. At the end of each step Reality announces the actual observation $y_n \in \mathbf{Y}$, and the futures contracts are settled. The adjective "additive" in the title of the protocol will be explained in Sect. 4.

The following result (game-theoretic counterpart of the Blackwell–Dubins theorem) uses the total variation distance

$$\|P - Q\| := 2 \sup_E |P(E) - Q(E)| \in [0, 2]$$

between probability measures on $\mathbf{Y}^\infty$ (equipped with the Borel σ-algebra), where E ranges over the events in the common domain of P and Q.

Theorem 2.2. *Sceptic has a strategy in Protocol 2.1 that guarantees the disjunction of*

- $\|P^{\mathrm{I}}_n - P^{\mathrm{II}}_n\| \to 0$ *as* $n \to \infty$,
- $\mathcal{K}^{\mathrm{I}}_n \to \infty$ *as* $n \to \infty$,
- $\mathcal{K}^{\mathrm{II}}_n \to \infty$ *as* $n \to \infty$.

Remark 2.3. Theorem 2.2 can be interpreted as establishing a connection between the correspondence (see, e.g., [14]) and convergence (see, e.g., [13]) theories of truth. In Peirce's words, "the settlement of opinion is the sole end of inquiry" [13]. Namely, according to Theorem 2.2, a version of the correspondence theory implies a version of the convergence theory.

Remark 2.4. According to Dawid [19, Sect. 7], much of statistics works with the data generating process. A possible interpretation of Theorem 2.2 is that the data generating process arises in the limit.

3 Proof of Theorem 2.2

We will identify a probability measure P on $\mathbf{Y}^\infty$ with a function mapping $x \in \mathbf{Y}^*$ to $P([x])$, $[x] \subset \mathbf{Y}^\infty$ being the set of the infinite sequences in $\mathbf{Y}^\infty$ that have x as their prefix. For simplicity we assume that $P^{\mathrm{I}}_n(x) > 0$ and $P^{\mathrm{II}}_n(x) > 0$ for all $x \in \mathbf{Y}^*$ ("Cromwell's rule").

In this proof we will construct a strategy for Sceptic that guarantees the disjunction of

- $\|P^{\mathrm{I}}_n - P^{\mathrm{II}}_n\| \to 0$ as $n \to \infty$,
- $\sqrt{\mathcal{K}^{\mathrm{I}}_n \mathcal{K}^{\mathrm{II}}_n}$ is unbounded as $n \to \infty$.

Algorithm 1 Gambling for a fixed $\epsilon \in (0,1)$

for $n = 1, 2 \ldots$:
  Observe P_n^{I} and P_n^{II} on step n of Protocol 2.1
  if $H(P_n^{\mathrm{I}}, P_n^{\mathrm{II}}) < 1 - \epsilon$:
    Find m such that $H_m(P_n^{\mathrm{I}}, P_n^{\mathrm{II}}) < 1 - \epsilon$
    Buy a forward contract expiring in m steps from Forecaster I that multiplies the current $\mathcal{K}_n^{\mathrm{I}}$ by $\dfrac{1}{H_m(P_n^{\mathrm{I}}, P_n^{\mathrm{II}})} \dfrac{\mathrm{d}\sqrt{P_n^{\mathrm{I}}|_m P_n^{\mathrm{II}}|_m}}{\mathrm{d}P_n^{\mathrm{I}}|_m}$
    Buy a forward contract expiring in m steps from Forecaster II that multiplies the current $\mathcal{K}_n^{\mathrm{II}}$ by $\dfrac{1}{H_m(P_n^{\mathrm{I}}, P_n^{\mathrm{II}})} \dfrac{\mathrm{d}\sqrt{P_n^{\mathrm{I}}|_m P_n^{\mathrm{II}}|_m}}{\mathrm{d}P_n^{\mathrm{II}}|_m}$
    Skip the next m steps

This is sufficient since we can apply Proposition 11.2 in [15] to turn an unbounded capital $\mathcal{K}_n^{\mathrm{I}}$ or $\mathcal{K}_n^{\mathrm{II}}$ into $\mathcal{K}_n^{\mathrm{I}} \to \infty$ or $\mathcal{K}_n^{\mathrm{II}} \to \infty$.

