Computable Randomness is Inherently Imprecise

Gert de Cooman   GERT.DECOOMAN@UGENT.BE
Jasper De Bock   JASPER.DEBOCK@UGENT.BE
Ghent University, IDLab
Technologiepark – Zwijnaarde 914, 9052 Zwijnaarde (Belgium)

Abstract

We use the martingale-theoretic approach of game-theoretic probability to incorporate imprecision into the study of randomness. In particular, we define a notion of computable randomness associated with interval, rather than precise, forecasting systems, and study its properties. The richer mathematical structure that thus arises lets us better understand and place existing results for the precise limit. When we focus on constant interval forecasts, we find that every infinite sequence of zeroes and ones has an associated filter of intervals with respect to which it is computably random. It may happen that none of these intervals is precise, which justifies the title of this paper. We illustrate this by showing that computable randomness associated with non-stationary precise forecasting systems can be captured by a stationary interval forecast, which must then be less precise: a gain in model simplicity is thus paid for by a loss in precision.

Keywords: computable randomness; imprecise probabilities; game-theoretic probability; interval forecast; supermartingale; computability.

1. Introduction

This paper documents the first steps in our attempt to incorporate indecision and imprecision into the study of randomness. Consider a infinite sequence $\omega = (z_1, \ldots, z_n, \ldots)$ of zeroes and ones; when do we call it *random*? There are many notions of randomness, and many of them have a number of equivalent definitions (Ambos-Spies and Kucera, 2000; Bienvenu et al., 2009). We focus here on *computable randomness*, mainly because its focus on computability—rather than, say, the weaker lower semicomputability—has allowed us in this first attempt to keep the mathematical nitpicking at arm's length. Randomness of a sequence ω is typically associated with a probability measure on the sample space of all infinite sequences, or—what is equivalent—with a *forecasting system* γ that associates with each finite sequence of outcomes $(x_1, \ldots, x_n)$ the (conditional) expectation $\gamma(x_1, \ldots, x_n)$ for the next (as yet unknown) outcome X_{n+1}. The sequence ω is then called *computably* random when it passes a (countable) number of *computable* tests of randomness, where the collection of randomness tests depends of the forecasting system γ. An alternative but equivalent definition, going back to Ville (1939), sees each forecast $\gamma(x_1, \ldots, x_n)$ as a fair price for—and therefore a commitment to bet on—the as yet unknown next outcome X_{n+1}. The sequence ω is then computably random when there is no computable strategy for getting infinitely rich by exploiting the bets made available by the forecasting system γ along the sequence, without borrowing. Technically speaking, all computable non-negative supermartingales should remain bounded on ω, and the forecasting system γ determines what a supermartingale is.

It is this last, martingale-theoretic approach which seems to lend itself most easily to allowing for imprecision in the forecasts, and therefore in the definition of randomness. As we explain in Sections 2 and 3, an 'imprecise' forecasting system γ associates with each finite sequence of outcomes

$(x_1, \dots, x_n)$ a (conditional) expectation *interval* $\gamma(x_1, \dots, x_n)$ for the next (as yet unknown) outcome X_{n+1}, whose lower bound represents a supremum acceptable buying price, and whose upper bound a infimum acceptable selling price for X_{n+1}. This idea rests firmly on the common ground between Walley's (1991) theory of coherent lower previsions and Shafer and Vovk's (2001) game-theoretic approach to probability that we have established in recent years, through our research on imprecise stochastic processes (De Cooman and Hermans, 2008; De Cooman et al., 2016). This allows us to associate supermartingales with an imprecise forecasting system, and therefore in Section 5 to extend the existing notion of computable randomness to allow for interval, rather than precise, forecasts—we discuss computability in Section 4. We show in Section 6 that our approach allows us to extend some of Dawid's (1982) well-known work on calibration, as well as an interesting 'limiting frequencies' or computable stochasticity result.

We believe the discussion becomes really interesting in Section 7, where we look at stationary interval forecasts to extend the classical account of randomness. That classical account typically considers a forecasting system with stationary expectation forecast $1/2$—corresponding to flipping a fair coin. As we have by now come to expect from our experience with imprecise probability models, a much more interesting mathematical picture appears when allowing for interval forecasts than the rather simple case of precise forecasts would lead us to suspect. In the precise case, a given sequence may not be (computably) random for any stationary forecast, but in the imprecise case there is always a set filter of intervals that a given sequence is computably random for. Furthermore, as we show in Section 8, this filter may not have a smallest element, and even when it does, this smallest element may be a non-vanishing interval: randomness may be inherently imprecise.

In order to comply with the page limit, proofs are omitted; we refer the reader to the appendix of (De Cooman and De Bock, 2017), an extended version of this paper that is available on arXiv.

2. A single interval forecast

The dynamics of making a single forecast can be made very clear by considering a simple game, with three players, namely Forecaster, Sceptic and Reality.

Game: single forecast of an outcome X
In a first step, Forecaster specifies an interval bound $I = [\underline{p}, \overline{p}]$ for the expectation of an as yet unknown outcome X in $\{0, 1\}$—or equivalently, for the probability that $X = 1$. We interpret this *interval forecast I* as a commitment, on the part of Forecaster, to adopt $\underline{p}$ as a *supremum buying price* and $\overline{p}$ as a *infimum selling price* for the gamble (with reward function) X. This is taken to mean that the second player, *Sceptic*, can now in a second step take Forecaster up on any (combination) of the following commitments:

  (i) for any $p \in [0, 1]$ such that $p \le \underline{p}$, and any $\alpha \ge 0$ Forecaster must accept the gamble $\alpha[X - p]$, leading to an uncertain reward $-\alpha[X - p]$ for Sceptic;[1]
  (ii) for any $q \in [0, 1]$ such that $q \ge \overline{p}$, and any $\beta \ge 0$ Forecaster accepts the gamble $\beta[q - X]$, leading to an uncertain reward $-\beta[q - X]$ for Sceptic.

