Imprecise Bernoulli processes

Jasper De Bock and Gert de Cooman

Ghent University, SYSTeMS Research Group
Technologiepark–Zwijnaarde 914, 9052 Zwijnaarde, Belgium.
{jasper.debock,gert.decooman}@UGent.be

Abstract In classical Bernoulli processes, it is assumed that a single Bernoulli experiment can be described by a precise and precisely known probability distribution. However, both of these assumptions can be relaxed. A first approach, often used in sensitivity analysis, is to drop only the second assumption: one assumes the existence of a precise distribution, but has insufficient resources to determine it precisely. The resulting imprecise Bernoulli process is the lower envelope of a set of precise Bernoulli processes. An alternative approach is to drop both assumptions, meaning that we don't assume the existence of a precise probability distribution and regard the experiment as inherently imprecise. In that case, a single imprecise Bernoulli experiment can be described by a set of desirable gambles. We show how this set can be extended to describe an imprecise Bernoulli process, by imposing the behavioral assessments of epistemic independence and exchangeability. The resulting analysis leads to surprisingly simple mathematical expressions characterizing this process, which turn out to be the same as the ones obtained through the straightforward sensitivity analysis approach.

Keywords: imprecise Bernoulli processes, sets of desirable gambles, epistemic independence, exchangeability, sensitivity analysis, Bernstein polynomials, IID processes, exchangeably independent natural extension.

1 Introduction

In classical probability theory, a Bernoulli process is defined as an infinite sequence of binary variables $X_1, \ldots, X_n, \ldots$ that are independent and identically distributed (IID). In this definition, a single Bernoulli experiment is implicitly assumed to have a precise and precisely known probability distribution. However this assumption can be relaxed. A first approach, used in sensitivity analysis, is to assume the existence of a precise probability distribution, but allowing it to be imprecisely known, for example due to limited resources. The resulting imprecise Bernoulli process is then the lower envelope of a set of precise Bernoulli processes. A second approach is to regard a single Bernoulli experiment as inherently imprecise, thereby dropping the assumption an underlying precise probability distribution. In such cases, using sensitivity analysis can no longer be justified and their is no known alternative method that is computationally tractable. In this paper, we offer a solution to this problem by introducing our notion of an *imprecise Bernoulli process*, defining it by imposing the behavioral assessments of exchangeability and epistemic independence. This is a generalisation of the precise-probabilistic definition, since applying our definition to precise distributions, is equivalent with imposing the IID property. We describe our imprecise Bernoulli processes using the language of coherent sets of desirable gambles [3,4,7], because these constitute the most general and powerful imprecise probability models we know of. We give a short introduction to the relevant theory in Section 2. In Section 3, we look at how the marginal model for one variable, describing a single Bernoulli experiment, can be represented as a coherent set of desirable gambles. Section 4 recalls how the assessment of exchangeability can be mathematically formulated in the theory of coherent sets of desirable gambles. In Section 5, we add the assessment of epistemic independence to that of exchangeability and extend the marginal model for a single variable to the smallest (most conservative) imprecise Bernoulli process satisfying those two requirements. We call this the exchangeably independent natural extension of the marginal model. We end by showing in Section 6 that the resulting imprecise Bernoulli process is identical to the one obtained by applying the sensitivity analysis approach mentioned above. This leads us to conclude that an assessment of exchangeability and epistemic independence serves as a behavioural justification for the rather strong assumptions associated with sensitivity analysis.

2 Desirability and coherence

Let us begin by giving a short introduction to the theory of coherent sets of desirable gambles, as it will be an important tool for our analysis. We refer to Refs. [3,4,7] for more details and further discussion. Consider a finite, non-empty set Ω, called the *possibility space*, which describes the possible and mutually exclusive outcomes of some experiment.

Sets of desirable gambles: A gamble f is a real-valued map on Ω which is interpreted as an uncertain reward. If the outcome of the experiment turns out to be ω, the (possibly negative) reward is $f(\omega)$. A non-zero gamble is called desirable if we accept the transaction in which (i) the actual outcome ω of the experiment is determined, and (ii) we receive the reward $f(\omega)$. The zero gamble is not considered to be desirable, mainly because we want desirability to represent a strict preference to the zero gamble.

We will model a subject's beliefs regarding the possible outcomes Ω of an experiment by means of a set $\mathcal{D}$ of desirable gambles, which will be a subset of the set $\mathcal{G}(\Omega)$ of all gambles on Ω. For any two gambles f and g in $\mathcal{G}(\Omega)$, we say that $f \geq g$ if $f(\omega) \geq g(\omega)$ for all ω in Ω and $f > g$ if $f \geq g$ and $f \neq g$.

