Cournot's principle revisited

Galvan Bruno

Abstract. Cournot's principle states that a typical event (i.e., an event with probability very close to 1) occurs nearly certainly in a single trial of an experiment. This principle has been considered by various authors as the only connection between mathematical probability and the real world of experiments.

To make the logical structure of the principle clearer, in this paper a reformulation of the principle is proposed. This reformulation is based on the following three elements: (1) The explicit definition of the empirical property of practical certainty, (2) the clear separation between probability measure and experiment, including the remark that typicality is a mathematical property defined by the probability measure while practical certainty is an empirical property defined by the experiment, and (3) the explicit formulation of the product rule for independent trials.

The novel formulation then states that a probability measure P *governs* an experiment E if the events that are typical according to P^n are practically certain according to E^n for all $n \geq 1$, where P^n is the n-fold product of P and E^n is the experiment whose trials are composed of n trials of E.

The novel formulation highlights the possible existence of two ambiguities in the principle, namely: (i) that different probability measures govern the same experiment and (ii) that the same probability measure governs different experiments. In this paper the first ambiguity is rigorously disproved, while the second is disproved provided that a suitable property characterizing the empirical equivalence of experiments is assumed.

Key words and phrases: Cournot's principle, Typicality, Practical certainty, Interpretation of probability.

1. INTRODUCTION

Cournot's principle states that a typical event (i.e., an event with probability very close to 1) singled out in advance occurs nearly certainly in a single trial of an experiment.

This principle was first formulated by Jacob Bernoulli in his *Ars Conjectandi* (1713) to derive the correspondence between the probability and the relative frequency of an event. However, Augustin Cournot seems to have been the first to say explicitly (1843) that the whole empirical meaning of classical probability derives from this principle. In the first half of the 20th century Borel, Lévy, and Kolmogorov all subscribed to Cournot's principle. See [7] for an extensive presentation of Cournot's principle and of its history.

In more recent years the principle has been mentioned in various papers by Goldstein et al. about the role of typicality in statistical and Bohmian mechanics [3, 4, 1, 2, 5].

To make the logical structure of the principle clearer, in this paper a reformulation of the principle is proposed. This reformulation is based on the following three elements: (1) The explicit definition of the empirical property of practical certainty, (2) the clear separation between probability measure and experiment, and (3) the explicit formulation of the product rule for independent trials. Let us explain.

(1) Practical certainty. Usually we say that an event that occurs nearly certainly in a single trial of an experiment is *practically certain*[1] (Bernoulli called these events *morally certain*). The idea is to explicitly recognize practical certainty as an objective empirical property of some events of an experiment. An operational definition of practical certainty is therefore proposed in this paper, and some of its properties are identified.

via Milano 60, 38122 Trento (Italy), (e-mail: b.galvan33@gmail.com).

[1] The term "almost certain" is also used in everyday language. We prefer not to use this term here because in mathematical probability it has a different meaning and this may generate some confusion.

Given the definition of practical certainty Cournot's principle can be reformulated as follows:

A typical event is practically certain.

(2) Separation between probability and experiment. Cournot's principle has been formulated in the context of classical probability. According to this approach a probability measure is naturally associated with an experiment, namely, classical probability, which is the ratio between favorable and possible outcomes. The situation is different in the modern measure-theoretic approach to probability, in which probability spaces and experiments are separate entities that must be related in some way, e.g., by Cournot's principle. This separation helps to clarify the different nature of typicality and practical certainty: the former is a mathematical property defined by the probability measure while the latter is an empirical property defined by the experiment.

To implement the separation between probability space and experiment in Cournot's principle let us reformulate it as follows. Let P and E be a probability measure and an experiment with the same event space, respectively. Typical events are defined by P while practically certain events are defined by E. Let $\mathcal{T}(P)$ and $\mathcal{C}(E)$ denote the classes of typical and of practically certain events, respectively. These classes are vaguely defined but, for simplicity, in this introductory section they are considered as exactly defined; the problem of vagueness is properly taken into consideration in the main part of the paper.

Cournot's principle can then be formulated as follows:

> A probability measure P *governs* an experiment E if $\mathcal{T}(P) \subseteq \mathcal{C}(E)$.

The verb "governs" may be replaced for example by "is a probabilistic model of", "represents", etc.

(3) Product rule for independent trials. The above formulation is not yet sufficient, for the following reason. We know that the classical probability of a sequence of events in a sequence of independent trials is the product of their probabilities. In the context of measure-theoretic probability this fact must be explicitly stated. This can be done by declaring that P governs E if P^n governs E^n for all $n \geq 1$, where P^n is the n-fold product of P and E^n is the experiment whose trials are composed of n trials of E. In other words, the condition $\mathcal{T}(P) \subseteq \mathcal{C}(E)$ must be extended to the condition $\mathcal{T}^{(n)}(P) \subseteq \mathcal{C}^{(n)}(E)$ for all $n \geq 1$, where $\mathcal{T}^{(n)}(P)$ and $\mathcal{C}^{(n)}(E)$ denote the classes of typical events defined by P^n and of practically certain events defined by E^n, respectively.