Finally, in a third step, the third player, *Reality*, determines the value x of X in $\{0, 1\}$. □

Elements x of $\{0, 1\}$ are called *outcomes*, and elements p of the real unit interval $[0, 1]$ are called (precise) *forecasts*. We denote by $\mathscr{C}$ the set of non-empty closed subintervals of the real unit

---

1. Because we allow $p \le \underline{p}$ rather than $p < \underline{p}$, we actually see $\underline{p}$ as a *maximum* buying price, rather than a supremum one. We do this because it does not affect the conclusions, but simplifies the mathematics. Similarly for $q \ge \overline{p}$.
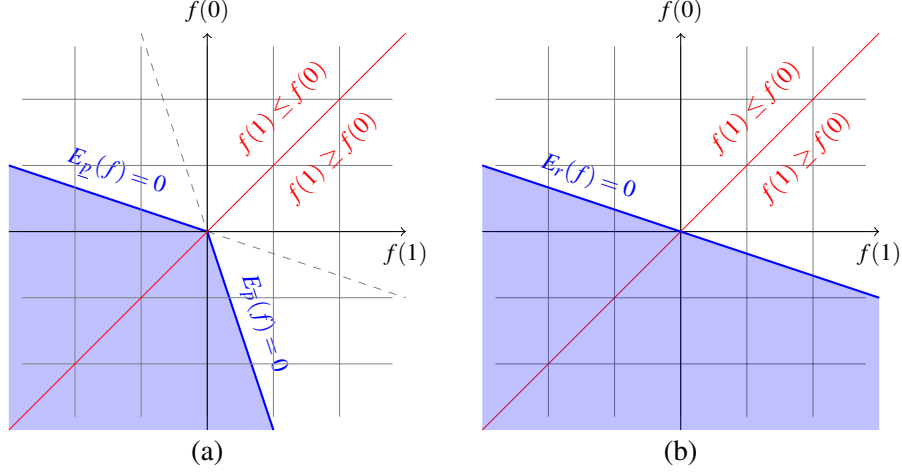
Figure 1: Gambles f available to Sceptic when (a) Forecaster announces $I \in \mathscr{C}$ with $\underline{p} < \overline{p}$; and when (b) Forecaster announces $I \in \mathscr{C}$ with $\underline{p} = \overline{p} =: r$.

interval $[0,1]$. Any element I of $\mathscr{C}$ is called an *interval forecast*. It has a smallest element $\min I$ and a greatest element $\max I$, so $I = [\min I, \max I]$. We will use the generic notation I for such an interval, and $\underline{p} := \min I$ and $\overline{p} := \max I$ for its lower and upper bounds, respectively.

After Forecaster announces a forecast interval I, what Sceptic can do is essentially to try and increase his capital by taking a gamble on the outcome X. Any such gamble can be considered as a map $f: \{0,1\} \to \mathbb{R}$, and can therefore be represented as a vector $(f(1), f(0))$ in the two-dimensional vector space $\mathbb{R}^2$; see also Figure 1. $f(X)$ is then the increase in Sceptic's capital after the game has been played, as a function of the outcome variable X. Of course, not every gamble $f(X)$ on the outcome X will be available to Sceptic: which gambles he can take is determined by Forecaster's interval forecast I. In their most general form, they are given by $f(X) = -\alpha[X - p] - \beta[q - X]$, where α and β are non-negative real numbers, $p \leq \underline{p}$ and $q \geq \overline{p}$. If we consider the so-called *lower expectation* (functional) $\underline{E}_I$ associated with an interval forecast I, defined by

$$\underline{E}_I(f) = \min_{p \in I} E_p(f) = \min_{p \in I}\big[pf(1) + (1-p)f(0)\big] = \begin{cases} E_{\underline{p}}(f) & \text{if } f(1) \geq f(0) \\ E_{\overline{p}}(f) & \text{if } f(1) \leq f(0) \end{cases} \tag{1}$$

for any gamble $f: \{0,1\} \to \mathbb{R}$, and similarly, the *upper expectation* (functional) $\overline{E}_I$, defined by

$$\overline{E}_I(f) = \max_{p \in I} E_p(f) = \begin{cases} E_{\overline{p}}(f) & \text{if } f(1) \geq f(0) \\ E_{\underline{p}}(f) & \text{if } f(1) \leq f(0) \end{cases} = -\underline{E}_I(-f), \tag{2}$$

then it is not difficult to see that *the cone of gambles $f(X)$ that are available to Sceptic after Forecaster announces an interval forecast I is completely determined by the condition* $\overline{E}_I(f) \leq 0$, as depicted by the blue regions in Figure 1. The functionals $\underline{E}_I$ and $\overline{E}_I$ are easily shown to have the following properties, typical for the more general lower and upper expectation operators defined on more general gamble spaces (Walley, 1991; Troffaes and De Cooman, 2014):
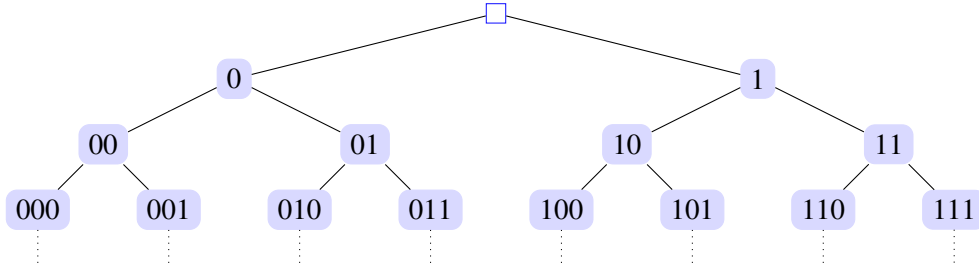
Proposition 1 *Consider any forecast interval* $I \in \mathscr{C}$. *Then for all gambles* f,g *on* $\{0,1\}$, $\mu \in \mathbb{R}$ *and non-negative* $\lambda \in \mathbb{R}$:

C1. $\min f \leq \underline{E}_I(f) \leq \overline{E}_I(f) \leq \max f$;   [bounds]

C2. $\underline{E}_I(\lambda f) = \lambda \underline{E}_I(f)$ and $\overline{E}_I(\lambda f) = \lambda \overline{E}_I(f)$;   [non-negative homogeneity]

C3. $\underline{E}_I(f+g) \geq \underline{E}_I(f) + \underline{E}_I(g)$ and $\overline{E}_I(f+g) \leq \overline{E}_I(f) + \overline{E}_I(g)$;   [super/subadditivity]

C4. $\underline{E}_I(f+\mu) = \underline{E}_I(f) + \mu$ and $\overline{E}_I(f+\mu) = \overline{E}_I(f) + \mu$.   [constant additivity]
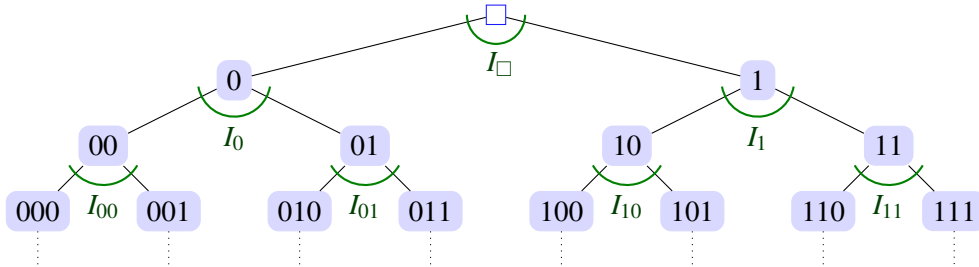
3. Interval forecasting systems and imprecise probability trees

We now consider a sequence of repeated versions of the forecast game in the previous section, where at each stage $k \in \mathbb{N}$, Forecaster presents an interval forecast $I_k = [\underline{p}_k, \overline{p}_k]$ for the unknown outcome variable X_k. This effectively allows Sceptic to choose any gamble $f_k(X_k)$ such that $\overline{E}_{I_k}(f_k) \leq 0$. Reality then chooses a value x_k for X_k, resulting in a gain, or increase in capital, $f_k(x_k)$ for Sceptic.

We call $(x_1, x_2, \ldots, x_n, \ldots)$ an *outcome sequence*, and collect all possible outcome sequences in the set $\Omega := \{0,1\}^{\mathbb{N}}$. We collect the finite outcome sequences $(x_1, \ldots, x_n)$ in the set $\Omega^{\Diamond} := \{0,1\}^* = \bigcup_{n \in \mathbb{N}_0} \{0,1\}^n$. Finite sequences s in $\Omega^{\Diamond}$ and infinite sequences ω in Ω are the nodes—called *situations*—and paths in an event tree with unbounded horizon, part of which is depicted below.

In this repeated game, Forecaster will only provide interval forecasts I_k after observing the actual sequence $(x_1, \ldots, x_{k-1})$ that Reality has chosen. This is the essence of so-called *prequential forecasting* (Dawid, 1982, 1984; Dawid and Vovk, 1999). But for technical reasons, it will be useful to consider the more involved setting where a forecast I_s is specified in each of the possible situations $s \in \Omega^{\Diamond}$; see the figure below.

Indeed, we can use this idea to generalise the notion of a forecasting system (Vovk and Shen, 2010).

Definition 2 (Forecasting system) *A forecasting system is a map* $\gamma \colon \Omega^{\Diamond} \to \mathscr{C}$, *that associates with any situation* s *in the event tree a forecast* $\gamma(s) \in \mathscr{C}$. *With any forecasting system* γ *we can associate two real-valued maps* $\underline{\gamma}$ *and* $\overline{\gamma}$ *on* $\Omega^{\Diamond}$, *defined by* $\underline{\gamma}(s) := \min \gamma(s)$ *and* $\overline{\gamma}(s) := \max \gamma(s)$ *for all* $s \in \Omega^{\Diamond}$. *A forecasting system* γ *is called* precise *if* $\underline{\gamma} = \overline{\gamma}$. Γ *denotes the set* $\mathscr{C}^{\Omega^{\Diamond}}$ *of all forecasting systems.*

Specifying such a forecasting system requires imagining in advance all moves that Reality could make, and devising in advance what forecasts to give in each imaginable situation s. In the precise case, that is typically what one does when specifying a probability measure on the so-called *sample space* Ω—the set Ω of all paths.

Since in each situation s the interval forecast $I_s = \gamma(s)$ corresponds to a local lower expectation $\underline{E}_{I_s}$, we can use the argumentation in our earlier papers (De Cooman and Hermans, 2008; De Cooman et al., 2016) on stochastic processes to let the forecasting system γ turn the event tree into a so-called *imprecise probability tree*, with an associated global lower expectation, and a corresponding notion of '(strictly) almost surely'. In what follows, we briefly recall how to do this; for more context, we also refer to the seminal work by Shafer and Vovk (2001).

For any path $\omega \in \Omega$, the initial sequence that consists of its first n elements is a situation in $\{0,1\}^n$ that is denoted by ω^n. Its n-th element belongs to $\{0,1\}$ and is denoted by ω_n. As a convention, we let its 0-th element be the *initial* situation $\omega^0 = \omega_0 = \square$. We write that $s \sqsubseteq t$, and say that the situation s *precedes* the situation t, when every path that goes through t also goes through s—so s is a precursor of t.

