

Article

Constructive Verification, Empirical Induction, and Falibilist Deduction: A Threefold Contrast

Julio Michael Stern

Department of Applied Mathematics, Institute of Mathematics and Statistics, University of São Paulo.

Rua do Matão 1010, Cidade Universitária. 05508-090, São Paulo, Brazil;

E-Mail: jmstern@hotmail.com; Fax: +55-11-3819-3922

Received: 9 August 2011; in revised form: 21 October 2011 / Accepted: 26 October 2011 /

Published: 31 October 2011

Abstract: This article explores some open questions related to the problem of verification of theories in the context of empirical sciences by contrasting three epistemological frameworks. Each of these epistemological frameworks is based on a corresponding central metaphor, namely: (a) Neo-empiricism and the gambling metaphor; (b) Popperian falsificationism and the scientific tribunal metaphor; (c) Cognitive constructivism and the object as eigen-solution metaphor. Each of one of these epistemological frameworks has also historically co-evolved with a certain statistical theory and method for testing scientific hypotheses, respectively: (a) Decision theoretic Bayesian statistics and Bayes factors; (b) Frequentist statistics and p -values; (c) Constructive Bayesian statistics and e -values. This article examines with special care the Zero Probability Paradox (ZPP), related to the verification of sharp or precise hypotheses. Finally, this article makes some remarks on Lakatos' view of mathematics as a quasi-empirical science.

Keywords: Bayesian statistics; cognitive constructivism; eigen-solutions; Imre Lakatos; neo-empiricism; Popperian falsificationism; sharp hypotheses; zero probability paradox

Gambling problems... seem to embrace the whole of theoretical statistics
Dubins and Savage (1965, Sec.12.8, p.229).

He who wishes to solve the problem of induction must beware of trying to prove too much.
Karl Popper; Replies to my Critics, in Schilpp (1974, Ch.32, p.1110).

With a positive solution to the problem of induction, however thin, methodological theories of demarcation can be turned from arbitrary conventions into rational metaphysics.
Imre Lakatos; A Plea to Popper for a Whiff of Inductivism, in Schilpp (1974, Ch.5, p.258).

(Omega:) I do not like this shift from truth to rationality.
Whose rationality? I sense conventionalist infiltration.
Lakatos (1976, p.104).

1. Introduction

The four quotations that open this article, from [1–3], demarcate the broad area of its scope and interest. They are related to the problem of verification of theories in the context of empirical sciences. Savage quotation states the neo-empiricist credo based on the gambling metaphor, according to which rational beliefs about scientific theories and the associated learning processes follow the logic rules of induction, the very same logic of traditional horse betting. Popper quotation alludes to well known difficulties in applying this concept of induction, specially in the case of sharp or precise hypotheses. Some of these difficulties are related to the ZPP—the Zero Probability Paradox. Nevertheless, Lakatos quotation on induction refers to his desire for some form of *aufhebung* from scientific methodology to rational metaphysics, that is, he seeks a way to positively verify theories in the practice of good science that justifies or explains why it is reasonable to accept these theories as truth bearing statements. This is the core of the classical problem of hypothesis verification in empirical science. Finally, Lakatos quotation on truth and rationality relates to his view of mathematics as a quasi-empirical science.

It is hard to choose a word among verification, confirmation, corroboration, or similar ones, because all of them are already heavily overloaded with very specific meanings. We choose verification for its direct etymological link to the truth bearing status of a statement. We will analyze the verification problem and other related questions from the perspective of Cog-Con—the Cognitive Constructivism epistemological framework. Cog-Con comes equipped with the statistical apparatus of FBST—the Full Bayesian Significance Test. FBST, in turn, defines a statistical support function for sharp hypothesis, namely, the *e*-value—*epistemic value* of a hypothesis given the observed data, or *evidence value* of the data in support of the hypothesis.

The FBST solution to the problem of verification is indeed very thin, in the sense that the proposed epistemic support function, the *e*-value, although based on a Bayesian posterior probability measure, provides only a possibilistic support measure for the hypothesis under scrutiny. We will see how this apparent weakness is in fact the key to overcome the barrage of classical impossibility results alluded by Popper. Nevertheless, the simultaneous Cog-Con characterization of the supported objects (sharp or precise hypotheses) as eigen-solutions implies such a strong and rich set of essential properties, that the

Cog-Con solution becomes also very positive. Moreover, the FBST formal apparatus naturally implies a logic, that is, an abstract belief calculus for composition of support functions and mechanisms for truth propagation in credal networks. In this context, we are able to uphold the cause for the Cog-Con framework as a solution to (a specific form of) the problem of verification that is capable of fulfilling (at least in substantial part) Lakatos' plea.

In the following sections we try to explain and clarify several aspects related to the Cog-Con approach to hypothesis verification, including its central metaphor of objects as token for eigen-solutions. We also contrast this approach with two standard alternatives, neoclassical empiricism and Popperian falsificationism, and their central metaphors, namely, gambling and the scientific tribunal. These two alternative epistemological frameworks are, in turn, associated to two alternative statistical methodologies, namely, frequentist (classical) and Bayesian statistics. In the limited space of this article, we cannot afford to explain any of the aforementioned statistical theories. [4,5] are the required readings and natural companions for this article. For a broader perspective on the FBST, see [6–10]. For the orthodox Bayesian approach, see [1,11–15]. For classical p -values, see [16].