As a second step, let us replace $\|P_n^{\mathrm{I}} - P_n^{\mathrm{II}}\| \to 0$ by $H(P_n^{\mathrm{I}}, P_n^{\mathrm{II}}) \to 1$, where H is the *Hellinger integral* [16, Definition 3.9.3]:

$$H(P,Q) := \int_{\mathbf{Y}^\infty} \sqrt{\mathrm{d}P \mathrm{d}Q} \in [0,1].$$

This can be done because of the standard connection between variation distance and Hellinger integral [16, Theorem 3.9.1]:

$$2\,(1 - H(P,Q)) \le \|P - Q\| \le \sqrt{8\,(1 - H(P,Q))}.$$

We will also need an approximation to $H(P,Q)$; namely, set

$$H_m(P,Q) := \int_{\mathbf{Y}^m} \sqrt{\mathrm{d}P|_m \mathrm{d}Q|_m},$$

where $P|_m$ and $Q|_m$ stand for the restrictions of P and Q to the first m observations.

Lemma 3.1. *As $m \to \infty$, $H_m(P,Q) \to H(P,Q)$.*

Proof. Applying Doob's martingale convergence theorem [17, Theorem 7.4.1] to $\mathrm{d}P|_m/\mathrm{d}R|_m$ and $\mathrm{d}Q|_m/\mathrm{d}R|_m$, where $R := (P+Q)/2$, we obtain their R-almost sure convergence to $\mathrm{d}P/\mathrm{d}R$ and $\mathrm{d}Q/\mathrm{d}R$, respectively (this step uses Carathéodory's theorem, as in the proof of [17, Theorem 7.6.1]). Therefore,

$$\sqrt{\frac{\mathrm{d}P|_m}{\mathrm{d}R|_m} \frac{\mathrm{d}Q|_m}{\mathrm{d}R|_m}} \to \sqrt{\frac{\mathrm{d}P}{\mathrm{d}R} \frac{\mathrm{d}Q}{\mathrm{d}R}}$$

almost surely and, since all these fractions take values in $[0,2]$, in L_1. □

Fix temporarily an $\epsilon > 0$ (we will later mix over a sequence of $\epsilon \to 0$). For each ϵ we gamble at the steps $n = 1, 2, \dots$ as in Algorithm 1. When the condition in line 3 is satisfied, the geometric mean $\sqrt{\mathcal{K}_n^{\mathrm{I}}\mathcal{K}_n^{\mathrm{II}}}$ of Sceptic's capitals over the m steps in line 7 will be multiplied by at least $\frac{1}{1-\epsilon}$. Therefore, $\sqrt{\mathcal{K}_n^{\mathrm{I}}\mathcal{K}_n^{\mathrm{II}}}$ will be unbounded if the condition in line 3 is satisfied infinitely often.

Let $\mathcal{K}_n^{\mathrm{I}}(\epsilon)$ and $\mathcal{K}_n^{\mathrm{II}}(\epsilon)$ be the capital processes that result from using Algorithm 1 with parameter ϵ. Then

$$\mathcal{K}_n^{\mathrm{I}} := \sum_{j=1}^{\infty} 2^{-j}\mathcal{K}_n^{\mathrm{I}}(2^{-j}) \quad \text{and} \quad \mathcal{K}_n^{\mathrm{II}} := \sum_{j=1}^{\infty} 2^{-j}\mathcal{K}_n^{\mathrm{II}}(2^{-j})$$

are also capital processes for a valid strategy for Sceptic, and $\sqrt{\mathcal{K}_n^{\mathrm{I}}\mathcal{K}_n^{\mathrm{II}}}$ is unbounded unless $H(P_n^{\mathrm{I}}, P_n^{\mathrm{II}}) \to 1$. This completes the proof.