Coherence: In order to represent a rational subject's beliefs regarding the outcome of an experiment, a set $\mathcal{D} \subseteq \mathcal{G}(\Omega)$ of desirable gambles should satisfy some rationality requirements. If these requirements are met, we call the set $\mathcal{D}$ coherent.

Definition 1 (Coherence). *A set of desirable gambles $\mathcal{D} \subseteq \mathcal{G}(\Omega)$ is called* coherent *if it satisfies the following requirements, for all gambles f, f_1, and f_2 in $\mathcal{G}(\Omega)$ and all real $\lambda > 0$:*

C1. *if $f = 0$ then $f \notin \mathcal{D}$;*
C2. *if $f > 0$ then $f \in \mathcal{D}$;*
C3. *if $f \in \mathcal{D}$ then $\lambda f \in \mathcal{D}$ [scaling];*
C4. *if $f_1, f_2 \in \mathcal{D}$ then $f_1 + f_2 \in \mathcal{D}$ [combination].*

Requirements C3 and C4 make $\mathcal{D}$ a convex cone: $\operatorname{posi}(\mathcal{D}) = \mathcal{D}$, where we have used the positive hull operator posi which generates the set of finite strictly positive linear combinations of elements of its argument set:

$$\operatorname{posi}(\mathcal{D}) := \left\{ \sum_{k=1}^{n} \lambda_k f_k : f_k \in \mathcal{D}, \lambda_k \in \mathbb{R}_0^+, n \in \mathbb{N}_0 \right\}.$$

Here $\mathbb{R}_0^+$ is the set of all positive real numbers, and $\mathbb{N}_0$ the set of all natural numbers (positive integers). The axioms also guarantee that if $f < 0$ then $f \notin \mathcal{D}$.

Weakly desirable gambles: We now define weak desirability, a concept that will lie at the basis of our discussion of exchangeability. Loosely speaking, a gamble is weakly desirable if adding anything desirable to it renders the result desirable.

Definition 2 (Weak desirability). *Consider a coherent set $\mathcal{D}$ of desirable gambles. Then a gamble f is called* weakly desirable *if $f + f'$ is desirable for all desirable f': $f + f' \in \mathcal{D}$ for all f' in $\mathcal{D}$. We use $\mathcal{W}_{\mathcal{D}}$ to denote the set of all weakly desirable gambles associated with $\mathcal{D}$.*

Coherent lower and upper previsions: With a set of gambles $\mathcal{D}$, we can associate a *lower prevision* $\underline{P}_{\mathcal{D}}$ and an *upper prevision* $\overline{P}_{\mathcal{D}}$, which can respectively be interpreted as a lower and upper expectation. For any gambles f we define:

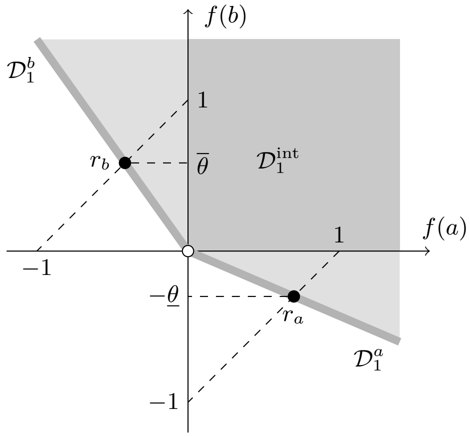
$$\underline{P}_{\mathcal{D}}(f) := \sup\{\mu \in \mathbb{R} : f - \mu \in \mathcal{D}\} \text{ and } \overline{P}_{\mathcal{D}}(f) := \inf\{\mu \in \mathbb{R} : \mu - f \in \mathcal{D}\}. \quad (1)$$

$\underline{P}_{\mathcal{D}}(f)$ is the subject's supremum acceptable price for buying the uncertain reward f, and $\overline{P}_{\mathcal{D}}(f)$ his infimum acceptable price for selling f. Observe that the so-called *conjugacy relation* $\underline{P}_{\mathcal{D}}(-f) = -\overline{P}_{\mathcal{D}}(f)$ is always satisfied. We call a real functional $\underline{P}$ on $\mathcal{G}(\Omega)$ a *coherent lower prevision* if there is some coherent set of desirable gambles $\mathcal{D}$ on $\mathcal{G}(\Omega)$ such that $\underline{P} = \underline{P}_{\mathcal{D}}$.