1.1. *Objects as Tokens for Eigen-Solutions*

The Cog-Con framework rests upon Heinz von Foerster's metaphor of *Objects as tokens for an eigen-solution* that, in turn, relies on Humberto Maturana and Francisco Varela's conceptual framework of *Autopoiesis and cognition*. These are the keys to Cog-Con ontology and metaphysics. This section presents these concepts in a nutshell and is based on books [17,18].

The notion of autopoietic system is an abstraction aiming to model the most essential properties of a living organism. Autopoiesis can be understood as an operationally based conceptual framework about the systemic nature of living beings. Autopoietic systems are non-equilibrium (dissipative) dynamical systems exhibiting (meta) stable structures, whose organization remains invariant over (long periods of) time, despite the frequent substitution of their components. Moreover, these components are produced by the same structures they regenerate. The regeneration processes in the autopoietic system production network always requires the acquisition of resources such as new materials, energy and neg-entropy (order), from the system's environment. Efficient acquisition of the needed resources demands selective (inter) actions which, in turn, must be based on suitable inferential processes (predictions). Hence, these inferential processes characterize the agent's domain of interaction as a cognitive domain.

In spite of the fact that autopoiesis was a conceptual framework developed to suit the essential characteristics of organic life, the concept of autopoietic system has been applied in the analysis of many other concrete or abstract autonomous systems such as social systems and corporate organizations. In particular, scientific research systems can be seen in this light, see [19,20].

The circular (cyclic or recursive) characteristic of autopoietic regenerative processes and their eigen- (auto, equilibrium, fixed, homeostatic, invariant, recurrent, recursive) states, both in concrete and abstract autopoietic systems, are investigated in von Foerster in [18,21]. The recursive nature of autopoietic systems produces recurrent states or stable solutions. Under appropriate conditions, such a solution, if presented to the system, will regenerate itself, as a fixed point, an equilibrium or homeostatic state. These are called eigen-values, eigen-vectors, eigen-functions, eigen-behaviors or, in general, eigen-solutions. The concept of eigen-solution is the key to distinguish specific objects in the cognitive

domain of an autopoietic system . Objects are “*tokens for eigen-solutions*”. (A soccer ball is something that interacts with a human in the exact way it is supposed to do for playing soccer.) Eigen-solutions can also be tagged or labeled by words, and these words can be articulated in language. Of course, the articulation rules defined for a given language, its grammar and semantics, only make the language useful if they somehow correspond to the composition rules for the objects the words stand for.

Moreover, von Foerster establishes four essential attributes of eigen-solutions: Eigen-values are ontologically discrete, stable, separable and composable. It is important to realize that, in the sequel, the term “discrete”, used by von Foerster to qualify eigen-solutions in general, should be replaced, depending on the specific context, by terms such as lower-dimensional, precise, sharp, singular *etc.* In several well known examples in exact sciences, these four essential properties lead to the concept of basis, basis of a finite vector space, like in linear algebra, basis of a Hilbert space, like in Fourier or Wavelet analysis, or more abstract basis, like in a matroid structure. Nevertheless, the concept of eigen-solution and its four essential properties is so important in the Cog-Con framework, that it is used as a fundamental metaphor in far more general, and not necessarily formalized, contexts.

For detailed interpretations of von Foerster’s four essential attributes of eigen-solutions, the best references are his original works in [18,21]. For some some examples concerning the applications at hand, see [22–25]. Based on the operational properties of von Foerster’s four essential attributes, these articles develop the thesis that, in the practice of empirical science, important (known) objects are adequately represented as statistical sharp or precise hypothesis. All these articles consider the scientific system as an abstract autopoietic system, as in [19]. Meanwhile, the articles [4,8–10], summarize the formal properties of the FBST, the Full Bayesian Significance Test, and explain the compatibility and mutually supportive relationships between the FBST’s formal properties and the aforementioned abstract attributes of eigen-solutions.

1.2. *The Gambling and the Scientific Tribunal Metaphors*

The gambling metaphor, accompanied by the colorful language of betting odds, is at the core of neoclassical empiricism. In p.152, V.2 of [26], Imre Lakatos states:

“Neoclassical empiricism had a central dogma: the dogma of the identity of (1) probabilities; (2) degree of evidential support (or confirmation); (3) degree of rational belief, and (4) rational betting quotients. This “neoclassical chain of identities” is not implausible. For a true empiricist the only source of rational belief is evidential support: thus he will equate the degree of rationality of a belief with the degree of its evidential support. But rational belief is plausibility measured by rational betting quotients. It was, after all, to determine rational betting quotients that the probability calculus was invented.”

In a game where there is a priori knowledge about the competitors, including perceived differences in strength, skill or other fair or unfair advantages, a score handicap compensation system can be used to equilibrate the winning chances of all the competitors. Gambling, with all its quirks and peculiarities, is the driving metaphor of decision theoretic (orthodox) Bayesian statistics. Several aspects and consequences of using this metaphor are analyzed in [1,11,12,22–25].

A modern tribunal follows the principle of *in dubio pro reo*, giving the defendant the benefit of the doubt, that is, the defendant is considered innocent until proven guilty. The benefit of the doubt is a consequence of the *onus probandi* or burden of proof principle, that states—*semper necessitas probandi incumbit ei qui agit*, that can be translated as—the burden of proof always rests upon the agent laying charges. On one hand, the benefit of the doubt makes it harder to condemn the defendant. On the other hand, the verdict of a judgment can never be “innocent”, but only “guilty” or “not-guilty”.