The extended condition is arguably an idealization (see for example the discussion in [7], at the end of pag. 91). Nevertheless, it is adopted in this paper because it allows us to highlight the logical structure of Cournot's principle and to deduce some rigorous results. Moreover, its study is certainly a prerequisite for a possible future study of more realistic cases.

A compact notation is obtained by introducing the classes

$$\bar{\mathcal{T}}(P) := \cup_{n=1}^{\infty} \mathcal{T}^{(n)}(P) \text{ and } \bar{\mathcal{C}}(E) := \cup_{n=1}^{\infty} \mathcal{C}^{(n)}(E).$$

With this notation the novel version of the principle, which is the definitive one[2], is the following:

COURNOT'S PRINCIPLE. A probability measure P governs an experiment E if

$$\bar{\mathcal{T}}(P) \subseteq \bar{\mathcal{C}}(E). \tag{1}$$

This formulation of Cournot's principle highlights the possible existence of two ambiguities in the principle. The first ambiguity, which we call *probabilistic ambiguity*, arises if there are two (or more) different probability measures governing the same experiment. In formal terms, this happens if

$$\bar{\mathcal{T}}(P_1) \cup \bar{\mathcal{T}}(P_2) \subseteq \bar{\mathcal{C}}(E) \tag{2}$$

for two different probability measures P_1 and P_2.

The second possible ambiguity, which we call *experimental ambiguity*, arises if there are two (or more) empirically distinguishable experiments governed by the same probability measure. In formal terms, this happens if

$$\bar{\mathcal{T}}(P) \subseteq \bar{\mathcal{C}}(E_1) \cap \bar{\mathcal{C}}(E_2) \tag{3}$$

for two empirically distinguishable experiments E_1 and E_2.

In this paper, it is proven that probabilistic ambiguity does not occur because one can prove the implication

$$\bar{\mathcal{T}}(P_1) \cup \bar{\mathcal{T}}(P_2) \subseteq \bar{\mathcal{C}}(E) \Rightarrow P_1 = P_2. \tag{4}$$

Moreover, by assuming that a suitable condition characterizes the empirical equivalence of two experiments, one can also prove the implication

$$\bar{\mathcal{T}}(P) \subseteq \bar{\mathcal{C}}(E_1) \cap \bar{\mathcal{C}}(E_2) \Rightarrow E_1 \sim E_2, \tag{5}$$

where $E_1 \sim E_2$ means that E_1 and E_2 are empirically equivalent (indistinguishable). This implication excludes experimental ambiguity.

The remainder of this paper is organized as follows. In Section 2 some purely mathematical notions are developed. In Section 3 the novel formulation of Cournot's principle is developed into the details and the announced results are proven. Section 4 concludes the paper.

[2]This formulation is, however, a simplified version of the complete formulation presented in Section 3, because here the problem of vagueness has been ignored.

2. MATHEMATICS FOR COURNOT'S PRINCIPLE

In this section some mathematical notions are developed. They are utilized in the next section to reformulate Cournot's principle and to prove the announced results.

2.1 Preliminary notions

The notions of sample space and event space (the σ-algebra of the events) are well known. Let Ω and $\mathcal{A}$ denote these two entities, respectively. The symbol $\mathcal{A}^n$ denotes the event space of Ω^n generated by the measurable rectangles of Ω^n.

Let us define the *extended* event space as

$$\bar{\mathcal{A}} := \cup_{n=1}^{\infty} \mathcal{A}^n. \tag{6}$$

Generic events of $\bar{\mathcal{A}}$ are denoted by $A, B, \ldots$, while the notation $A^{(n)}, B^{(n)}, \ldots$ is adopted to specify that the events belong to $\mathcal{A}^n$. If $\bar{\mathcal{C}} \subseteq \bar{\mathcal{A}}$, let us call *components* of $\bar{\mathcal{C}}$ the classes of the type $\mathcal{C}^{(n)} := \bar{\mathcal{C}} \cap \mathcal{A}^n$.

To help keep the notions clear, sets of events are called *classes* and sets of classes are called *superclasses*.