4 Agnostic probabilism

According the idea of radical probabilism, put forward very clearly by Jeffrey [6], empirical evidence is never certain. In particular, we never learn the outcomes $y_1, y_2, \dots$ for sure. Radical probabilism can be regarded as extension of Cromwell's rule, which is an assumption about the synchronic picture at each step n, to the diachronic picture.

Jeffrey referred to the opposite point of view, in which we do observe the true y_n (in our current context), as dogmatic probabilism. We would like our mathematical results to cover it as a special case, and the title of this section, *agnostic probabilism*, refers to its prediction protocols eventually disclosing, or never disclosing, the true outcomes y_n, as the case may be for different n. In particular, the following protocol includes Protocol 2.1 as a special case.

Protocol 4.1. Agnostic additive protocol:

$\mathcal{K}_0^{\mathrm{I}} = \mathcal{K}_0^{\mathrm{II}} := 1$
FOR $n = 1, 2, \dots$:
  Forecaster I announces $P_n^{\mathrm{I}} \in \mathfrak{P}(\mathbf{Y}^\infty)$
  Forecaster II announces $P_n^{\mathrm{II}} \in \mathfrak{P}(\mathbf{Y}^\infty)$
  IF $n > 1$:

$$\begin{aligned} \mathcal{K}_{n-1}^{\mathrm{I}} &:= \mathcal{K}_{n-2}^{\mathrm{I}} + \textstyle\sum_{x\in\mathbf{Y}^*} f_{n-1}^{\mathrm{I}}(x)(P_n^{\mathrm{I}}(x) - P_{n-1}^{\mathrm{I}}(x)) \\ \mathcal{K}_{n-1}^{\mathrm{II}} &:= \mathcal{K}_{n-2}^{\mathrm{II}} + \textstyle\sum_{x\in\mathbf{Y}^*} f_{n-1}^{\mathrm{II}}(x)(P_n^{\mathrm{II}}(x) - P_{n-1}^{\mathrm{II}}(x)) \end{aligned} \qquad (1)$$

  Sceptic announces $f_n^{\mathrm{I}}, f_n^{\mathrm{II}} \in \mathbb{R}^{\mathbf{Y}^*}$ such that
    $f_n^{\mathrm{I}}(x) = f_n^{\mathrm{II}}(x) = 0$ for all but finitely many $x \in \mathbf{Y}^*$

Protocol 4.1 does not include Reality as a separate player. Her role is played by the two Forecasters: in order to model Reality, they should choose P_n^{I} and P_n^{II} concentrated on $[(y_1, \dots, y_n)]$ for some observations $y_1, \dots, y_n \in \mathbf{Y}$. Such a choice already implies a certain agreement between the Forecasters; in particular, $P_n^{\mathrm{I}} - P_n^{\mathrm{II}} \to 0$ as $n \to \infty$ *weakly*, meaning that, for any $x \in \mathbf{Y}^*$,

$$P_n^{\mathrm{I}}([x]) - P_n^{\mathrm{II}}([x]) \to 0 \quad \text{as } n \to \infty.$$

Remark 4.2. Jeffrey traces the idea of radical probabilism back to Ramsey and de Finetti. In [8, p. 66], he says, "this is his radical probabilism—Ramsey denies that our probable knowledge need be based on certainties". And in [7, p. 2], he says, "De Finetti's probabilism is 'radical' in the sense of going all the way down to the roots: he sees probabilities as ultimate forms of judgment which need not be based on deeper all-or-none knowledge."