In the tribunal metaphor, a scientific law is (provisionally) accepted to be truthful, until it is refuted or proved wrong by pertinent evidence. In the court of science, pertinent evidence that can be used to refute a theory has the form of empirical observations that disagree with the consequences or predictions made by the theory on trial. Hence, a fair trial in the scientific tribunal can assure the validity of the deduction process leading to a proof a falsehood, but cannot give any positive certification or assurance concerning a theory’s validity or good quality.

Empirical sciences, specially the so-called exact sciences, like physics, chemistry or engineering, deal with quantitative entities. Moreover, the standard practice of these sciences also requires the truth content of scientific hypotheses to be handled in a quantitative fashion, that is, to undergo a quantitative form of judgment for accuracy and precision. Furthermore, the Cog-Con framework allows us to model some aspects of the development of science in the context of dynamical systems and evolutionary processes, see [27] and Ch.5 of [25]. Nevertheless, measurements of fitness, adaptation and progress in these evolutionary processes also require metrics for rating the objectivity of a concept, the epistemic value of an hypothesis, the statistical significance of a theory, *etc.*

Given the importance the metrics used to evaluate the scientific statements, and the many roles they play in the practice of science, we must choose these metrics with extreme attention, care and caution, designing their structure and regulating their strength and balance. The standard metrics used in empirical science are based on mathematical statistics. Although many alternative belief calculi have been able to successfully occupy local niches or find special applications, modern statistical data analysis finds no rival for its elegance, robustness, flexibility, computational power, and the generality of its scope of application. Nevertheless, there are many long standing issues and unresolved problems related to the use of statistical metrics in the context of hypothesis verification. This is the topic on the next section.

2. ZPP—The Zero Probability Paradox

In order to fully appreciate the Cog-Con+FBST framework, and make further contrasts with other approaches, we use two celebrated examples given in the XIX century by Charles Saunders Peirce. These examples concern the abduction and induction of hypotheses, studying possible procedures to guess, justify and test statements in statistical models. As used in these two examples, Peirce’s idea of “induction” would nowadays be called *parameter estimation*. Meanwhile Peirce’s idea of “abduction” would now-a-days be called *model selection*. In the context of these two examples, induction and abduction would also relate to the contemporary concept of *hypothesis testing*. For the sake of simplicity, instead of using Peirce’s original terminology, we present his examples translating them to contemporary statistical language. We hope to “tradurre senza tradire”, that is, to make this translation without betraying Peirce’s original meaning or losing his intuition. We also review some of the standard modern treatments for these two prototypical examples. As we will see, many aspects of modern treatments, as

well as many of their inherent difficulties, have been foreseen, one way or another, in Peirce's work. The ZPP is at the core some of these alluded difficulties, and will be studied in some detail.

The statistical work of Charles Saunders Peirce has several aspects that deserve further scrutiny. For example, [23,28,29] analyze his pioneering use of randomization in experiment design. Peirce's work on kernel methods and other statistical ideas are almost forgotten. More historical investigations on how, and to what extent his ideas came to influence modern statistical science are long overdue. Peirce's philosophical system and the compatibility or incompatibility of his ideas with any of the epistemological frameworks mentioned in this article is another field that deserves further attention. Section 6 of [23] makes a humble attempt to investigate if the epistemological framework of cognitive constructivism is compatible with, or if it can benefit from, the concepts of semiotics and Peircean philosophy. However, most of these important and interesting topics must wait for future research. In the present article we focus our full attention solely on the two examples of inference described in the next sub-section.

2.1. Two Examples of Inference by Ch.S.Peirce

The first example, published by Peirce in 1868, [30], concerns the induction of letter frequencies and the abduction of cipher codes. The Cipher example, described in contemporary statistical language, is as follows:

- Given the English books B_1, B_2, \dots, B_k , we compile the vectors $\lambda^1, \lambda^2, \dots, \lambda^k$ with the frequencies in which every letter in the alphabet occurs in the text. We realize that they all (approximately) agree with with the mean or average frequencies in vector λ^a .
- Given a new English book, B_{k+1} , we may, by *induction*, state that its letter frequency vector, λ^{k+1} (not yet compiled) will also be (approximately) equal to λ^a .
- Given a coded book C , whose text was encrypted using a simple substitution cipher, we compile its letter frequency vector, λ^c . We realize that there is one and only one permutation vector, π , that can be used to (approximately) match vectors λ^a and λ^c , that is, there is a unique bijection $\pi = [\pi(1), \pi(2), \dots, \pi(m)]$, where m is the number of letters in the alphabet, such that $\lambda^a(j) \approx \lambda^c(\pi(j))$, for $1 \leq j \leq m$. In this case we may, by *abduction*, state the hypothesis that vector π is the correct key for the cipher.

A standard formulation for the induction part of this example includes parameter estimation (posterior distribution, likelihood or, at least, a point estimate and confidence interval) in an n-dimensional Dirichlet-Multinomial model, where m is the number of letters in the alphabet, see [31]. The parameter space of this model is the $(m - 1)$ -simplex, $\Lambda = \{\lambda \in [0, 1]^m \mid \lambda' \mathbf{1} = 1\}$; For a particular case, see the Hardy–Weinberg example in [4]. A possible formulation for the abduction part involves expanding the parameter space of the basic model to $\Theta = \Lambda \times \Pi$, where Π , the discrete space of m -permutations, encodes the key to the cipher.