Probability measures and probability spaces are also well-known notions. The symbols P, P_1, and P_2 always denote probability measures on the event space $\mathcal{A}$. The symbol P^n denotes the usual n-fold product of P, i.e., the unique probability measure on $\mathcal{A}^n$ such that $P^n(A_1 \times \cdots \times A_n) = P(A_1) \cdots P(A_n)$ for all $A_1, \ldots, A_n \in \mathcal{A}$.

Let us define the *extended* probability $\bar{P} : \bar{\mathcal{A}} \to [0,1]$ as

$$\bar{P}(A^{(n)}) := P^n(A^{(n)}), \tag{7}$$

and the class

$$\bar{\mathcal{T}}(P, \delta) := \{A \in \bar{\mathcal{A}} : \bar{P}(A) \geq \delta\}, \tag{8}$$

where $\delta \in [0,1]$.

Let us now define a type of events that is used often in the following. For $A \in \bar{\mathcal{A}}$, I interval of $[0,1]$, and $k \in \mathbb{N}^+$, let us define the event $S(A, I, k) \in \bar{\mathcal{A}}$ as follows:

$$S(A^{(n)}, I, k) := \left\{ (\omega_1^{(n)}, \ldots, \omega_k^{(n)}) \in \Omega^{n \times k} : \frac{1}{k} \sum_{i=1}^{k} \chi_{A^{(n)}}(\omega_i^{(n)}) \in I \right\}, \tag{9}$$

where $\omega_i^{(n)} \in \Omega^n$ and $\chi_{A^{(n)}}$ is the characteristic function of $A^{(n)}$. In other words, the event $S(A^{(n)}, I, k)$ contains all the elements $(\omega_1^{(n)}, \ldots, \omega_k^{(n)}) \in \Omega^{n \times k}$ for which the relative frequency of $A^{(n)} \in \mathcal{A}^n$ in the sequence $(\omega_1^{(n)}, \ldots, \omega_k^{(n)})$ belongs to I.

The set $S(A, I, k)$ is measurable and, if I_1 and I_2 are disjoint, then the events $S(A, I_1, k)$ and $S(A, I_2, k)$ are disjoint. To see this, it is sufficient to note that $S(A^{(n)}, I, k)$ is the inverse image of I under the $\mathcal{A}^{n \times k}$-measurable function

$$f_{A^{(n)}}(\omega_1^{(n)}, \ldots, \omega_k^{(n)}) := \frac{1}{k} \sum_{i=1}^{k} \chi_{A^{(n)}}(\omega_i^{(n)}) \tag{10}$$

from $\Omega^{n \times k}$ to $\mathbb{R}$.

In the next subsection the following two limits are used:

$$\lim_{k \to \infty} \bar{P}[S(A, [\sigma, 1], k)] = \begin{cases} 1 \text{ if } \sigma < \bar{P}(A), \\ 0 \text{ if } \sigma > \bar{P}(A), \end{cases} \tag{11}$$

$$\lim_{k \to \infty} \bar{P}[S(A, [0, \sigma), k)] = \begin{cases} 0 \text{ if } \sigma < \bar{P}(A), \\ 1 \text{ if } \sigma > \bar{P}(A). \end{cases} \tag{12}$$

where $\sigma \in (0,1]$. These limits easily follow from Bernoulli's theorem (i.e., the weak law of large number; see for example [6]).

If $\bar{\mathcal{C}} \subseteq \bar{\mathcal{A}}$, we use the notation

$$S(A, I, k) \in_d \bar{\mathcal{C}}$$

to represent the fact that $S(A, I, k) \in \bar{\mathcal{C}}$ *definitively*, i.e., for any k greater than a suitable k_0. For example, according to limit (11), $S(A, [\sigma, 1], k) \in_d \bar{\mathcal{T}}(P, \delta)$ for $\delta < 1$ and $\sigma < \bar{P}(A)$.

The last notation: the expression $\delta \approx 1$ means

$$1 - \epsilon \leq \delta \leq 1$$

for some $\epsilon \ll 1$.

2.2 Novel notions

In this subsection some mathematical notions specifically related to Cournot's principle are introduced.

The first notion is that of C-class, which represents the structure of the classes of typical and practically certain events:

DEFINITION (C-class). A class $\bar{\mathcal{C}} \subseteq \bar{\mathcal{A}}$ is said to be a *C-class* if, for all $n \geq 1$, the component $\mathcal{C}^{(n)} := \bar{\mathcal{C}} \cap \mathcal{A}^n$ satisfies the following properties:

(a) $\Omega^n \in \mathcal{C}^{(n)}$;
(b) if $A^{(n)} \in \mathcal{C}^{(n)}$ and $A^{(n)} \subseteq B^{(n)}$, then $B^{(n)} \in \mathcal{C}^{(n)}$;
(c) $\mathcal{C}^{(n)}$ does not contain two disjoint events.