A *process* F is a map defined on $\Omega^\lozenge$. A *real process* is a real-valued process: it associates a real number $F(s) \in \mathbb{R}$ with every situation $s \in \Omega^\lozenge$. With any real process F, we can always associate a process ΔF, called the *process difference*. For every situation $(x_1, \ldots, x_n)$ with $n \in \mathbb{N}_0$, $\Delta F(x_1, \ldots, x_n)$ is a gamble on $\{0,1\}$ defined by $\Delta F(x_1, \ldots, x_n)(x_{n+1}) := F(x_1, \ldots, x_{n+1}) - F(x_1, \ldots, x_n)$ for all $x_{n+1} \in \{0,1\}$. In the imprecise probability tree associated with a *given* forecasting system γ, a *submartingale* M for γ is a real process such that $\underline{E}_{\gamma(x_1, \ldots, x_n)}(\Delta M(x_1, \ldots, x_n)) \geq 0$ for all $n \in \mathbb{N}_0$ and $(x_1, \ldots, x_n) \in \{0,1\}^n$. A real process M is a *supermartingale* for γ if $-M$ is a submartingale, meaning that $\overline{E}_{\gamma(x_1, \ldots, x_n)}(\Delta M(x_1, \ldots, x_n)) \leq 0$ for all $n \in \mathbb{N}_0$ and $(x_1, \ldots, x_n) \in \{0,1\}^n$: all supermartingale differences have non-positive upper expectation, so supermartingales are real processes that Forecaster expects to decrease. We denote the set of all submartingales for a given forecasting system γ by $\underline{\mathbb{M}}^\gamma$—whether a real process is a submartingale depends of course on the forecasts in the situations. Similarly, the set $\overline{\mathbb{M}}^\gamma := -\underline{\mathbb{M}}^\gamma$ is the set of all supermartingales for γ.

It is clear from the discussion in Section 2 that the supermartingales are effectively all the possible capital processes $\mathscr{K}$ for a Sceptic who starts with an initial capital $\mathscr{K}(\square)$, and in each possible subsequent situation s selects a gamble $f_s = \Delta\mathscr{K}(s)$ that is available there because Forecaster specifies the interval forecast $I_s = \gamma(s)$ and because $\overline{E}_{I_s}(f_s) = \overline{E}_{\gamma(s)}(\Delta\mathscr{K}(s)) \leq 0$. If Reality chooses outcomes $s = (x_1, \ldots, x_n)$, then Sceptic ends up with capital $\mathscr{K}(x_1, \ldots, x_n) = \mathscr{K}(\square) + \sum_{k=0}^{n-1} \Delta\mathscr{K}(x_1, \ldots, x_k)(x_{k+1})$. A *non-negative* supermartingale M is non-negative in all situations, which corresponds to Sceptic never borrowing any money. We call *test supermartingale* any non-negative supermartingale M that starts with unit capital $M(\square) = 1$. We collect all test supermartingales for γ in the set $\overline{\mathbb{T}}^\gamma$.

In the context of probability trees, we call *variable* any function defined on the sample space Ω. When this variable is real-valued and bounded, we call it a *gamble* on Ω. An *event* A in this context is a subset of Ω, and its indicator $\mathbb{I}_A$ is a gamble on Ω assuming the value 1 on A and 0 elsewhere. The following expressions define lower and upper expectations on such gambles g on Ω:

$$\underline{E}^\gamma(g) := \sup\left\{ M(\square) \colon M \in \underline{\mathbb{M}}^\gamma \text{ and } \limsup_{n \to +\infty} M(\omega^n) \leq g(\omega) \text{ for all } \omega \in \Omega \right\} \tag{3}$$

$$\overline{E}^\gamma(g) := \inf\left\{ M(\square) \colon M \in \overline{\mathbb{M}}^\gamma \text{ and } \liminf_{n \to +\infty} M(\omega^n) \geq g(\omega) \text{ for all } \omega \in \Omega \right\} = -\underline{E}^\gamma(g). \tag{4}$$

They satisfy coherence properties similar to those in Proposition 1. We refer to extensive discussions elsewhere (De Cooman et al., 2016; Shafer and Vovk, 2001) about why these expressions are interesting and useful. For our present purposes, it may suffice to mention that for precise forecasts, they lead to models that coincide with the ones found in measure-theoretic probability theory (Shafer and Vovk, 2001, Chapter 8). *In particular, when all $I_s = \{1/2\}$, they coincide with the usual uniform (Lebesgue) expectations on measurable gambles.*

We call an event $A \subseteq \Omega$ *null* if $\overline{P}^\gamma(A) := \overline{E}^\gamma(\mathbb{I}_A) = 0$, or equivalently $\underline{P}^\gamma(A^c) := \underline{E}^\gamma(\mathbb{I}_{A^c}) = 1$, and *strictly null* if there is some test supermartingale $T \in \overline{\mathbb{T}}^\gamma$ that converges to $+\infty$ on A, meaning that $\lim_{n \to +\infty} T(\omega^n) = +\infty$ for all $\omega \in A$. Any strictly null event is null, but null events need not be strictly null (Vovk and Shafer, 2014; De Cooman et al., 2016). Because it is easily checked that $\overline{P}^\gamma(\emptyset) = \underline{P}^\gamma(\emptyset) = 0$, the complement A^c of a (strictly) null event A is never empty. As usual, any property that holds, except perhaps on a (strictly) null event, is said to hold (strictly) *almost surely*.

4. Basic computability notions

We recall a few notions and results from computability theory that are relevant to the discussion. For a much more extensive treatment, we refer for instance to the books by Pour-El and Richards (1989) and Li and Vitányi (1993).

A *computable* function $\phi \colon \mathbb{N}_0 \to \mathbb{N}_0$ is a function that can be computed by a Turing machine. All further notions of computability that we will need, build on this basic notion. It is clear that it in this definition, we can replace any of the $\mathbb{N}_0$ with any other countable set.