Peirce's (abductive) hypothesis about the cipher proclaims the "correct" or "true" permutation vector, π^0 . This hypothesis has an interesting peculiarity: The parameter space, $\Theta = \Lambda \times \Pi$, has a continuous sub-space, Λ , and a discrete (actually, finite) sub-space, Π . However, the hypothesis only (directly) involves the finite part. This peculiarity makes this hypothesis very simple, and amenable to the treatment

given by Peirce. However, over-simplification can be a dangerous thing, as shown by an example published by Peirce in 1883 [32], concerning the abduction of a hypothesis with continuous parameters.

“[Kepler] traced out the miscellaneous consequences of the supposition that Mars moved in an ellipse, with the sun at the focus, and showed that both the longitudes and the latitudes resulting from this theory were such as agreed with observation. ...The term Hypothesis [means] a proposition believed in because its consequences agree with experience.”

Instead of formulating Kepler’s hypothesis in a contemporary statistical model, we can make use of an equivalent example already at hand, namely, the Hardy–Weinberg model formulated in [4]. For a sharp hypotheses H stated in a continuous parameter space, Peirce perceives that we cannot speak about the probability of H given the observed data or, more exactly, that $\Pr(H | X) = 0$, that is, that the probability of such a statement is always zero. This is the origin of the ZPP, a symptom that is part of a complex syndrome. As a consequence, Peirce “shifts the problem” (to use a Lakatosian expression) of testing the hypothesis to an assessment of its predictive power or accuracy.

Peirce’s idea for testing the cipher hypothesis is a forerunner of modern decision theoretic Bayesian procedures that compute support values like *posterior probabilities* or *betting odds*. Peirce’s idea for testing Kepler’s or other continuous hypothesis is a forerunner of statistics procedures that compute support values like the classical statistics p -value (alas, there are now Bayesian versions of that too). In the next subsection we try to give an intuitive and non-technical review of these two prototypical solutions, contrasting them with the FBST. For far more detailed analysis, see [25].

2.2. Frequentist and Decision Theoretic Orthodox Bayesian Statistics for Peirce’s Examples

Peirce’s approach to the first example leads, in modern Statistical theory, to a decision theoretic posterior probability for the hypothesis, given the observed data bank. This approach works very well for the Cipher problem. In fact, as the number of observations increase, posterior probabilities will automatically converge, concentrating full support (probability 1) into the true hypothesis. Hence, in this simple problem, we can in fact confuse the problems of induction and abduction. In a context with finite set of alternative hypotheses, one can equivalently speak about the posterior probability of hypothesis H_i , namely $p_i = \Pr(H_i | X)$, or the betting odds for hypothesis H_i , that is, $b_i = p_i / (1 - p_i)$.

The posterior probability solution can be adapted for testing hypotheses on continuous parameters if we only consider partitions of the parameter space into a finite number of non-zero measure sets, corresponding to coarse, un-sharp or inexact hypotheses. However, this approach breaks down as soon as we consider sharp hypotheses. The reason for this collapse is the zero-probability trap: A sharp hypothesis has zero-probability measure and, accordingly, zero prior probability. Moreover, the multiplicative nature of probabilistic scaling will never update a zero probability to a non-zero value, see [4,33]. This is the origin of the ZPP paradox. If we now consider the Cog-Con framework, we can understand the ZPP syndrome in its full extent:

- (1) The “object as eigen-solution” metaphor implies the sharpness of the corresponding hypotheses;
- (2) Hypotheses sharpness implies zero prior probability (in the natural Lebesgue measure);
- (3) Zero prior probability implies perpetual null support. This is one way of understanding the following conclusion stated by Lakatos in p. 154, V.2 of [26]:

“But then degrees of evidential support cannot be the same as degrees of probability [of a theory] in the sense of the probability calculus. All this would be trivial if not for the powerful time-honored dogma of what I called the ‘neoclassical chain’ identifying, among other things, rational betting quotients with degrees of evidential support. This dogma confused generations of mathematicians and of philosophers.”

There are two obvious ways out of this conundrum:

- (A) Fixing the mathematics to avoid the ZPP; or
- (B) Forbidding the use of sharp hypotheses.

(A) Fixing the mathematics in the standard decision theoretic or orthodox Bayesian framework in order to avoid the ZPP is something that is more easily said than done. Modern Bayesian statistics devised several technical maneuvers to circumvent the ZPP. Some of the best known among these techniques are Jeffreys tests, and other handicapped or relative betting odds for scoring competing sharp hypotheses. These techniques provide ad-hoc procedures for practical use, but are plagued by internal inconsistencies, like Lindley’s paradox, or by the need of justifying auxiliary ad-hoc assumptions, like the choice of prior betting odds or the design characteristics of artificial prior densities (an obvious oxymoron), see Section 10.3 of [14]. This precarious state of affairs is fully recognized and admitted in decision theoretic statistical theory. In fact, the orthodox position is that it is not statistical science that is to blame, but rather the paradigm of sharp hypotheses prevalent in exact science. This attitude leads to justifications to the second solution.