Two examples of C-class are: $\bar{\mathcal{T}}(P, \delta)$ for $\delta > \frac{1}{2}$ and the *trivial* C-class $\{\Omega, \Omega^2, \Omega^3, \ldots\}$.

Hereafter, the symbols $\bar{\mathcal{C}}, \bar{\mathcal{C}}_1$, and $\bar{\mathcal{C}}_2$ denote C-classes contained in $\bar{\mathcal{A}}$.

The second notion is that of C-measure, which is used to define an equivalence relation between C-classes. In this paper this measure is used only as a mathematical tool. However, see the end of this subsection for a discussion about a possible interpretation of the C-measure.

DEFINITION (C-measure). The *C-measure* generated by a C-class $\bar{\mathcal{C}}$ is the set function $\bar{M}_{\bar{\mathcal{C}}} : \bar{\mathcal{A}} \to [0,1]$ defined as follows:

$$\bar{M}_{\bar{\mathcal{C}}}(A) := \sup\{\sigma \in [0,1] : S(A, [\sigma, 1], k) \in_d \bar{\mathcal{C}}\}. \tag{13}$$

Sometime the subscript $\bar{\mathcal{C}}$ of $\bar{M}_{\bar{\mathcal{C}}}$ is omitted if this does not create ambiguity.

We note that $\bar{M}_{\bar{\mathcal{C}}}$ does not depend of any finite set of components of $\bar{\mathcal{C}}$. Unlike a probability measure, the C-measure is not necessarily additive (see Example 1). However, the C-measure $\bar{M}$ generated by a generic C-class $\bar{\mathcal{C}}$ satisfies the following properties:

PROPOSITION 1. *For all $n \geq 1$ we have:*

(a) $\bar{M}(\emptyset^{(n)}) = 0$*;*
(b) $\bar{M}(\Omega^n) = 1$*;*
(c) $\bar{M}(A^{(n)}) \leq \bar{M}(B^{(n)})$ *if* $A^{(n)} \subseteq B^{(n)}$*.*

PROOF. (a) $S(\emptyset^{(n)}, [0,1], k) = \Omega^{n\times k} \in \bar{\mathcal{C}}$ and $S(\emptyset^{(n)}, [\sigma,1], k) = \emptyset^{(n\times k)} \notin \bar{\mathcal{C}}$ for $\sigma \in (0,1]$, which implies $\bar{M}(\emptyset^{(n)}) = 0$.

(b) $S(\Omega^n, [\sigma,1], k) = \Omega^{n\times k} \in \bar{\mathcal{C}}$ for all $\sigma \in [0,1]$, which implies $\bar{M}(\Omega^n) = 1$.

(c) $A^{(n)} \subseteq B^{(n)}$ implies

$$S(A^{(n)}, [\sigma,1], k) \subseteq S(B^{(n)}, [\sigma,1], k),$$

so that

$$S(A^{(n)}, [\sigma,1], k) \in_d \bar{\mathcal{C}} \Rightarrow S(B^{(n)}, [\sigma,1], k) \in_d \bar{\mathcal{C}},$$

which implies $\bar{M}(A^{(n)}) \leq \bar{M}(B^{(n)})$. □

Let us give two examples of C-measures.

EXAMPLE 1. Let $\bar{\mathcal{C}} := \{\Omega, \Omega^2, \Omega^3, \dots\}$ be the trivial C-class. Then:

$$\bar{M}_{\bar{\mathcal{C}}}(A^{(n)}) = \begin{cases} 0 & \text{for } A^{(n)} \neq \Omega^n, \\ 1 & \text{for } A^{(n)} = \Omega^n. \end{cases} \tag{14}$$

PROOF. If $A^{(n)} \neq \Omega^n$, we have:

$$S(A^{(n)}, [0,1], k) = \Omega^{n\times k},$$

$$S(A^{(n)}, [\sigma,1], k) \neq \Omega^{n\times k} \text{ for } \sigma \in (0,1],$$

$$S(\Omega^n, [\sigma,1], k) = \Omega^{n\times k} \text{ for } \sigma \in [0,1],$$

from which equation (14) easily follows. □

EXAMPLE 2. Let $\delta \in (\frac{1}{2}, 1)$. Then the C-measure generated by $\bar{\mathcal{T}}(P,\delta)$ is $\bar{P}$.