3 Imprecise Bernoulli experiments

In order for the infinite sequence $X_1, \ldots, X_n, \ldots$ of variables to represent an imprecise Bernoulli process, a necessary requirement is that all individual variables have the same marginal model, describing our subject's uncertainty about a single Bernoulli experiment. In our framework, this is a coherent set of desirable gambles $\mathcal{D}_1$. Let us take a closer look at what this model looks like.

Consider a binary variable X taking values in the set $\mathcal{X} = \{a, b\}$. A gamble f on $\mathcal{X}$ can be identified with a point $(f(a), f(b))$ in two-dimensional Euclidean space. A coherent set of desirable gambles $\mathcal{D}_1$ is a convex cone in this space (the grey area in the figure below), which has to include all gambles $f > 0$ (the dark grey area) but cannot include the zero gamble (the white dot).



Such a cone can be characterised using its extreme rays $\mathcal{D}_1^a$ and $\mathcal{D}_1^b$ (the thick, gray lines in the figure above), which in turn are characterised by the gambles $r_a = (1 - \underline{\theta}, -\underline{\theta})$ and $r_b = (\overline{\theta} - 1, \overline{\theta})$ (the black dots):

$$\mathcal{D}_1^a := \{\lambda_a r_a : \lambda_a > 0\} \text{ and } \mathcal{D}_1^b := \{\lambda_b r_b : \lambda_b > 0\}.$$

It follows from coherence that $0 \leq \underline{\theta} \leq \overline{\theta} \leq 1$.

Since the cone $\mathcal{D}_1$ need not be closed, each of its extreme rays might be included or not. We use δ_a (δ_b) to indicate wether $\mathcal{D}_1^a$ ($\mathcal{D}_1^b$) is included in $\mathcal{D}_1$ or not, by setting it equal to 1 or 0 respectively. Coherence imposes some restrictions on the possible values of δ_a and δ_b. For instance, δ_a must equal 1 if $\underline{\theta} = 0$ and 0 if $\underline{\theta} = 1$. Similarly, δ_b has to be 1 if $\overline{\theta} = 1$ and 0 if $\overline{\theta} = 0$. Finally, δ_a and δ_b cannot both equal 1 if $\underline{\theta} = \overline{\theta}$.

Define $\delta\mathcal{D}_1^a$ to be $\mathcal{D}_1^a$ if $\delta_a = 1$ and to be the empty set $\emptyset$ if $\delta_a = 0$. Analogous definitions hold for $\delta\mathcal{D}_1^b$ and for other sets defined further on. We use $\mathcal{D}_1^{\text{int}}$ to denote the set of all gambles $f \in \mathcal{D}_1$ that are not part of one of the extreme rays $\mathcal{D}_1^a$ or $\mathcal{D}_1^b$ and thus lie in the interior of $\mathcal{D}_1$:

$$\mathcal{D}_1^{\text{int}} := \{\lambda + \lambda_a r_a + \lambda_b r_b : \lambda > 0, \lambda_a \geq 0, \lambda_b \geq 0\}.$$

We can now generally define an arbitrary coherent set of desirable gambles describing our subject's beliefs about a single binary variable as follows:

$$\mathcal{D}_1 := \mathcal{D}_1^{\text{int}} \cup \delta\mathcal{D}_1^a \cup \delta\mathcal{D}_1^b. \tag{2}$$

All that is needed to uniquely determine a set of desirable gambles described by this equation, is the values of $\underline{\theta}$, $\overline{\theta}$, δ_a and δ_b.

4 Exchangeability

A sequence of variables is called exchangeable if, simply put, the order of the variables is irrelevant. In classical Bernoulli processes, exchangeability is a direct consequence of the IID property. In an imprecise-probabilistic context, it turns out that this is not necessarily the case (see, for instance, [1,2] for an approach to IID processes without exchangeability), and we therefore impose exchangeability explicitly as one of the defining properties. This leads to a generalisation of the precise-probabilistic definition, since exchangeability implies that the individual variables have identical marginal models (and are therefore identically distributed in the precise case).

Defining exchangeability in terms of desirable gambles: Consider a finite sequence of variables $X_1, \ldots, X_n$ and an associated set $\mathcal{D}_n$ of desirable gambles on $\mathcal{X}^n$. This sequence assumes values $x = (x_1, \ldots, x_n)$ in $\mathcal{X}^n$. We use $\mathcal{P}_n$ to denote the set of all permutations π of the index set $\{1, \ldots, n\}$. With any such permutation $\pi \in \mathcal{P}_n$, we associate a permutation of $\mathcal{X}^n$, defined by $\pi x = \pi(x_1, \ldots, x_n) := (x_{\pi(1)}, \ldots, x_{\pi(n)})$. Similarly, for any gamble f in $\mathcal{G}(\mathcal{X}^n)$, we define the permuted gamble $\pi^t f = f \circ \pi$, so $(\pi^t f)(x) = f(\pi x)$.