(B) Forbidding the use of sharp hypotheses may be very tempting from an orthodox decision theoretic (Bayesian) point of view, however, it is unfeasible in statistical practice: Scientists and other customers of statistical science just insist on using sharp hypotheses, as if they were magnetically attracted to them, and demand appropriate statistical methods. From the Cog-Con perspective, these scientist are, of course, just doing the right thing. As a compromise solution, some influential statistical textbooks offer methods like Jeffreys tests, taking however the care of posting a scary caveat emptor, warning the user that he is entering theoretically unsound territory at his own risk, see for example p. 234 in [34]. Savage, in p. 254, Section 16.3 of [35], realizes that sharp hypotheses, even if important, make little sense in this paradigm, a position that is accepted throughout decision theoretic Bayesian statistics:

“The unacceptability of extreme (sharp) null hypotheses is perfectly well known; it is closely related to the often heard maxim that science disproves, but never proves, hypotheses, The role of extreme (sharp) hypotheses in science and other statistical activities seems to be important but obscure. In particular, though I, like everyone who practice statistics, have often “tested” extreme (sharp) hypotheses, I cannot give a very satisfactory analysis of the process, nor say clearly how it is related to testing as defined in this chapter and other theoretical discussions.”

Peirce’s intuition for testing sharp hypotheses in his second example leads to the p -value of classical statistics. The p -value is defined as the cumulative probability of data banks that are “more extreme” than the one observed, that is, the p -value integrates (adds over) the probabilities of all possible data banks resulting from the experiment that have a smaller probability of outcome than the probability of

the data bank actually obtained. The p -value is a practical solution that works reasonably well for a *singular* or *point* hypothesis, that is, a hypothesis stating that the true value of a model's parameter, π^0 , is a specific value, π^1 . The p -value has some desirable asymptotic properties, for example: The p -value converges to zero if the hypothesis is false, $\pi^0 \neq \pi^1$, and has uniform limit distribution if the hypothesis is true, $\pi^0 = \pi^1$. These properties are very convenient, since they can be used to obtain numerical approximations relatively easy to compute. Nowadays it is hard to appreciate the importance of these properties in a world where digital computers were not (easily) available, and statistical modeling had to be done using tools like slide-rules, numerical tables and graphical charts, see [36].

Lakatos, in pp. 31–32, V.2 of [26], makes very interesting comments concerning the conceptual and historical relation between Popperian falsificationism and Neyman–Person–Wald statistical theory of p -values for hypothesis testing. For example:

“Since the difficulties with induction had been known for a long time, it is remarkable that independently and nearly simultaneously Neyman and Popper found a revolutionary way to finesse the issue by replacing inductive reasoning with a deductive process of hypothesis testing. They then proceeded to develop this shared central idea in different directions, with Popper pursuing it philosophically while Neyman (in his joint work with Pearson) showed how to implement it in scientific practice.”

However useful as a practical technique, even in the case of point hypotheses, the p -value solution can be criticized on some technical issues. For example, it does not conform with the *likelihood principle* of good statistical inference, see [37,38]. The p -value also offers a deceiving answer, because it “translates” a question related to the parameter space into a completely different question stated in the sample space. This leads to several interpretation difficulties, see for example [39]. However, the p -value solution really starts to break down in the case of composite hypotheses, that is, proper sub-manifolds of the parameter space. (For example, the Hardy–Weinberg equilibrium hypotheses constitutes a 1-dimensional sub-manifold in a 2-dimensional parameter space.) The main reason for that break-down is that the aforementioned “definition” of p -value is really not a definition at all. In the case of composite hypotheses, there is no pre-established order in the sample space, hence no natural notion of more extreme. A standard way to fix the p -value definition is to test the auxiliary point hypotheses $\pi^0 = \pi^*$, where π^* is the maximum likelihood (or MAP—maximum a posteriori) estimator under the original hypothesis, given the observed data. But the maximum likelihood auxiliary hypothesis is post-hoc, and therefore cannot adequately represent the original hypothesis. Other alternatives consider a priori reductions or projections of the composite hypothesis into a point hypotheses by “nuisance parameter elimination” procedures. An excellent survey, containing more than 10 different techniques with this purpose, is given in [40]. Nevertheless, these procedures generate case by case solutions, may become technically convoluted, and are not even always available.

Decision theoretic (Bayesian) posterior probabilities and classical statistics p -values, as well as a host of variations of these two paradigms share one thing in common. The maneuvers used to circumvent underlying technical difficulties create case-by-case solutions. Hence, solutions given for different problems cannot be directly compared or readily combined. Therefore, in these paradigms, it is impossible to define general logical rules or abstract belief calculi for the composition and propagation of support functions, as the rules defined for the FBST in [4].

The FBST solution for testing sharp hypotheses can be seen as a “dual” of the p -value, in the sense that the e -value accumulates the probability mass of more extreme points in the parameter space, just as the p -value accumulates the mass of more extreme points in the sample space. Surprisingly, the use of e -value and related ideas has only been proposed very late in the game of statistical science, in [41]. The FBST makes a clear distinction between the hypothesis space and the parameter space, adopting distinct measures at each one of them, namely, the natural posterior probability Bayesian measure at the parameter space, and the e -value possibilistic measure at the hypothesis space. There have been many proposals for using alternative measures, that however did not make the distinction between parameter and hypothesis space so clear, keeping the same measure for both spaces, as it has been usual in statistical theory. Excellent surveys for these theories are given in [42–47].