PROOF. Let $\bar{M}$ denote the C-measure generated by $\bar{\mathcal{T}}(P,\delta)$, and let $A \in \bar{\mathcal{A}}$. If $\sigma < \bar{P}(A)$, then $S(A, [\sigma,1], k) \in_d \bar{\mathcal{T}}(P,\delta)$ due to limit (11), so that $\bar{M}(A) \geq \bar{P}(A)$. On the contrary, if $\sigma > \bar{P}(A)$, then $S(A, [\sigma,1], k) \notin_d \bar{\mathcal{T}}(P,\delta)$, so that $\bar{M}(A) \leq \bar{P}(A)$. In conclusion, $\bar{M}(A) = \bar{P}(A)$. □

The C-measure allows us to define the following equivalence relation:

DEFINITION. We say that two C-classes $\bar{\mathcal{C}}_1, \bar{\mathcal{C}}_2$ are *(asymptotically) equivalent* if $\bar{M}_{\bar{\mathcal{C}}_1} = \bar{M}_{\bar{\mathcal{C}}_2}$. In this case we write

$$\bar{\mathcal{C}}_1 \sim \bar{\mathcal{C}}_2. \tag{15}$$

The equivalence relation $\sim$ induces a partition in the superclass of the C-classes contained in $\bar{\mathcal{A}}$. We call *equivalence superclasses* the elements of this partition.

All the results relative to Cournot's principle presented in the next section are based on the following theorem:

THEOREM 1. *Let* $\delta \in (\frac{1}{2}, 1)$*. Then:*

$$\bar{\mathcal{T}}(P,\delta) \subseteq \bar{\mathcal{C}} \Rightarrow \bar{\mathcal{T}}(P,\delta) \sim \bar{\mathcal{C}}. \tag{16}$$

PROOF. Let $\bar{M}$ denote the C-measure generated by $\bar{\mathcal{C}}$. Since $\bar{P}$ is the C-measure generated by $\bar{\mathcal{T}}(P,\delta)$ (see Example 2) we have to prove that $\bar{M} = \bar{P}$.

The inequality $\bar{P}(A) \leq \bar{M}(A)$ follows trivially from the fact that $S(A, [\sigma,1], k) \in_d \bar{\mathcal{T}}(P,\delta) \Rightarrow S(A, [\sigma,1], k) \in_d \bar{\mathcal{C}}$ (because $\bar{\mathcal{T}}(P,\delta) \subseteq \bar{\mathcal{C}}$).

Let us prove now that $\bar{M}(A) \leq \bar{P}(A)$, or equivalently, that $\bar{M}(A) > \bar{P}(A)$ leads to a contradiction. Let us assume therefore the latter inequality, and let σ belong to the interval $(\bar{P}(A), \bar{M}(A))$. From the definition of $\bar{M}$ it follows that $S(A, [\sigma,1], k) \in_d \bar{\mathcal{C}}$, while from limit (12) it follows that $S(A, [0,\sigma), k) \in_d \bar{\mathcal{T}}(P,\delta) \subseteq \bar{\mathcal{C}}$. This implies that both disjoint events $S(A, [0,\sigma), k)$ and $S(A, [\sigma,1], k)$ belong to $\bar{\mathcal{C}}$ for k large enough, which is impossible because $\bar{\mathcal{C}}$ is a C-class. □

This theorem has many corollaries, for example:

COROLLARY 1. *Let* $\delta, \delta_1, \delta_1 \in (\frac{1}{2}, 1)$*. Then:*

(a) $\bar{\mathcal{T}}(P_1,\delta_1) \cup \bar{\mathcal{T}}(P_2,\delta_2) \subseteq \bar{\mathcal{C}} \Rightarrow P_1 = P_2$*;*
(b) $\bar{\mathcal{T}}(P,\delta) \subseteq \bar{\mathcal{C}}_1 \cap \bar{\mathcal{C}}_2 \Rightarrow \bar{\mathcal{C}}_1 \sim \bar{\mathcal{C}}_2$*.*

These and other corollaries can be easily proved by using the transitive property of $\sim$ and by considering that the C-measure generated by $\bar{\mathcal{T}}(P,\delta)$ is $\bar{P}$.

REMARK. As previously mentioned, in this paper the C-measure is considered only as a mathematical tool. However, with some adjustment, it could be arguably interpreted as a measure of certainty, something like the non-probabilistic typicality measure evoked by Goldstein [4]. Let us explain.

In Subsection 3.4 a practically certain event is defined as an event whose long-run relative frequency is close to 1. Suppose that a suitable C-class $\bar{\mathcal{C}}$ represents the class of practically certain events of some experiment, and that $\bar{M}_{\bar{\mathcal{C}}}(A) \approx 1$ for some event A. This means that $S(A, [\sigma,1], k) \in_d \bar{\mathcal{C}}$ for some $\sigma \approx 1$, i.e., $S(A, [\sigma,1], k)$ is practically certain for k large enough. As a consequence (nearly certainly) the long-run relative frequency of A is close to 1, and therefore, A is practically certain as well. In other words, $\bar{M}_{\bar{\mathcal{C}}}(A)$ measures the degree of certainty of A.