If a subject assessess the sequence $X_1, \ldots, X_n$ to be exchangeable, this means that for any gamble $f \in \mathcal{G}(\mathcal{X}^n)$ and any permutation $\pi \in \mathcal{P}_n$, he is indifferent between the gambles $\pi^t f$ and f, which we translate by saying that he regards exchanging $\pi^t f$ for f as weakly desirable, see [6, Section 4.1.1] and [3] for more motivation and extensive discussion. Equivalently, we require that the gamble $f - \pi^t f$ is weakly desirable.[1] We define $\mathcal{W}_{\mathcal{P}_n} := \{f - \pi^t f : f \in \mathcal{G}(\mathcal{X}^n) \text{ and } \pi \in \mathcal{P}_n\}$.

Definition 3 (Exchangeability). *A coherent set $\mathcal{D}_n$ of desirable gambles on $\mathcal{X}^n$ is called* exchangeable *if all gambles in $\mathcal{W}_{\mathcal{P}_n}$ are weakly desirable: $\mathcal{W}_{\mathcal{P}_n} \subseteq \mathcal{W}_{\mathcal{D}_n}$.*

An infinite sequence of variables $X_1, \ldots, X_n, \ldots$ is called *exchangeable* if each of its finite subsequences is, or equivalently, if for all $n \in \mathbb{N}_0$ the variables $X_1, \ldots, X_n$ are exchangeable. This is modelled as follows: the subject has an exchangeable coherent set of desirable gambles on $\mathcal{X}^n$, for all $n \in \mathbb{N}_0$.

For such a family of sets $\mathcal{D}_n$ of desirable gambles to consistently represent beliefs about an infinite sequence of variables, it should also be *time consistent*. This means that, with $n_1 \leq n_2$, if we consider a gamble h on $\mathcal{X}^{n_2}$ that really only depends on the first n_1 variables, it should not matter, as far as its desirability is concerned, whether we consider it to be a gamble on $\mathcal{X}^{n_1}$ or a gamble on $\mathcal{X}^{n_2}$: $h \in \mathcal{D}_{n_2} \Leftrightarrow h \in \mathcal{D}_{n_1}$. See Ref. [3] for a formal definition of this intuitive property.

As a direct consequence of exchangeability, for any gamble $f \in \mathcal{G}(\mathcal{X}^n)$ and any permutation $\pi \in \mathcal{P}_n$, the gamble $\pi^t f$ is desirable if and only if f is. Limiting ourselves to those permutations in which only the indexes 1 and n are switched, and gambles f that only depend on X_1 or X_n, we see that exchangeability implies that the marginal modal describing X_n is essentially identical to the one describing X_1, and therefore equal to $\mathcal{D}_1$, for all $n \in \mathbb{N}_0$.

[1] We do not require it to be actually desirable, as it can be zero, and the zero gamble is not regarded as desirable.

Representation in terms of polynomials: Consider the set $\mathcal{V}$ of all polynomial functions on $[0,1]$. Subscripting this set with an integer $n \in \mathbb{N}$ means that we limit ourselves to the set of polynomials of degree up to n. The *Bernstein basis polynomials* $B_{k,n}(\theta) := \binom{n}{k}\theta^k(1-\theta)^{n-k}$ form a basis for the linear space $\mathcal{V}_n$ [5]: for each polynomial p whose degree $\deg(p)$ does not exceed n, there is a unique n-tuple $b_p^n = (b_0, b_1, \dots, b_n)$ such that $p = \sum_{k=0}^{n} b_k B_{k,n}(\theta)$. We call a polynomial p *Bernstein positive* if there is some $n \geq \deg(p)$ such that $b_p^n > 0$, meaning that $b_i \geq 0$ for all $i \in \{0, \dots, n\}$ and $b_i > 0$ for at least one $i \in \{0, \dots, n\}$. The set of all Bernstein positive polynomials is denoted by $\mathcal{V}^+$. We are now ready to introduce the concept of Bernstein coherence for polynomials:

Definition 4 (Bernstein coherence). *We call a set $\mathcal{H}$ of polynomials in $\mathcal{V}$* Bernstein coherent *if for all p, p_1, and p_2 in $\mathcal{V}$ and all real $\lambda > 0$:*