Remark 4.3. Jeffrey often uses "radical probabilism" in the sense of our "agnostic probabilism": "Radical probabilism doesn't insist that probabilities be based on certainties" [8, p. 11]. However, in other places he appears to deny the existence of certain empirical evidence; e.g., in one of his earliest (1968) publications on radical probabilism he says, "Radical probabilism adds [to "probabilism"] the "nonfoundational" thought that there is no bedrock of certainty underlying our probabilistic judgments" [8, pp. 44–45]. One of the reasons I am using my terminology, even if it may distort Jeffrey's meaning, is that I do not see anything radical in failing to observe a few of y_n.

Example 4.4. Let us check that Theorem 2.2 does not hold under radical probabilism (without any further conditions). Suppose $P_n^{\mathrm{I}} = P^{\mathrm{I}}$ and $P_n^{\mathrm{II}} = P^{\mathrm{II}}$ do not depend on n. Then we have $\mathcal{K}_n^{\mathrm{I}} = \mathcal{K}_n^{\mathrm{II}} = 1$ for all n, and so we have no convergence of opinion for successful Forecasters even if P^{I} and P^{II} are very different.

Theorem 4.5. *Sceptic has a strategy in Protocol 4.1 that guarantees* $\|P_n^{\mathrm{I}} - P_n^{\mathrm{II}}\| \to 0$ *as* $n \to \infty$ *whenever the following three conditions are satisfied:*

- $P_n^{\mathrm{I}} - P_n^{\mathrm{II}} \to 0$ *weakly as* $n \to \infty$,
- $\mathcal{K}_n^{\mathrm{I}} \not\to \infty$ *as* $n \to \infty$, *and*
- $\mathcal{K}_n^{\mathrm{II}} \not\to \infty$ *as* $n \to \infty$.

Theorem 4.5 includes Theorem 2.2 as a special case. According to Example 4.4, we need to impose some conditions of agreement between the two Forecasters before we can claim that they agree in the strong sense of convergence in total variation. In Theorem 2.2 both Forecasters agree with Reality (and therefore, between themselves): $P_n^{\mathrm{I}}([x])$ and $P_n^{\mathrm{II}}([x])$ become equal (and both equal to 0 or to 1) as soon as n exceeds the length of x. In Theorem 4.5 we require a much weaker agreement between Forecasters (and do not require any agreement with Reality, who is not even a player).

Remark 4.6. In Remark 2.3 we interpreted Jeffreys's law in the form of Theorem 2.2 as establishing a connection between the correspondence and convergence theories of truth. There is no absolute notion of truth under radical probabilism, and so this interpretation is not applicable to Theorem 4.5. Theorem 4.5 merely provides a means of boosting agreement between successful Forecasters: agreement in the sense of weak convergence implies agreement in the sense of convergence in total variation.

Two simplified pictures of agnostic probabilism

The following is a simplified version of Protocol 4.1, where we consider a finite sample space $\mathbf{X}$ instead of $\mathbf{Y}^\infty$ (as in [18]). In the case of agnostic probabilism, we can have a play of infinite duration even for a finite $\mathbf{X}$.

Protocol 4.7. Agnostic additive protocol:

$\mathcal{K}_0^{\mathrm{I}} = \mathcal{K}_0^{\mathrm{II}} := 1$
FOR $n = 1, 2, \dots$:
  Forecaster I announces $P_n^{\mathrm{I}} \in \mathfrak{P}(\mathbf{X})$
  Forecaster II announces $P_n^{\mathrm{II}} \in \mathfrak{P}(\mathbf{X}))$
  IF $n > 1$:
    $\mathcal{K}_{n-1}^{\mathrm{I}} := \mathcal{K}_{n-2}^{\mathrm{I}} + \sum_{x\in\mathbf{X}} f_{n-1}^{\mathrm{I}}(x)(P_n^{\mathrm{I}}(x) - P_{n-1}^{\mathrm{I}}(x))$
    $\mathcal{K}_{n-1}^{\mathrm{II}} := \mathcal{K}_{n-2}^{\mathrm{II}} + \sum_{x\in\mathbf{X}} f_{n-1}^{\mathrm{II}}(x)(P_n^{\mathrm{II}}(x) - P_{n-1}^{\mathrm{II}}(x))$
  Sceptic announces $f_n^{\mathrm{I}}, f_n^{\mathrm{II}} \in \mathbb{R}^{\mathbf{X}}$

The following is the multiplicative version of Protocol 4.7; its name (*double Cover–Bayes game*) will be discussed after the protocol.