However, the above interpretation has the following drawback: even if A is practically certain because its long-run relative frequency is close to 1, we have no guarantee that $A \in \bar{\mathcal{C}}$, as expected. In order to interpret $\bar{M}_{\bar{\mathcal{C}}}$ as a measure of certainty it is necessary that $\bar{\mathcal{C}}$ and $\bar{M}_{\bar{\mathcal{C}}}$ satisfy some further property that removes this drawback, for example the following property:

$$\bar{M}_{\bar{\mathcal{C}}}(A) \geq \inf\{\bar{M}_{\bar{\mathcal{C}}}(B) : B \in \bar{\mathcal{C}}\} \Rightarrow A \in \bar{\mathcal{C}}.$$

We note that the C-class $\bar{\mathcal{T}}(P,\delta)$ and the associated C-measure $\bar{P}$ satisfy this property.

This subject is left for possible future research and is not further considered in this paper.

3. COURNOT'S PRINCIPLE REVISITED

In this more conceptual section, the notions of typical and practically certain events are introduced and the novel formulation of Cournot's principle is presented.

3.1 The problem of vagueness

Typical and practically certain events are vaguely defined classes of events. Therefore, they must be managed with some caution. For this purpose, we adopt the strategy of representing a vague class with the vague superclass of its instances. Let us explain.

Let $\bar{\mathcal{V}}$ denote a vague class of events of $\bar{\mathcal{A}}$. An *instance* of $\bar{\mathcal{V}}$ is an *exact* class of events that can be considered as an acceptable exact version of the vague class $\bar{\mathcal{V}}$. A vague class has many instances, and the superclass of the instances, which we denote by $\mathfrak{V}$, is a vague superclass.

Representing a vague class $\bar{\mathcal{V}}$ by means of the vague superclass $\mathfrak{V}$ of its instances has basically two advantages: (i) if $\mathfrak{V}$, however it is defined, is certainly contained in an exact superclass whose elements satisfy a suitable property, then certainly the elements of $\mathfrak{V}$ satisfy that property, so we can say that the vague class $\bar{\mathcal{V}}$ satisfies that property as well. (ii) Cournot's principle relates typical and practically certain events. By using superclasses, this relation can be expressed as a relation between exact instances of two vague superclasses rather than between two vague classes, and the former approach is clearer than the latter from a logical point of view.

This method is exemplified below by typical and practically certain events.

3.2 Typical events

Typicality is a vague property of the events that is derived from a probability measure: An event $A \in \mathcal{A}$ is said to be *typical* according a probability measure P (or P-*typical*) if

$$P(A) \approx 1. \tag{17}$$

More generally, we say that $A \in \bar{\mathcal{A}}$ is $\bar{P}$-*typical* if

$$\bar{P}(A) \approx 1. \tag{18}$$

REMARK. We note that one of the corollaries of Theorem 1 is:

$$\bar{\mathcal{T}}(P_1,\delta) = \bar{\mathcal{T}}(P_2,\delta) \Rightarrow P_1 = P_2 \tag{19}$$

for any $\delta \in (\frac{1}{2}, 1)$. This means that the $\bar{P}$-typical events are sufficient to determine P.

Let $\bar{\mathcal{T}}(P)$ denote the vague class of $\bar{P}$-typical events and let $\bar{\mathfrak{T}}(P)$ denote the vague superclass of its instances. The superclass $\bar{\mathfrak{T}}(P)$ can be better defined by introducing the notion of *threshold*: A threshold is a number δ_T that can be conventionally chosen to properly discriminate between $\bar{P}$-typical ($\bar{P}(A) \geq \delta_T$) and non-$\bar{P}$-typical ($\bar{P}(A) < \delta_T$) events. Not any $\delta \approx 1$ can be chosen as a threshold: While δ_T must certainly be close to 1, it cannot be too close, because in this case some events that are certainly typical would be defined as non-typical.

The superclass $\bar{\mathfrak{T}}(P)$ can then be defined as follows:

$$\bar{\mathfrak{T}}(P) := \{\bar{\mathcal{T}}(P,\delta_T) : \delta_T \text{ is a threshold}\}. \tag{20}$$

This vague superclass is certainly contained in the exact superclass $\{\bar{\mathcal{T}}(P,\delta) : \delta \in (\frac{1}{2}, 1)\}$. From this fact and from Theorem 1 one easily deduces the following properties:

PROPOSITION 2.

(a) Every element of $\bar{\mathfrak{T}}(P)$ is a C-class;
(b) $\bar{\mathfrak{T}}(P)$ is contained in an equivalence superclass;
(c) $P_1 = P_2$ if and only if $\bar{\mathfrak{T}}(P_1)$ and $\bar{\mathfrak{T}}(P_2)$ are contained in the same equivalence superclass.