We start with the definition of a computable real number. We call a sequence of rational numbers r_n *computable* if there are three computable functions a, b, σ from $\mathbb{N}_0$ to $\mathbb{N}_0$ such that $b(n) > 0$ and $r_n = (-1)^{\sigma(n)} \frac{a(n)}{b(n)}$ for all $n \in \mathbb{N}_0$, and we say that it *converges effectively* to a real number x if there is some computable function $e \colon \mathbb{N}_0 \to \mathbb{N}_0$ such that $n \geq e(N) \Rightarrow |r_n - x| \leq 2^{-N}$ for all $n, N \in \mathbb{N}_0$. A real number is then called *computable* if there is a computable sequence of rational numbers that converges effectively to it. Of course, every rational number is a computable real.

We also need a notion of computable real processes, or in other words, computable real-valued maps $F \colon \Omega^\diamond \to \mathbb{R}$ defined on the set $\Omega^\diamond$ of all situations. Because there is an obvious computable bijection between $\mathbb{N}_0$ and $\Omega^\diamond$, whose inverse is also computable, we can in fact identify real processes and real sequences, and simply import, *mutatis mutandis*, the definitions for computable real sequences common in the literature (Li and Vitányi, 1993, Chapter 0). Indeed, we call a net of rational numbers $r_{s,n}$ *computable* if there are three computable functions a, b, s from $\Omega^\diamond \times \mathbb{N}_0$ to $\mathbb{N}_0$ such that $b(s,n) > 0$ and $r_{s,n} = (-1)^{\sigma(s,n)} \frac{a(s,n)}{b(s,n)}$ for all $s \in \Omega^\diamond$ and $n \in \mathbb{N}_0$. We call a real process $F \colon \Omega^\diamond \to \mathbb{R}$ *computable* if there is a computable net of rational numbers $r_{s,n}$ and a computable function $e \colon \Omega^\diamond \times \mathbb{N}_0 \to \mathbb{N}_0$ such that $n \geq e(s,N) \Rightarrow |r_{s,n} - F(s)| \leq 2^{-N}$ for all $s \in \Omega^\diamond$ and $n, N \in \mathbb{N}_0$. Obviously, it follows from this definition that in particular $F(t)$ is a computable real number for any $t \in \Omega^\diamond$: fix $s = t$ and consider the sequence $r_{t,n}$ that converges to $F(s)$ as $n \to +\infty$. Also, a constant real process is computable if and only if its constant value is.

The following definitions are now obvious. A gamble f on $\{0,1\}$ is called *computable* if both its values $f(0)$ and $f(1)$ are computable real numbers. An interval forecast $I = [\underline{p}, \overline{p}] \in \mathscr{C}$ is called *computable* if both its lower bound $\underline{p}$ and upper bound $\overline{p}$ are computable real numbers. A forecasting system γ is called *computable* if the associated real processes $\underline{\gamma}$ and $\overline{\gamma}$ are.

5. Random sequences in an imprecise probability tree

We will now associate a notion of randomness with a forecasting system γ—or in other words, with an imprecise probability tree. In what follows, we will often consider computable test supermartingales. These computable test supermartingales for a forecasting system are countable in number, because the computable processes are (Li and Vitányi, 1993; Vovk and Shen, 2010).

Definition 3 (Computable randomness) *Consider any forecasting system $\gamma\colon \Omega^\diamond \to \mathscr{C}$. We call an outcome sequence ω computably random for γ if all computable test supermartingales T remain bounded above on ω, meaning that there is some $B \in \mathbb{R}$ such that $T(\omega^n) \le B$ for all $n \in \mathbb{N}$, or equivalently, that $\sup_{n\in\mathbb{N}} T(\omega^n) < +\infty$. We then also say that the forecasting system γ makes ω computably random. We denote by $\Gamma_C(\omega) := \{\gamma \in \Gamma\colon \omega \text{ is computably random for } \gamma\}$ the set of all forecasting systems for which the outcome sequence ω is computably random.*

Computable randomness of an outcome sequence means that there is no computable strategy that starts with capital 1 and avoids borrowing, and allows Sceptic to increase his capital without bounds by exploiting the bets on these outcomes that are made available to him by Forecaster's specification of the forecasting system γ. When the forecasting system γ is precise and computable, our notion of computable randomness reduces to the classical notion of computable randomness (Ambos-Spies and Kucera, 2000; Bienvenu et al., 2009).

The (computable) *vacuous* forecasting system γ_v assigns the vacuous forecast $\gamma_v(s) := [0,1]$ to all situations $s \in \Omega^\diamond$. The following proposition implies that no $\Gamma_C(\omega)$ is empty.

Proposition 4 *All paths are computably random for the vacuous forecasting system: $\gamma_v \in \Gamma_C(\omega)$ for all $\omega \in \Omega$.*

More conservative (or imprecise) forecasting systems have more computably random sequences.

Proposition 5 *Let ω be computably random for a forecasting system γ. Then ω is also computably random for any forecasting system γ^* such that $\gamma \subseteq \gamma^*$, meaning that $\gamma(s) \subseteq \gamma^*(s)$ for all $s \in \Omega^\diamond$.*

6. Consistency results

We first show that any Forecaster who specifies a forecasting system is consistent in the sense that he believes himself to be *well calibrated*: in the imprecise probability tree generated by his own forecasts, (strictly) almost all paths will be computably random, so he is sure that Sceptic will not be able to become infinitely rich at his expense, by exploiting his—Forecaster's—forecasts. This also generalises the arguments and conclusions in a paper by Dawid (1982).