2.3. Some Conclusions about Constructive Verification

In the last sections we saw that the Cog-Con was able to tame the ZPP. This process involved taking three basic conceptual steps:

- 1- Adopting the standard Bayesian statistical model setup, including the posterior probability measure in the parameter space.
- 2- Making a clear distinction between the parameter and the hypothesis spaces.
- 3- Defining the e -value as a possibilistic measure for the hypothesis space.

The e -value has several remarkable properties, including the following:

- 4- The use of the e -value possibilistic measure in the hypothesis space is fully compatible and coherent with the use of the posterior probability measure in the parameter space, $p_n = p(\theta | X)$. In fact the FBST is built upon the posterior measure, since the e -value is defined as an integral over the measure $p_n(\theta)d\theta$.
- 5- The definition of the e -value (and e -functions) engenders a logic, that is, compositionality rules for computing and propagating e -values, from elementary constituents to complex statements.
- 6- Moreover, the e -value possibilistic logic has classical logic as its limit in the case of Boolean (0 or 1, false or truth, null or full) support values, see [4].

These properties allow the e -value measure to accomplish two wonderful deeds:

- 7- Solve the ZPP for sharp hypothesis , and
- 8- Work like a bridge, harmonizing probability—the underlying logic of statistical inference, the paradigm of belief calculus for empirical science—and classical logic—the prototypical rule of deductive inference.

Step 7 represents an absolution. Sharp hypotheses are freed from the zero-support syndrome, and admitted as full citizens in the hypothesis space. However, Step 7 does not warrant that there will ever be a sharp hypothesis in an empirical science with good support. In fact, considering the original ZPP, finding such an outstanding (sharp) hypothesis should be really surprising, the scientific equivalent of a miracle! What else should we call showing possible what is almost surely (in the probability measure)

infeasible? Nevertheless, we know that miracles do exist. (Non-believers are encouraged to take some good classes in experimental physics, including a fair amount of laboratory work.)

In the Cog-Con framework, the certification of a sharp hypothesis by e -values close to unity is a strong form of verification, akin to empirical confirmation or pragmatic authentication. In contrast, Popperian corroboration is only fail to refute. Nevertheless, the e -value does not provide the inductive engine or truth-pump dreamed by the empiricist school. There is a lot more to the understanding of science as an evolutionary process than the passive waiting for truthful theories to mushroom-up from well harvested data, see [48] and Ch.4 of [25]. (Actually, such an engine could become a real nightmare, draining all soul and conscience from research activity and extinguishing the creative spirit of scientific life.) Hence, we sustain that the Cog-Con framework follows the golden path, finding a well-suited equilibrium between the opposite extremes of excess, aimed by empiricism, and scarcity, offered by falsificationism. In so doing, the FBST e -value provides exactly the right measure needed for hypothesis verification, answering Imre Lakatos' plea for a "whiff of inductivism".

From this perspective, the Cog-Con framework not only redeems sharp or precise hypotheses from statistical damnation, but places them at the center stage of scientific activity. (The star role of any exact science will always be played by eigen-solutions represented by a statement called *somebody's equation*). Therefore, we believe that Step 7 makes clear the way in which the Cog-Con framework provides important insights about the nature of empirical sciences, insights that, in important issues, penetrate deeper than some of the standard alternative epistemological frameworks.

3. Mathematical Ontology

Mathematics is the common language used for the expression and manipulation of symbolic entities associated with the quantities of interest pertinent to the scope of each particular empirical science. Hence, we are particularly interested in the nature of mathematical language. In this section we will argue that, in the Cog-Con framework, mathematics can be regarded as a quasi-empirical science, an idea developed at length by the philosopher Imre Lakatos. The key of our arguments is Step 8 in the last Section. Step 8 constitutes a bridge from physics to mathematics, from empirical to quasi-empirical science. From this perspective, mathematics is seen as an idealized world of absolutely verified theories populated by hypotheses with full (or null) support.

Ontologies are controlled languages used in the practice of science. They are developed as tools for scientific communication. This communication has typical external and internal aspects: we need language in order to communicate with others and with ourselves. We use language as a tool for effective coordination of action and as a tool for efficient structuring of understanding. Equipped with the appropriate ontologies, scientists are supposed to build models capable of providing reliable predictions and insightful explanations. Moreover, at least in the domain of exact sciences, these models are required to have a formal and quantitative nature. Hence, the approach we follow naturally highlights the special role played by formal or mathematical languages, our main interest in this section.

Formal or "abstract" mathematics, including several of its less formal or popular dialects, is the common language used for the expression and manipulation of symbolic entities associated with the quantities of interest pertinent to the scope of each particular empirical science. Hence, we will be particularly interested in the nature of mathematical language. In fact, we will argue that mathematics

should be regarded as an ontology for a class of concepts relevant to all exact sciences, namely, those related to the intuitive ideas of counting, symmetry, number, infinity, measure, dimension, and the continuum. From this point of view, mathematics can be regarded as a quasi-empirical, as opposed to an Euclidean science, according to the classical distinction defined by Lakatos in p.40, V.2 of [26].

“Whether a deductive system is Euclidean or quasi-empirical is decided by the pattern of truth value flow in the system. The system is Euclidean if the characteristic flow is the transmission of truth from the set of axioms “downwards” to the rest of the system—logic here is the organon of proof; it is quasi-empirical if the characteristic flow is “upwards” towards the “hypothesis”—logic here is an organon of criticism. We may speak (even more generally) of Euclidean vs. quasi-empirical theories independently of what flows in the logical channels: certain or fallible truth or falsehoods, probability or improbability, moral desirability or undesirability, etc. It is the how of the flow that is decisive.”