B1. *if $p = 0$ then $p \notin \mathcal{H}$;*
B2. *if $p \in \mathcal{V}^+$, then $p \in \mathcal{H}$;*
B3. *if $p \in \mathcal{H}$ then $\lambda p \in \mathcal{H}$;*
B4. *if $p_1, p_2 \in \mathcal{H}$ then $p_1 + p_2 \in \mathcal{H}$.*

With any $\theta \in [0,1]$, we can associate a binary probability mass function on $\mathcal{X} = \{a, b\}$ by letting $\theta_a := \theta$ and $\theta_b := 1 - \theta$. Such a mass function uniquely determines a binomial distribution on $\mathcal{X}^n$. For every sequence of observations $x \in \mathcal{X}^n$, its probability of occurrence is given by $P_\theta(x) := \theta^{C_a(x)}(1-\theta)^{C_b(x)}$, where $C_a(x)$ and $C_b(x)$ respectively denote the number of occurrences of a and b in the sequence x. The expectation associated with the binomial distribution with parameters n and θ is then given by $\mathrm{Mn}_n(f|\theta) := \sum_{x \in \mathcal{X}^n} P_\theta(x) f(x)$, for all gambles f on $\mathcal{X}^n$.

We can now define a linear map Mn_n from $\mathcal{G}(\mathcal{X}^n)$ to $\mathcal{V}$, defining it by $\mathrm{Mn}_n(f) = \mathrm{Mn}_n(f|\cdot)$. In other words, if we let $p = \mathrm{Mn}_n(f)$, then $p(\theta) = \mathrm{Mn}_n(f|\theta)$ for all $\theta \in [0,1]$. To conclude, we let $\mathrm{Mn}_n(\mathcal{D}) := \{\mathrm{Mn}_n(f) : f \in \mathcal{D}\}$ for all $\mathcal{D} \subseteq \mathcal{G}(\mathcal{X}^n)$ and $(\mathrm{Mn}_n)^{-1}(\mathcal{H}) := \{f \in \mathcal{G}(\mathcal{X}^n) : \mathrm{Mn}_n(f) \in \mathcal{H}\}$ for all $\mathcal{H} \subseteq \mathcal{V}$.

Recent work [3] has shown that de Finetti's famous representation result for exchangeable events (binary variables) can be significantly generalised as follows:

Theorem 1 (Infinite Representation). *A family $\mathcal{D}_n$, $n \in \mathbb{N}_0$ of sets of desirable gambles on $\mathcal{X}^n$ is time consistent, coherent and exchangeable if and only if there is some Bernstein coherent set $\mathcal{H}_\infty$ of polynomials in $\mathcal{V}$ such that $\mathcal{D}_n = (\mathrm{Mn}_n)^{-1}(\mathcal{H}_\infty)$ for all $n \in \mathbb{N}_0$. In that case this $\mathcal{H}_\infty$ is uniquely given by $\mathcal{H}_\infty = \bigcup_{n \in \mathbb{N}_0} \mathrm{Mn}_n(\mathcal{D}_n)$.*

We call $\mathcal{H}_\infty$ the *frequency representation* of the coherent, exchangeable and time consistent family of sets of desirable gambles $\mathcal{D}_n$, $n \in \mathbb{N}_0$.

5 Imprecise Bernoulli processes

We now have a way of representing our uncertainty regarding an infinite sequence of variables $X_1, \dots, X_n, \dots$ that we assess to be exchangeable, by means of a frequency representation $\mathcal{H}_\infty$. The only remaining property we have to impose in order to arrive at an imprecise Bernoulli process, is epistemic independence. We call an infinite sequence of variables epistemically independent if learning the value of any finite number of variables does not change our beliefs about any finite subset of the remaining, unobserved ones. It is proven in Ref. [3] that imposing this type of independence on an exchangeable sequence of variables becomes really easy if we use its frequency representation.

Theorem 2 (Independence). *Consider an exchangeable sequence of binary variables $X_1, \ldots, X_n, \ldots$, with frequency representation $\mathcal{H}_\infty$. These variables are epistemically independent if and only if*

$$(\forall k, n \in \mathbb{N}_0 \colon k \leq n)(\forall p \in \mathcal{V}) \; (p \in \mathcal{H}_\infty \Leftrightarrow B_{k,n}p \in \mathcal{H}_\infty). \tag{3}$$