Theorem 6 *Consider any forecasting system $\gamma\colon \Omega^\diamond \to \mathscr{C}$. Then (strictly) almost all outcome sequences are computably random for γ in the imprecise probability tree that corresponds to γ.*

This result is quite powerful, and it guarantees in particular that:

Corollary 7 *For any sequence of interval forecasts $(I_1, \ldots, I_n, \ldots)$ there is a forecasting system given by $\gamma(x_1, \ldots, x_n) := I_{n+1}$ for all $(x_1, \ldots, x_n) \in \{0,1\}^n$ and all $n \in \mathbb{N}_0$, and associated imprecise probability tree such that (strictly) almost all—and therefore definitely at least one—outcome sequences are computably random for γ in the associated imprecise probability tree.*

The following weaker consistency result deals with limits (inferior and superior) of relative frequencies, taken with respect to a so-called *selection process* $S\colon \Omega^\diamond \to \{0,1\}$. It is a counterpart in our more general context of the notions of *computable stochasticity* or *Church randomness* in the precise case with $I = \{1/2\}$ ([Ambos-Spies and Kucera, 2000](#)).

Theorem 8 (Church randomness) *Let* $\gamma\colon \Omega^\diamond \to \mathscr{C}$ *be any* computable *forecasting system, let* $\omega = (x_1,\dots,x_n,\dots) \in \Omega$ *be any outcome sequence that is computably random for* γ, *and let* f *be any* computable *gamble on* $\{0,1\}$. *If* $S\colon \Omega^\diamond \to \{0,1\}$ *is any computable selection process such that* $\sum_{k=0}^{n} S(x_1,\dots,x_k) \to +\infty$, *then also*

$$\liminf_{n\to+\infty} \frac{\sum_{k=0}^{n-1} S(x_1,\dots,x_k)\left[f(x_{k+1}) - \underline{E}_{\gamma(x_1,\dots,x_k)}(f)\right]}{\sum_{k=0}^{n-1} S(x_1,\dots,x_k)} \geq 0.$$

7. Constant interval forecasts

We now introduce a significant simplification. For any interval $I \in \mathscr{C}$, we let γ_I be the corresponding *stationary forecasting system* that assigns the same interval forecast I to all nodes: $\gamma_I(s) \coloneqq I$ for all $s \in \Omega^\diamond$. In this way, with any outcome sequence ω, we can associate the collection of all interval forecasts for which the corresponding stationary forecasting system makes ω computably random:

$$\mathscr{C}_C(\omega) \coloneqq \{I \in \mathscr{C} : \gamma_I \in \Gamma_C(\omega)\} = \{I \in \mathscr{C} : \gamma_I \text{ makes } \omega \text{ computably random}\}.$$

As an immediate consequence of Propositions [4](#) and [5](#), we find that this set of intervals is non-empty and increasing.

Proposition 9 (Non-emptiness) *For all* $\omega \in \Omega$, $[0,1] \in \mathscr{C}_C(\omega)$, *so any sequence of outcomes* ω *has at least one stationary forecast that makes it computably random:* $\mathscr{C}_C(\omega) \neq \emptyset$.

Proposition 10 (Increasingness) *Consider any* $\omega \in \Omega$ *and any* $I,J \in \mathscr{C}$. *If* $I \in \mathscr{C}_C(\omega)$ *and* $I \subseteq J$, *then also* $J \in \mathscr{C}_C(\omega)$.

Theorem [8](#) implies the following property. However, quite remarkably, and seemingly in contrast with Theorem [8](#), it does not require any computability assumptions on the (stationary) forecasts.

Corollary 11 (Church randomness) *Consider any outcome sequence* $\omega = (x_1,\dots,x_n,\dots)$ *in* Ω *and any stationary interval forecast* $I = [\underline{p},\overline{p}] \in \mathscr{C}_C(\omega)$ *that makes* ω *computably random. Then for any computable selection process* $S\colon \Omega^\diamond \to \{0,1\}$ *such that* $\sum_{k=0}^{n} S(x_1,\dots,x_k) \to +\infty$:

$$\underline{p} \leq \liminf_{n\to+\infty} \frac{\sum_{k=0}^{n-1} S(x_1,\dots,x_k)x_{k+1}}{\sum_{k=0}^{n-1} S(x_1,\dots,x_k)} \leq \limsup_{n\to+\infty} \frac{\sum_{k=0}^{n-1} S(x_1,\dots,x_k)x_{k+1}}{\sum_{k=0}^{n-1} S(x_1,\dots,x_k)} \leq \overline{p}.$$

The following proposition can of course be straightforwardly extended to any finite number of interval forecasts, and guarantees, together with Proposition [10](#), that $\mathscr{C}_C(\omega)$ is a *set filter*.

Proposition 12 *For any* $\omega \in \Omega$ *and any two interval forecasts* I *and* J: *if* $I \in \mathscr{C}_C(\omega)$ *and* $J \in \mathscr{C}_C(\omega)$ *then* $I \cap J \neq \emptyset$, *and* $I \cap J \in \mathscr{C}_C(\omega)$.

This result also tells us that the collection $\mathscr{C}_C(\omega)$ of closed subsets of the compact set $[0,1]$ has the finite intersection property, and its intersection is therefore a non-empty closed interval: $\bigcap \mathscr{C}_C(\omega) = [\underline{p}_C(\omega), \overline{p}_C(\omega)]$. Propositions 10 and 12 guarantee that all intervals $[\underline{p}_C(\omega) - \varepsilon_1, \overline{p}_C(\omega) + \varepsilon_2]$ in $\mathscr{C}$ with $\varepsilon_1, \varepsilon_2 > 0$ belong to $\mathscr{C}_C(\omega)$. But we will see in the next section that this does not generally hold for $\varepsilon_1 = 0$ and/or $\varepsilon_2 = 0$. For this reason, we now define the following two subsets of $[0,1]$:

$$L_C(\omega) := \{\min I \colon I \in \mathscr{C}_C(\omega)\} \text{ and } U_C(\omega) := \{\max I \colon I \in \mathscr{C}_C(\omega)\}.$$