Of course, in the Cog-Con framework it is verification (or not), measured by e -values, that flows upwards towards the hypotheses, as discussed in the previous sections. At this point, we can see how the Cog-Con framework may lead to a renewed appreciation and a fresh understanding of two celebrated statements by Albert Einstein, p.28 of [49], and Imre Lakatos, p.102 of [3]:

“As far as the statements of mathematics refer to actual truth, they are not certain; and as far as they are certain, they do not refer to actual truth”.

(Kappa:) “If you want mathematics to be meaningful, you must resign of certainty. If you want certainty, get rid of meaning. You cannot have both.”

3.1. Why did Mathematics Become a Deductive Science?

Lakatos' view of mathematics as a quasi-empirical science is not the most widespread current opinion, even though a few authors explore similar ideas, see for example [50] for a perspective of mathematics as an experimental science. On the contrary, since the times of the compilation of the Elements of Geometry, Mathematics is usually seen as an Euclidean science. According to Árpád Szabó, Imre Lakatos teacher in Hungary, this change of perception of the nature of mathematics corresponds to long historical processes that can be traced, among other things, by the transformation of technical words and specialized terms used in mathematical texts. His masterpiece, [51], Szabó studies in great depth the history of this transformation in the beginning of Greek mathematics, see also [52].

For example, in the beginning of Greek mathematics, the use of the word $\delta\epsilon\iota\chi\alpha\iota$ —to demonstrate, corresponds to its colloquial meaning—to show or to display. At early stages, mathematical demonstrations seek the goal of explanation, and are developed as visualization techniques or intuitive gedanken-experiments. Later on, the same word became a technical term in Greek mathematics, as in the expression— $\sigma\pi\epsilon\rho\ \epsilon\delta\epsilon\iota\ \delta\epsilon\iota\chi\alpha\iota$, our familiar *quod erat demonstrandum*. Seeking greater generality, arguments are presented in increasingly abstract or logical form. This later use corresponds to the ideal of arithmetization of mathematical deductions. In this regard, Socrates, in Plato's Republic, 525–526, as quoted in p.194, p.197 of [51], states: *“the subject of arithmetic lay within the domain of pure thought.”* Also the word axiom, $\alpha\chi\iota\omicron\upsilon\upsilon$ —to be worthy, originally meant a proposition put forward

for critical discussion or dialectical debate. Ironically, later on, the same word came to indicate the obvious or self-evident.

Following Szabó arguments, it is possible answer his famous interrogation: How did mathematics become a deductive science? Nevertheless, in this paper, we are more interested in a closely related question, not how, but why: Why did mathematics become a deductive science? Why was that transformation of the mathematical ideal, from down to earth science to aethereal philosophy, even possible?

Once again, the miraculous or wonderful nature of a well supported eigen-solutions, discussed in the previous sections, can offer a positive answer to the last questions. After all, it is only natural to believe that miraculous theorems are born in heaven. I will not venture into the discussion of whether or not good mathematics comes from heaven or “straight from The Book”, as Pál Erdős used to say. I will only celebrate the revelation of this mystery. It represents the ultimate transmutation of the ZPP, from bad omen of confusion, to good augury of universal knowledge.

4. Further Research and Final Remarks

The history of mathematics provides many interesting themes of study that we hope to explore in forthcoming papers. For example, modern logic and set-theory seem to have followed some trends that converge to the Cog-Con perspective. As an illustration, in p. 27, V.2 of [26], Gödel states:

“the role of the alleged ‘foundation’ is rather comparable to the function discharged, in physical theory, by explanatory hypotheses... the actual function of axioms is to explain the phenomena described by the theorems of this system rather than to provide a genuine ‘foundation’ for such theorems.”

Eugen Wigner, [53], and Richard Hamming [54], are astonished by “*the unreasonable effectiveness of mathematics in the natural sciences.*” Examining this mystery from the Cog-Con perspective, we understand that nothing is more natural that the effectiveness of mathematics in the natural sciences, for mathematics is nothing but the order of the natural world (including ourselves) expressed in language (as well as we currently can). Of course, a deeper mystery remains untouched, namely, the existence of an orderly cosmos and not only chaos. Actually, not only the existence of any cosmos, but the existence of a “good” one, in which we can find the sharply defined, stable, separable and composable eigen-solutions we need to use as build blocks in the construction of knowledge. Nevertheless, I believe that even this small change in perspective is, in itself, a nice accomplishment of the Cog-Con framework.

Acknowledgements

The author is grateful for the support of the Department of Applied Mathematics of the Institute of Mathematics and Statistics of the University of São Paulo, FAPESP—Fundação de Amparo à Pesquisa do Estado de São Paulo, and CNPq—Conselho Nacional de Desenvolvimento Científico e Tecnológico (grant PQ-306318-2008-3). This paper was first presented at EBL-2011, the XVI Brazilian Logic Conference, held on May 9–13 at LNCC—Laboratório Nacional de Computação Científica, Petrópolis, Brazil, and also presented at COBAL-2011, the III Latin American Meeting on Bayesian Statistics, held on October 23–27 at UFRO—Universidad de La Frontera, Pucón, Araucaía, Chile. The author is grateful

for the advice received from anonymous referees, and also for helpful discussions with several of his colleagues at the Bayesian research group at University of São Paulo, especially its head, Carlos Alberto de Bragança Pereira, who is always poking and probing our ideas on the foundations of probability and statistics. Finally, the author is grateful for some comments concerning Lakatos' late works made by Gábor Kutrovátz of Loránd Eötvös Budapest University.