We shall call such models exchangeably independent*.*

It follows that an imprecise Bernoulli process, defined by the properties of exchangeability and epistemic independence, can be described mathematically using a Bernstein coherent set $\mathcal{H}_\infty$ of polynomials that satisfies Eq. (3). By Theorem 1, $\mathcal{H}_\infty$ is equivalent with a time consistent and exhangeable family of coherent sets of desirable gambles $\mathcal{D}_n = (\mathrm{Mn}_n)^{-1}(\mathcal{H}_\infty)$, $n \in \mathbb{N}_0$. In order for this imprecise Bernoulli process to marginalise to a given set of desirable gambles $\mathcal{D}_1$, representing the marginal model we want to extend, we should have that $\mathcal{D}_1 = (\mathrm{Mn}_1)^{-1}(\mathcal{H}_\infty)$, or equivalently that $\mathcal{H}_1 := \mathrm{Mn}_1(\mathcal{D}_1) = \mathcal{H}_\infty \cap \mathcal{V}_1$. We start by investigating what the set of polynomials $\mathcal{H}_1$ looks like.

Polynomial representation of the marginal model: For a given marginal model $\mathcal{D}_1$, the corresponding set of polynomials is given by

$$\mathcal{H}_1 := \mathrm{Mn}_1(\mathcal{D}_1) = \{\mathrm{Mn}_1(f) : f \in \mathcal{D}_1\}, \tag{4}$$

where $\mathrm{Mn}_1(f) = \theta f(a) + (1-\theta) f(b)$. Due to the linearity of the transformation Mn_1, and considering that $\mathrm{Mn}_1(r_a) = \theta - \underline{\theta}$ and $\mathrm{Mn}_1(r_b) = \overline{\theta} - \theta$, it follows from Eqs. (2) and (4) that

$$\mathcal{H}_1 = \mathcal{H}_1^{\mathrm{int}} \cup \delta\mathcal{H}_1^a \cup \delta\mathcal{H}_1^b, \tag{5}$$

where we defined

$$\mathcal{H}_1^{\mathrm{int}} := \{\lambda + \lambda_a(\theta - \underline{\theta}) + \lambda_b(\overline{\theta} - \theta) : \lambda > 0, \lambda_a \geq 0, \lambda_b \geq 0\};$$
$$\mathcal{H}_1^a := \{\lambda_a(\theta - \underline{\theta}) : \lambda_a > 0\}; \tag{6}$$
$$\mathcal{H}_1^b := \{\lambda_b(\overline{\theta} - \theta) : \lambda_b > 0\}. \tag{7}$$

Proposition 1. *$\mathcal{H}_1^{\mathrm{int}}$ is the set of all linear polynomials h in θ that are strictly positive over $[\underline{\theta}, \overline{\theta}]$: $h \in \mathcal{H}_1^{\mathrm{int}} \Leftrightarrow h \in \mathcal{V}_1$ and $h(\theta) > 0$ for all $\theta \in [\underline{\theta}, \overline{\theta}]$.*

The next task is now to find the smallest Bernstein coherent set of polynomials that satisfies Eq. (3) and contains the representation $\mathcal{H}_1$ of a given marginal model $\mathcal{D}_1$. We will call this set the *exchangeably independent natural extension* of $\mathcal{H}_1$.

Polynomial representation of the global model: We start by defining the following set of polynomials:

$$\mathcal{H}_\infty := \mathrm{posi}\{hp : h \in \mathcal{H}_1 \text{ and } p \in \mathcal{V}^+\}, \tag{8}$$

which will be closely related to the sets

$$\begin{aligned} \mathcal{H}_\infty^{\mathrm{int}} &:= \mathrm{posi}\{hp : h \in \mathcal{H}_1^{\mathrm{int}} \text{ and } p \in \mathcal{V}^+\}; \\ \mathcal{H}_\infty^{a} &:= \mathrm{posi}\{hp : h \in \mathcal{H}_1^{a} \text{ and } p \in \mathcal{V}^+\} = \{(\theta - \underline{\theta})p : p \in \mathcal{V}^+\}; \qquad (9) \\ \mathcal{H}_\infty^{b} &:= \mathrm{posi}\{hp : h \in \mathcal{H}_1^{b} \text{ and } p \in \mathcal{V}^+\} = \{(\overline{\theta} - \theta)p : p \in \mathcal{V}^+\}. \qquad (10) \end{aligned}$$

For $\mathcal{H}_1^a$ and $\mathcal{H}_1^b$, the alternative characterisations that are given above are easy to prove. For $\mathcal{H}_\infty^{\mathrm{int}}$, finding an alternative characterisation turns out to be more involved.