Protocol 4.8. Double Cover–Bayes game:

$\mathcal{K}_0^{\mathrm{I}} = \mathcal{K}_0^{\mathrm{II}} := 1$
FOR $n = 1, 2, \dots$:
  Forecaster I announces $P_n^{\mathrm{I}} \in \mathfrak{P}(\mathbf{X})$
  Forecaster II announces $P_n^{\mathrm{II}} \in \mathfrak{P}(\mathbf{X}))$
  IF $n > 1$:
    $\mathcal{K}_{n-1}^{\mathrm{I}} := \mathcal{K}_{n-2}^{\mathrm{I}} \sum_{x\in\mathbf{X}} G_{n-1}^{\mathrm{I}}(x) P_n^{\mathrm{I}}(x)/P_{n-1}^{\mathrm{I}}(x)$
    $\mathcal{K}_{n-1}^{\mathrm{II}} := \mathcal{K}_{n-2}^{\mathrm{II}} \sum_{x\in\mathbf{X}} G_{n-1}^{\mathrm{II}}(x) P_n^{\mathrm{II}}(x)/P_{n-1}^{\mathrm{II}}(x)$
  Sceptic announces $G_n^{\mathrm{I}}, G_n^{\mathrm{II}} \in \mathfrak{P}(\mathbf{X})$

It is easy to check that Protocols 4.7 and 4.8 are equivalent, in the sense of leading to the same positive capital processes $\mathcal{K}_n^{\mathrm{I}}$ and $\mathcal{K}_n^{\mathrm{II}}$.

Protocol 4.8 can be interpreted as representing each Forecaster as a financial market. For concreteness, let us talk about Forecaster I. There are $|\mathbf{X}|$ securities traded in the market, and they are indexed by $\mathbf{X}$. The prices $P^{\mathrm{I}}(x)$ of the securities always sum to 1, which will be the case if the prices are measured as fractions of the total market capitalization, as in [15, Sect. 17.1]. The interpretation of $G_n^{\mathrm{I}}(x)$ is as the fraction of $\mathcal{K}_{n-1}^{\mathrm{I}}$ invested in security x when playing against Forecaster I. Such a financial game without the restriction of the prices summing to 1 was considered by Cover [2]. I call it "Cover–Bayes" because of the restriction, and it is double since we have two Forecasters.

5 Proof of Theorem 4.5

We follow the proof of Theorem 2.2 in Sect. 3. Starting from step 5 of Algorithm 1 the capital $\mathcal{K}_n^{\mathrm{I}}$ will be multiplied by

$$\frac{1}{H_m(P_n^{\mathrm{I}}, P_n^{\mathrm{II}})} \frac{\mathrm{d}\sqrt{P_n^{\mathrm{I}}|_m P_n^{\mathrm{II}}|_m}}{\mathrm{d}P_n^{\mathrm{I}}|_m} \cdot (P_N^{\mathrm{I}}|_m) \tag{2}$$

by step N (later we will make N large), and starting from step 6 the capital $\mathcal{K}_n^{\mathrm{II}}$ will be multiplied by

$$\frac{1}{H_m(P_n^{\mathrm{I}}, P_n^{\mathrm{II}})} \frac{\mathrm{d}\sqrt{P_n^{\mathrm{I}}|_m P_n^{\mathrm{II}}|_m}}{\mathrm{d}P_n^{\mathrm{II}}|_m} \cdot (P_N^{\mathrm{II}}|_m) \tag{3}$$

by step N. The dot product "$\cdot$" in (2) and (3) refers to the representation of each probability measure Q on $\mathbf{Y}^m$ as a $|\mathbf{Y}|^m$-dimensional vector (with the components $Q([x])$, $x \in \mathbf{Y}^m$).