The proof is omitted.

3.3 Experiments

The empirical notions of experiment and trial are well known. Every experiment is associated with an event space. The symbols E, E_1, and E_2 always denote experiments with event space $\mathcal{A}$. The symbol E^n denotes the experiment with event space $\mathcal{A}^n$ whose trials are composed of n trials of E.

We say that two experiments E_1 and E_2 are *empirically equivalent* if they cannot be distinguished by observing the outcomes they produce. In this case we write $E_1 \sim E_2$.

3.4 Practically certain events

As previously stated, we assume that some events of an experiment may possess an empirical objective property that we call *practical certainty*. This property depends only on the structure of the experiment. Let us propose the following operational definition:

DEFINITION (Practical certainty).

(a) Practical certainty is defined operationally as follows: We single out an event and then perform a long sequence of trials; the event is practically certain if and only if its relative frequency in the sequence is very close to 1.
(b) Like any experimental procedure, the above procedure may sometimes produce the wrong result. However, this does not prevent us from considering practical certainty as an objective property of some events.
(c) A natural consequence of the above definition is that if we single out a practically certain event and then perform a single trial of the experiment, the event occurs nearly certainly in the trial.

It is easy to recognize that the class of practically certain events of an experiment satisfies the same properties as the components of a C-class, namely:

(a) The sample space is practically certain;
(b) if A is practically certain and $A \subseteq B$, then B is practically certain;
(c) two disjoint events cannot both be practically certain.

Let us justify for example (c): We single out two disjoint events A and B and then perform a long sequence of trials. If A is practically certain, its relative frequency is close to 1; this implies that the relative frequency of B is not close to 1, and therefore, B is not practically certain. Properties (a) and (b) can be justified even more easily.

However, the reasoning for deducing these properties from the definition of practical certainty is not sufficiently rigorous to be presented as a formal proposition, and therefore, we formulate it as a postulate.

We say that an event $A^{(n)} \in \bar{\mathcal{A}}$ is $\bar{E}$*-practically certain* if it is a practically certain event of the experiment E^n. Let $\bar{\mathcal{C}}(E)$ denote the vague class of the $\bar{E}$-practically certain events and let

$$\bar{\mathfrak{C}}(E) \tag{21}$$

denote the vague superclass of its instances. According to what was said above, we postulate that:

POSTULATE 1.

(a) Every element of $\bar{\mathfrak{C}}(E)$ *is a C-class.*

The above property corresponds to property (a) of Proposition 2 relative to typical events. A version of properties (b) and (c) of that proposition can also be formulated for practically certain events. Since these properties are less self-evident than property (a), we prefer to formulate them as conjectures:

CONJECTURE 1.

(b) $\bar{\mathfrak{C}}(E)$ *is contained in an equivalence superclass;*
(c) $E_1 \sim E_2$ *if and only if* $\bar{\mathfrak{C}}(E_1)$ *and* $\bar{\mathfrak{C}}(E_2)$ *are contained in the same equivalence superclass.*

Property (b) implies that a C-measure can be associated with an experiment. In the next subsection it is shown that if E is governed by a probability measure this property can be deduced rather than conjectured.

Property (c) can be considered in some way as the mathematical definition of the empirical equivalence of two experiments. It is used in the next subsection to prove the non-existence of experimental ambiguity.

It is possible that a better understanding of the structure and interpretation of the C-measure will allow us to reformulate Conjecture 1 as a postulate.

3.5 Cournot's principle revisited

The revised version of Cournot's principle that we propose is the following:

COURNOT'S PRINCIPLE. A probability measure P *governs* an experiment E if for any instance $\bar{\mathcal{C}} \in \bar{\mathfrak{C}}(E)$ there is an instance $\bar{\mathcal{T}} \in \bar{\mathfrak{T}}(P)$ such that

$$\bar{\mathcal{T}} \subseteq \bar{\mathcal{C}}. \tag{22}$$

The following proposition contains the results announced in the Introduction.

PROPOSITION 3.