Theorem 3. *The following statements are equivalent:*

(i) $h \in \mathcal{H}_\infty^{\mathrm{int}}$*;*
(ii) $(\exists \epsilon > 0)(\forall \theta \in [\underline{\theta} - \epsilon, \overline{\theta} + \epsilon] \cap (0,1))\ \ h(\theta) > 0$*;*
(iii) $h = ph'$ *for some* $p \in \mathcal{V}^+$ *and* h' *such that* $(\forall \theta \in [\underline{\theta}, \overline{\theta}])\ \ h'(\theta) > 0$*.*

When both $\underline{\theta} \neq 1$ *and* $\overline{\theta} \neq 0$*, these tree statements are also equivalent with:*

(iv) $(\forall \theta \in [\underline{\theta}, \overline{\theta}] \setminus \{0,1\})\ \ h(\theta) > 0$.

It turns out that the set of polynomials $\mathcal{H}_\infty$ is indeed very closely related to the sets $\mathcal{H}_\infty^{\mathrm{int}}$, $\mathcal{H}_\infty^{a}$ and $\mathcal{H}_\infty^{b}$, since instead of using Eq. (8), $\mathcal{H}_\infty$ can be equivalently characterised by

$$\mathcal{H}_\infty = \mathcal{H}_\infty^{\mathrm{int}} \cup \delta\mathcal{H}_\infty^{a} \cup \delta\mathcal{H}_\infty^{b}. \tag{11}$$

We are now ready to formulate the most important result of this paper, which says that a set $\mathcal{H}_\infty$, given by Eq. (8), represents an imprecise Bernoulli process.

Theorem 4. $\mathcal{H}_\infty$ *is the smallest Bernstein coherent superset of* $\mathcal{H}_1$ *that satisfies the epistemic independence condition* (3)*.*

This means that the set $\mathcal{H}_\infty$ given by Eq. (8) is the exchangeably independent natural extension of $\mathcal{H}_1$. It follows from Theorem 1 that $\mathcal{H}_\infty$ represents an imprecise Bernoulli proces that marginalises to $\mathcal{D}_1$ if $\mathcal{D}_1 = (\mathrm{Mn}_1)^{-1}(\mathcal{H}_\infty)$. This is equivalent to demanding that $\mathcal{H}_\infty$ should contain no other polynomials in $\mathcal{V}_1$ than those in $\mathcal{H}_1$. Due to Eq. (11), it suffices to check this property separately for each of the three subsets of $\mathcal{H}_\infty$. For $\mathcal{H}_\infty^{a}$ and $\mathcal{H}_\infty^{b}$ this property follows from Eqs. (5)–(7) and (9)–(10). For $\mathcal{H}_\infty^{\mathrm{int}}$, it follows from Proposition 1, Theorem 3 and Eqs. (5)–(7). We conclude that $\mathcal{H}_\infty$ is the smallest (most conservative) representation of an imprecise Bernoulli process that marginalises to a set of desirable gambles $\mathcal{D}_1$ that is given by Eq. (2).

6 Justifying a sensitivity analysis approach

It now only remains to explain how the results obtained above relate to our previous statement that a sensitivity analysis approach to dealing with imprecise Bernoulli processes can be justified using an assessment of epistemic independence and exchangeability.

Consider an arbitrary gamble $f \in \mathcal{G}(\mathcal{X}^n)$, $n \in \mathbb{N}_0$ and a probability $\theta \in [0,1]$ that characterises a probability mass function on $\mathcal{X} = \{a, b\}$. As is shown in Section 4, $\mathrm{Mn}_n(f|\theta)$ is the expectation of f, associated with the binomial distribution with parameters n and θ. If one defines an imprecise Bernoulli process using the sensitivity analysis approach, this results in letting θ vary over an interval $[\underline{\theta}, \overline{\theta}]$ and the lower (upper) expectation of f associated with such an imprecise Bernoulli process is then the minimum (maximum) of $\mathrm{Mn}_n(f|\theta)$ as θ ranges over this interval.We will now show that this intuitive result is also obtained using the type of imprecise Bernoulli process we considered in the previous sections.