Then Proposition 10 guarantees that $L_C(\omega)$ is a decreasing set, and that $U_C(\omega)$ is increasing. They are therefore both subintervals of $[0,1]$. Obviously, $\underline{p}_C(\omega) = \sup L_C(\omega)$ and $\overline{p}_C(\omega) = \inf U_C(\omega)$. On the one hand clearly $L_C(\omega) = [0, \underline{p}_C(\omega))$ or $L_C(\omega) = [0, \underline{p}_C(\omega)]$, and on the other hand $U_C(\omega) = (\overline{p}_C(\omega), 1]$ or $U_C(\omega) = [\overline{p}_C(\omega), 1]$. Proposition 12 easily allows us to give the following simple description of the set $\mathscr{C}_C(\omega)$ in terms of $L_C(\omega)$ and $U_C(\omega)$:

$$I \in \mathscr{C}_C(\omega) \Leftrightarrow \Big(\min I \in L_C(\omega) \text{ and } \max I \in U_C(\omega) \Big).$$

A trivial example is given by:

Proposition 13 *If the sequence ω is computable with infinitely many zeroes and ones, then $\mathscr{C}_C(\omega) = \{[0,1]\}$, and therefore $L_C(\omega) = \{0\}$, $U_C(\omega) = \{1\}$, $\underline{p}_C(\omega) = 0$ and $\overline{p}_C(\omega) = 1$.*

At the other extreme, there are the sequences ω that are computably random for some *precise* stationary forecasting system $\gamma_{\{p\}}$, with $p \in [0,1]$. They are amongst the random sequences that have received most attention in the literature, thus far. For any such sequence, $\mathscr{C}_C(\omega) = \{I \in \mathscr{C} \colon p \in I\}$, $L_C(\omega) = [0,p]$ and $U_C(\omega) = [p,0]$, and therefore also $\underline{p}_C(\omega) = \overline{p}_C(\omega) = p$.

We show in the next section that, in between these extremes of total imprecision and maximal precision, there lies a—to the best of our knowledge—previously uncharted realm of sequences, with similar (and even in some sense 'larger') unpredictability than the ones traditionally called 'computably random', for which $L_C(\omega)$ and $U_C(\omega)$ need not always be closed, and more importantly, for which $0 < \underline{p}_C(\omega) < \overline{p}_C(\omega) < 1$. This is what we mean when we claim that 'computable randomness is inherently imprecise'.

8. Randomness is inherently imprecise

Our work on imprecise Markov chains (De Cooman et al., 2016) has taught us that in some cases, we can very efficiently compute tight bounds on expectations in non-stationary precise Markov chains, by replacing them with their stationary imprecise versions. Similarly, in statistical modelling, when learning from data sampled from a distribution with a varying (non-stationary) parameter, it seems hard to estimate the exact time sequence of its values. But we may be more successful in learning about its (stationary) interval *range*. This idea was also considered earlier by Fierens et al. (2009), when they argued for a frequentist interpretation of imprecise probability models based on non-stationarity.

In this section, we exploit this idea, by showing that randomness associated with non-stationary precise forecasting systems can be captured by a stationary forecasting system, which must then be less precise: we gain simplicity of representation, but pay for it by losing precision.

We begin with a simple example. Consider any p and q in $[0,1]$ with $p \leq q$, and any outcome sequence $\omega = (x_1, \ldots, x_n, \ldots)$ that is computably random for the forecasting system $\gamma_{p,q}$ that is

defined by

$$\gamma_{p,q}(z_1,\ldots,z_n) := \begin{cases} p & \text{if } n \text{ is odd} \\ q & \text{if } n \text{ is even} \end{cases} \quad \text{for all } (z_1,\ldots,z_n) \in \Omega^\Diamond.$$

We know from Corollary 7 that there is at least one such outcome sequence. It turns out that the stationary forecasting systems that make such ω computably random have a simple characterisation:

Proposition 14 *Consider any ω that is computably random for the forecasting system $\gamma_{p,q}$. Then for all $I \in \mathscr{C}$, $I \in \mathscr{C}_C(\omega) \Leftrightarrow [p,q] \subseteq I$.*

Its proof relies on a very simple argument involving Corollary 11. This result implies in particular also that $L_C(\omega) = [0,p]$, $U_C(\omega) = [q,1]$, $\underline{p}_C(\omega) = p$ and $\overline{p}_C(\omega) = q$.

Next, we turn to a more complicated example, where we look at sequences that are 'nearly' computably random for the stationary precise forecast $1/2$, but not quite. This example was inspired by the ideas involving Hellinger-like divergences in a beautiful paper by Vovk (2009).

Consider the following sequence $\{p_n\}_{n \in \mathbb{N}}$ of precise forecasts:

$$p_n := \frac{1}{2} + (-1)^n \delta_n, \text{ with } \delta_n := e^{-\frac{1}{n+1}} \sqrt{e^{\frac{1}{n+1}} - 1} \text{ for all } n \in \mathbb{N},$$

converging to $1/2$. Observe that the sequence δ_n is decreasing towards its limit 0 and that $\delta_n \in (0, 1/2)$ and $p_n \in (0,1)$, for all $n \in \mathbb{N}$. Now consider any outcome sequence $\omega = (x_1, \ldots, x_n, \ldots)$ that is computably random for the precise forecasting system $\gamma_{\sim 1/2}$ that is defined by

$$\gamma_{\sim 1/2}(z_1,\ldots,z_{n-1}) := p_n \text{ for all } n \in \mathbb{N} \text{ and } (z_1,\ldots,z_{n-1}) \in \Omega^\Diamond.$$

We know from Corollary 7 that there is at least one such outcome sequence. It turns out that the stationary forecasting systems that make such ω computably random have a simple characterisation:

Proposition 15 *Consider any ω that is computably random for the forecasting system $\gamma_{\sim 1/2}$. Then for all $I \in \mathscr{C}$, $I \in \mathscr{C}_C(\omega)$ if and only if $\min I < 1/2$ and $\max I > 1/2$.*

This result implies in particular that $L_C(\omega) = [0, 1/2)$, $U_C(\omega) = (1/2, 1]$ and $\underline{p}_C(\omega) = \overline{p}_C(\omega) = 1/2$.