Using the weak convergence $P_N^{\mathrm{I}} - P_N^{\mathrm{II}} \to 0$ and the uniform continuity of (2) as function of $P_N^{\mathrm{I}}|_m$ and of (3) as function of $P_N^{\mathrm{II}}|_m$, we can replace $P_N^{\mathrm{I}}|_m$ and $P_N^{\mathrm{II}}|_m$ by the same probability measure P on $\mathbf{Y}^m$, and so the geometric mean of (2) and (3) can be bounded below as

$$\begin{aligned}
&\frac{1}{H_m(P_n^{\mathrm{I}}, P_n^{\mathrm{II}})} \sqrt{\left(\frac{\mathrm{d}\sqrt{P_n^{\mathrm{I}}|_m P_n^{\mathrm{II}}|_m}}{\mathrm{d}P_n^{\mathrm{I}}|_m} \cdot P\right)\left(\frac{\mathrm{d}\sqrt{P_n^{\mathrm{I}}|_m P_n^{\mathrm{II}}|_m}}{\mathrm{d}P_n^{\mathrm{II}}|_m} \cdot P\right)} \\
&\ge \frac{1}{H_m(P_n^{\mathrm{I}}, P_n^{\mathrm{II}})} \sqrt{\frac{\mathrm{d}\sqrt{P_n^{\mathrm{I}}|_m P_n^{\mathrm{II}}|_m}}{\mathrm{d}P_n^{\mathrm{I}}|_m} \frac{\mathrm{d}\sqrt{P_n^{\mathrm{I}}|_m P_n^{\mathrm{II}}|_m}}{\mathrm{d}P_n^{\mathrm{II}}|_m}} \cdot P \\
&= \frac{1}{H_m(P_n^{\mathrm{I}}, P_n^{\mathrm{II}})} > \frac{1}{1-\epsilon},
\end{aligned}$$

where the "$\ge$" follows from the concavity of the geometric mean function $(u, v) \in [0, \infty)^2 \mapsto \sqrt{uv}$ and Jensen's inequality [4, Lemma 2.8.1]. Therefore, we can wait until the geometric mean of the capitals increases $\frac{1}{1-\epsilon}$-fold. The proof is completed as before.

Finally, let us check formally that at step n Forecaster I has a strategy leading to a (2)-fold increase in his capital at any step $N \ge n$. (The argument for Forecaster II is analogous.) For $N := n$, this statement is true since (2) is 1. Let us now check that any strategy satisfying this statement for some value of N can be extended to satisfy it for the next value of N. Denoting by g the expression to the left of the "$\cdot$" in (2), we have

$$g \cdot (P_{N+1}^{\mathrm{I}}|_m) - g \cdot (P_N^{\mathrm{I}}|_m) = \sum_{x \in \mathbf{Y}^m} g(x)(P_{N+1}^{\mathrm{I}}(x) - P_N^{\mathrm{I}}(x)),$$

and so the required increment in Sceptic's capital is covered by (1).

6 Conclusion

A big disadvantage of Theorems 2.2 and 4.5, inherited from their measure-theoretic prototype [1], is that they are merely asymptotic; it is not obvious how to make them quantitative. The versions given in [15, Sect. 10.7] are much more precise, but they cover only one-step-ahead forecasting. Bridging the gap between these two very different kinds of results (depicted in Fig. 2) looks to me an interesting direction of further research.