Theorem 5. *Consider the set of polynomials $\mathcal{H}_\infty$ defined by Eq.* (8)*. Then for any polynomial function p on* $[0,1]$*:*

$$\underline{P}_{\mathcal{H}_\infty}(p) := \sup\{\mu \in \mathbb{R} : p - \mu \in \mathcal{H}_\infty\} = \min\{p(\theta) : \theta \in [\underline{\theta}, \overline{\theta}]\}. \tag{12}$$

By Theorem 1 and Eq. (1), the lower prevision (or minimum expected value) of a gamble $f \in \mathcal{G}(\mathcal{X}^n)$, $n \in \mathbb{N}_0$, corresponding with an imprecise Bernoulli process represented by a Bernstein coherent set $\mathcal{H}_\infty$ of polynomials, is given by

$$\begin{aligned}\underline{E}(f) := \underline{P}_{\mathcal{D}_n}(f) &= \underline{P}_{(\mathrm{Mn}_n)^{-1}(\mathcal{H}_\infty)}(f) \\ &= \sup\{\mu \in \mathbb{R} : f - \mu \in (\mathrm{Mn}_n)^{-1}(\mathcal{H}_\infty)\} \\ &= \sup\{\mu \in \mathbb{R} : \mathrm{Mn}_n(f) - \mu \in \mathcal{H}_\infty\} = \underline{P}_{\mathcal{H}_\infty}(\mathrm{Mn}_n(f)),\end{aligned}$$

thereby implying the following de Finetti-like representation result for lower previsions: $\underline{P}_{\mathcal{D}_n} = \underline{P}_{\mathcal{H}_\infty} \circ \mathrm{Mn}_n$. Using Theorem 5, we find that

$$\underline{E}(f) = \min\{\mathrm{Mn}_n(f|\theta) : \theta \in [\underline{\theta}, \overline{\theta}]\},$$

which is exactly what we would get using the sensitivity analysis approach. Notice also that $\overline{E}(f) := \overline{P}_{(\mathrm{Mn}_n)^{-1}(\mathcal{H}_\infty)}(f) = -\underline{E}(-f)$ because of the conjugacy relation between lower and upper previsions. As a direct consequence, we find that, similarly:

$$\overline{E}(f) = \max\{\mathrm{Mn}_n(f|\theta) : \theta \in [\underline{\theta}, \overline{\theta}]\}.$$

7 Conclusions

The existence of a precise probability distribution describing the outcomes of a single Bernoulli experiment is not crucial to the definition of a Bernoulli process. It can be relaxed by replacing it with an assessment of exchangeability, which means that we consider the order of the different Bernoulli experiments to be irrelevant. Taken together with epistemic independence, exchangeability then becomes a defining property for an imprecise Bernoulli process. Using this approach, we have derived an expression for the most conservative imprecise Bernoulli process, corresponding with a given marginal model. The resulting imprecise Bernoulli process is exactly the same as the one obtained using a sensitivity analysis approach. An assessment of exchangeability and epistemic independence can therefore be used as a behavioural justification for the strong assumptions associated with the latter approach.

Although we have not discussed this here, we have also looked at how to make multinomial processes imprecise, and we are confident that our results for binomial processes can be generalised. We will report these results elsewhere, together with proofs for (generalisations of) the theorems mentioned above.

We have used the very general theory of sets of desirable gambles to develop our notion of an imprecise Bernoulli process. The important sensitivity analysis result at the end, however, is stated purely in terms of lower and upper previsions, which constitute a less general model than sets of desirable gambles. It would be interesting to see if and how this result can be obtained directly using the language of previsions, without using sets of desirable gambles.

Given the importance of binomial (and multinomial) processes in practical statistics, we hope that our results can lead to a better understanding, and perhaps to much needed practical applications, of imprecise probability theory in statistics.

Acknowledgments. Jasper De Bock is a Ph.D. Fellow of the Fund for Scientific Research – Flanders (FWO) and wishes to acknowledge the financial support of the FWO. Research for this paper was also partially funded by the IWT SBO project 60043, and by FWO project G.0125.12.

References

1. De Cooman, G., Miranda, E.: Weak and strong laws of large numbers for coherent lower previsions. Journal of Statistical Planning and Inference 138(8), 2409–2432 (2008)
2. De Cooman, G., Miranda, E.: Forward irrelevance. Journal of Statistical Planning and Inference 139(2), 256–276 (2009)
3. De Cooman, G., Quaeghebeur, E.: Exchangeability and sets of desirable gambles. International Journal of Approximate Reasoning (2010), in print. Special issue in honour of Henry E. Kyburg, Jr.
4. Couso, I., Moral, S.: Sets of desirable gambles: conditioning, representation, and precise probabilities. International Journal of Approximate Reasoning 52(7), 1034–1055 (2011)
5. Prautzsch, H., Boehm, W., Paluszny, M.: Bézier and B-Spline Techniques. Springer, Berlin (2002)
6. Walley, P.: Statistical Reasoning with Imprecise Probabilities, Monographs on Statistics and Applied Probability, vol. 42. Chapman and Hall, London (1991)
7. Walley, P.: Towards a unified theory of imprecise probability. International Journal of Approximate Reasoning 24, 125–148 (2000)