(a) If P *governs* E*, then* $\bar{\mathfrak{T}}(P)$ *and* $\bar{\mathfrak{C}}(E)$ *are contained in the same equivalence superclass;*
(b) if P_1 *and* P_2 *govern the same experiment* E*, then* $P_1 = P_2$*;*
(c) if P *governs two experiments* E_1 *and* E_2*, then* $E_1 \sim E_2$*.*

PROOF. (a) Let $\bar{\mathcal{C}}_1$ and $\bar{\mathcal{C}}_2$ be two instances of $\bar{\mathfrak{C}}(E)$. According to the hypothesis, there are two instances $\bar{\mathcal{T}}_1, \bar{\mathcal{T}}_2 \in \bar{\mathfrak{T}}(P)$ such that $\bar{\mathcal{T}}_1 \subseteq \bar{\mathcal{C}}_1$ and $\bar{\mathcal{T}}_2 \subseteq \bar{\mathcal{C}}_2$. From Theorem 1 it follows that $\bar{\mathcal{C}}_1 \sim \bar{\mathcal{T}}_1 \sim \bar{\mathcal{T}}_2 \sim \bar{\mathcal{C}}_2$.

(b) Let $\bar{\mathcal{C}} \in \bar{\mathfrak{C}}(E)$. According to the hypothesis, there are $\bar{\mathcal{T}}_1 \in \bar{\mathfrak{T}}(P_1)$ and $\bar{\mathcal{T}}_2 \in \bar{\mathfrak{T}}(P_2)$ such that $\bar{\mathcal{T}}_1 \cup \bar{\mathcal{T}}_2 \subseteq \bar{\mathcal{C}}$. This implies $\bar{\mathcal{T}}_1 \sim \bar{\mathcal{T}}_2$, and therefore, $P_1 = P_2$.

(c) Let $\bar{\mathcal{C}}_1 \in \bar{\mathfrak{C}}(E_1)$ and $\bar{\mathcal{C}}_2 \in \bar{\mathfrak{C}}(E_2)$. According to the hypothesis, there are $\bar{\mathcal{T}}_1, \bar{\mathcal{T}}_2 \in \bar{\mathfrak{T}}(P)$ such that $\bar{\mathcal{T}}_1 \subseteq \bar{\mathcal{C}}_1$ and

$\bar{\mathcal{T}}_2 \subseteq \bar{\mathcal{C}}_2$. This implies $\bar{\mathcal{C}}_1 \sim \bar{\mathcal{C}}_2$, and therefore $\bar{\mathfrak{C}}(E_1)$ and $\bar{\mathfrak{C}}(E_2)$ belong to the same equivalence superclass. From Conjecture 1c it follows that $E_1 \sim E_2$. □

We note that property (a) implies Conjecture 1b, and the C-measure generated by $\bar{\mathfrak{C}}(E)$ is exactly $\bar{P}$. Properties (b) and (c) exclude probabilistic and experimental ambiguity, respectively. Finally, we point out that properties (a) and (b) follow from Postulate 1, while property (c) follows from Conjecture 1c.

4. CONCLUSION

Let us recall the main motivations for the reformulation of Cournot's principle proposed in this paper: (1) In its original formulation, the principle evokes the empirical property of practical certainty without really defining it. In this paper, this property has been explicitly recognized, operationally defined, and some of its properties identified. (2) According to the modern measure-theoretic approach to probability, in the novel formulation, a probability measure and an experiment are recognized as separated entities that are related by Cournot's principle.

The novel formulation makes the logical structure of the principle clearer and emphasizes the possible existence of two ambiguities, namely: (1) that different probability measures govern the same experiment, and (2) that the same probability measure governs different experiments. The first ambiguity is excluded in any case, while the second is excluded provided that a suitable condition characterizing the empirical equivalence of two experiments is assumed.

In this paper, some novel mathematical notions have been introduced, most notably the notion of C-measure. This set function has been considered here as a simple mathematical tool, but it is possible that further research will allow us to interpret it as a non-additive measure of the degree of certainty of the events.

REFERENCES

[1] ALLORI, V. (2020). *Statistical mechanics and scientific explanation: Determinism, indeterminism and laws of nature*. World Scientific.
[2] DÜRR, D. and STRUYVE, W. (2021). Typicality in the foundations of statistical physics and Born's rule. In *Do Wave Functions Jump?* 35–43. Springer.
[3] GOLDSTEIN, S. (2001). Boltzmann's Approach to Statistical Mechanics. In *Chance in Physics: Foundations and Perspectives*, (J. Bricmont, D. Dürr, M. C. Galavotti, G. Ghirardi, F. Petruccione and N. Zanghi, eds.). *Lecture Notes in Physics* **574** Springer-Verlag.
[4] GOLDSTEIN, S. (2012). Typicality and notions of probability in physics. In *Probability in physics* 59–71. Springer.
[5] LAZAROVICI, D. (2023). *Typicality Reasoning in Probability, Physics, and Metaphysics*. Springer Nature.
[6] PAPOULIS, A. (1990). *Probability and statistics*. Prentice-Hall, Inc.
[7] SHAFER, G. and VOVK, V. (2006). The sources of Kolmogorov's Grundbegriffe. *Statistical Science* **21** 70–98.