9. Conclusion

Even with the limited number of examples we have been able to examine in this paper, it becomes apparent that incorporating imprecision in the study of randomness allows for much more mathematical structure to arise, which we would argue lets us better understand and place the existing results in the precise limit.

In our argumentation that 'randomness is inherently imprecise', we are well aware that we are restricting ourselves to stationary forecasts. Our examples in Section 8 all involve sequences that are computably random for a precise non-stationary forecasting system, but no longer computably random for any stationary precise variant. To make our claim irrefutable, we would have to show that there are sequences that are computably random for forecasting systems more precise than the vacuous one, but not for any (computable) precise forecasting system. Or in other words, that there is 'randomness' or 'unpredictability' that cannot be 'explained' by any non-stationary (computable) precise forecasting system. We will of course keep this challenge foremost in our minds.

Nevertheless, the examples in Section 8 do indicate that it is in some ways possible to replace an 'explanation' by a complex non-stationary precise forecasting model by a(n infinite filter of) more imprecise stationary one(s).

This work may seem promising, but we are well aware that it is only a humble beginning. We see many extensions in many directions. First of all, we want to find out if our approach can also be used to find interval versions of *Martin-Löf* and *Schnorr randomness* (Ambos-Spies and Kucera, 2000; Bienvenu et al., 2009) with similarly interesting properties and conclusions. Secondly, our preliminary exploration suggests that it will be possible to formulate equivalent randomness definitions in terms of *randomness tests*, rather than supermartingales, but this needs to be checked in much more detail. Thirdly, the approach we follow here is not prequential: we assume that our Forecaster specifies an entire forecasting system γ, or in other words an interval forecast in all possible situations $(x_1, \dots, x_n)$, rather than only interval forecasts in those situations $z_1, \dots, z_n$ of the sequence $\omega = (z_1, \dots, z_n, \dots)$ whose potential randomness we are considering. The *prequential approach*, which we eventually will want to come to, looks at the randomness of a sequence of interval forecasts and outcomes $(I_1, z_1, I_2, z_2, \dots, I_n, z_n, \dots)$, where each I_k is an interval forecast for the as yet unknown X_k, which is afterwards revealed to be z_k, without the need of specifying forecasts in situations that are never reached; see the paper by Vovk and Shen (2010) for an account of how this works for precise forecasts. Fourthly, we need to connect our work with earlier approaches to associating imprecision with randomness (Walley and Fine, 1982; Fierens et al., 2009; Fierens, 2009; Gorban, 2016). And finally, and perhaps most importantly, we believe this research could be a very early starting point for an approach to statistics that takes imprecise or set-valued parameters more seriously, when learning from finite amounts of data.

Acknowledgments

References

K. Ambos-Spies and A. Kucera. Randomness in computability theory. *Contemporary Mathematics*, 257:1–14, 2000.

L. Bienvenu, G. Shafer, and A. Shen. On the history of martingales in the study of randomness. *Electronic Journal for History of Probability and Statistics*, 5, 2009.

A. P. Dawid. The well-calibrated Bayesian. *Journal of The American Statistical Association*, 77 (379):605–610, 1982.

A. P. Dawid. Statistical theory: The prequential approach. *Journal of the Royal Statistical Society, Series A*, 147:278–292, 1984.

A. P. Dawid and V. G. Vovk. Prequential probability: principles and properties. *Bernoulli*, 5: 125–162, 1999.

G. de Cooman and F. Hermans. Imprecise probability trees: Bridging two theories of imprecise probability. *Artificial Intelligence*, 172(11):1400–1427, 2008. doi:10.1016/j.artint.2008.03.001.

G. de Cooman and J. De Bock. Computable randomness is inherently imprecise. ArXiv report 1703.00931 [math.PR], 2017. https://arxiv.org/abs/1703.00931.

G. de Cooman, J. De Bock, and S. Lopatatzidis. Imprecise stochastic processes in discrete time: global models, imprecise Markov chains, and ergodic theorems. *International Journal of Approximate Reasoning*, 76:18–46, 2016.

P. I. Fierens. An extension of chaotic probability models to real-valued variables. *International Journal of Approximate Reasoning*, 50(4):627–641, 2009.

P. I. Fierens, L. C. Rego, and T. L. Fine. A frequentist understanding of sets of measures. *Journal of Statistical Planning and Inference*, 139(6):1879–1892, 2009.

I. I. Gorban. *The Statistical Stability Phenomenon*. Springer, 2016.

M. Li and P. M. B. Vitányi. *An Introduction to Kolmogorov Complexity and Its Applications*. Springer, New York, 1993.

M. B. Pour-El and J. I. Richards. *Computability in Analysis and Physics*. Springer Verlag, 1989.

G. Shafer and V. Vovk. *Probability and Finance: It's Only a Game!* Wiley, New York, 2001.

M. C. M. Troffaes and G. de Cooman. *Lower Previsions*. Wiley, 2014.

J. Ville. *Étude critique de la notion de collectif*. Gauthier-Villars, Paris, 1939.

V. Vovk and G. Shafer. Game-theoretic probability. In T. Augustin, F. P. A. Coolen, G. de Cooman, and M. C. M. Troffaes, editors, *Introduction to Imprecise Probabilities*. John Wiley & Sons, 2014.

V. G. Vovk. On a criterion of randomness. *Doklady Akademii Nauk SSSR*, 294(6):1298–1302, 1987.

V. G. Vovk. Merging of opinions in game-theoretic probability. *Annals of the Institute of Statistical Mathematics*, 61(4):969–993, 2009.

V. G. Vovk and A. Shen. Prequential randomness and probability. *Theoretical Computer Science*, 411(29-30):2632–2646, 2010.

P. Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1991.

P. Walley and T. L. Fine. Towards a frequentist theory of upper and lower probability. *Annals of Statistics*, 10:741–761, 1982.