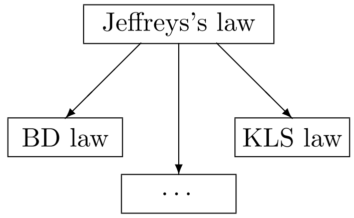


Figure 2: Two special cases of Jeffreys's law

Figure 2 represents two known special cases of Jeffreys's law. The Blackwell–Dubins theorem is an instance of one of them; more generally, Fig. 2 refers to the *BD law* as an asymptotic qualitative statement about convergence between forecasts for the infinite future (or, in the case of radical probabilism, infinite past, present, and future). The results of Kabanov, Liptser, and Shiryaev ([11], generalizing Kakutani [12]) are very different: they are quantitative (albeit with interesting qualitative implications); their major limitation is that they only cover one-step ahead prediction. They have similar disadvantages to those of the Blackwell–Dubins result discussed at the beginning of Sect. 1, but those were eliminated in the game-theoretic versions described in [15, Sect. 10.7] (and developed in the references given in [15, Sect. 10.9]). Figure 2 refers to this class of results as the *KLS law*. The ellipsis represents new special cases of Jeffreys's law, those yet to be discovered.

Acknowledgments

Many thanks to A. Philip Dawid for his questions and comments. Research on this paper has been partially supported by Mitie.

References

[1] David Blackwell and Lester Dubins. Merging of opinions with increasing information. *Annals of Mathematical Statistics*, 33:882–886, 1962.

[2] Thomas Cover. Universal portfolios. *Mathematical Finance*, 1:1–29, 1991.

[3] A. Philip Dawid. Statistical theory: the prequential approach (with discussion). *Journal of the Royal Statistical Society A*, 147:278–292, 1984.

[4] Thomas S. Ferguson. *Mathematical Statistics: A Decision Theoretic Approach*. Academic Press, New York, 1967.

[5] Simon M. Huttegger. Merging of opinions and probability kinematics. *Review of Symbolic Logic*, 8:611–648, 2015.

[6] Richard C. Jeffrey. Probable knowledge. In Imre Lakatos, editor, *The Problem of Inductive Logic*, pages 166–190. Elsevier, 1968. This paper is included as Chap. 3 (pp. 30–43) of [8].

[7] Richard C. Jeffrey. De Finetti's radical probabilism. Richard C. Jeffrey Papers archive at the University of Pittsburgh. Accessed online in December 2023, 1988.

[8] Richard C. Jeffrey. *Probability and the Art of Judgment*. Cambridge University Press, Cambridge, 1992.

[9] Harold Jeffreys. Science, logic and philosophy. *Nature*, 141:672–676 and 716–719, 1938.

[10] Harold Jeffreys. *Theory of Probability*. Oxford University Press, Oxford, third edition, 1961.

[11] Yury M. Kabanov, Robert Sh. Liptser, and Albert N. Shiryaev. To the question of absolute continuity and singularity of probability measures. *Mathematics of the USSR—Sbornik*, 33:203–221, 1977.

[12] Shizuo Kakutani. On equivalence of infinite product measures. *Annals of Mathematics, Second Series*, 49:214–224, 1948.

[13] Charles Sanders Peirce. The fixation of belief. *Popular Science Monthly*, 12:1–15, November 1877.

[14] Karl R. Popper. *The Open Society and Its Enemies*. Routledge, London, 1945. In 2 volumes. Fifth edition (revised): 1966.

[15] Glenn Shafer and Vladimir Vovk. *Game-Theoretic Foundations for Probability and Finance*. Wiley, Hoboken, NJ, 2019.

[16] Albert N. Shiryaev. *Probability-1*. Springer, New York, third edition, 2016.

[17] Albert N. Shiryaev. *Probability-2*. Springer, New York, third edition, 2019.

[18] Vladimir Vovk. The diachronic Bayesian. Technical Report arXiv:2308.13069 [stat.ME], arXiv.org e-Print archive, November 2023.

[19] Vladimir Vovk and Glenn Shafer. A conversation with A. Philip Dawid. *Statistical Science*, 2023. See "Future Papers" on the journal web site. An extended version is published as arXiv:2312.00632.