References

1. Dubins, L.E.; Savage, L.J. *Inequalities for Stochastic Processes: How to Gamble If You Must*; McGraw-Hill: New York, NY, USA, 1965.
2. Schilpp, P.A. *The Philosophy of Karl Popper*; Open Court: La Salle, Canada, 1974.
3. *Proofs and Refutations: The Logic of Mathematical Discovery*; Lakatos, I., Worall, J., Zahar, E., Eds.; Cambridge University Press: Cambridge, UK, 1976.
4. Borges, W.; Stern, J.M. The rules of logic composition for the bayesian epistemic e -values. *Log. J. IGPL* **2007**, *15*, 401–420.
5. Stern, J.M. Symmetry, invariance and ontology in physics and statistics. *Symmetry* **2011**, *3*, 611–635.
6. Lauretto, M.; Pereira, C.A.B.; Stern, J.M.; Zacks, S. Full bayesian significance test applied to multivariate normal structure models. *Braz. J. Probab. Stat.* **2003**, *17*, 147–168.
7. Madruga, M.R.; Esteves, L.G.; Wechsler, S. On the bayesianity of pereira-stern tests. *Test* **2001**, *10*, 291–299.
8. Pereira, C.A.B.; Wechsler, S.; Stern, J.M. Can a significance test be genuinely Bayesian? *Bayesian Anal.* **2008**, *3*, 79–100.
9. Stern, J.M. Significance tests, belief calculi, and burden of proof in legal and scientific discourse. Laptec-2003. *Front. Artif. Intell. Appl.* **2003**, *101*, 139–147.
10. Stern, J.M. Paraconsistent sensitivity analysis for bayesian significance tests. *SBIA'04* **2004**, *3171*, 134–143.
11. de Finetti, B. *Probability, Induction and Statistics*; Wiley: New York, NY, USA, 1972.
12. de Finetti, B. *Theory of Probability, VI and V2*; Wiley: London, UK, 1974.
13. DeGroot, M.H. *Optimal Statistical Decisions*; McGraw-Hill: New York, NY, USA, 1970.
14. Zellner, A. *Introduction to Bayesian Inference in Econometrics*; Wiley: New York, NY, USA, 1971.
15. Gelman, A.; Carlin, J.B.; Stern, H.S.; Rubin, D.B. *Bayesian Data Analysis*, 2nd ed.; Chapman and Hall—CRC: New York, NY, USA, 2003.
16. Pereira, C.A.B.; Wechsler, S. On the Concept of p -value. *Braz. J. Probab. Stat.* **1993**, *7*, 159–177.
17. Maturana, H.R.; Varela, F.J. *Autopoiesis and Cognition. The Realization of the Living*; Reidel: Dordrecht, The Netherlands, 1980.
18. von Foerster, H. *Understanding Understanding: Essays on Cybernetics and Cognition*; Springer Verlag: New York, NY, USA, 2003.
19. Krohn, W.; Küppers, G. The Selforganization of Science—Outline of a Theoretical Model. In *Selforganization: Portrait of a Scientific Revolution*, Krohn, W., Küppers, G., Nowotny, H., Eds.; Kluwer: Dordrecht, The Netherlands, 1990; pp. 208–222.
20. Luhmann, N. *Ecological Communication*; Chicago University Press: Chicago, IL, USA, 1989.

21. Segal, L. *The Dream of Reality. Heintz von Foerster's Constructivism*; Springer: New York, NY, USA, 2001.
22. Stern, J.M. Cognitive constructivism, eigen-solutions, and sharp statistical hypotheses. *Cybern. Hum. Knowing* **2007**, *14*, 9–36.
23. Stern, J.M. Language and the self-reference paradox. *Cybern. Hum. Knowing* **2007**, *14*, 71–92.
24. Stern, J.M. Decoupling, sparsity, randomization, and objective bayesian inference. *Cybern. Hum. Knowing* **2008**, *15*, 49–68.
25. Stern, J.M. *Cognitive Constructivism and the Epistemic Significance of Sharp Statistical Hypotheses*. Tutorial book for MaxEnt 2008, The 28th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, Boracéia, São Paulo, Brazil, 8-11 July 2008.
26. Lakatos, I. *Philosophical Papers. V.1—The Methodology of Scientific Research Programmes. V.2.—Mathematics, Science and Epistemology*; Cambridge University Press: Cambridge, UK, 1978.
27. Inhasz, R.; Stern, J.M. Emergent Semiotics in Genetic Programming and the Self-Adaptive Semantic Crossover. In *Model-Based Reasoning in Science & Technology*; Magnani, L., Carnielli, W., Eds.; Springer: Berlin, Heidelberg, Germany, 2010; pp. 381–392.
28. Hacking, I. Telepathy: Origins of randomization in experimental design. *Isis* **1988**, *79*, 427–451.
29. Stigler, S.M. *The History of Statistics: The Measurement of Uncertainty before 1900*; Harvard University Press: Cambridge, MA, USA, 1986.
30. Peirce, Ch.S. Questions concerning certain faculties claimed for man. *J. Specul. Philos.* **1868**, *2*, 103–114.
31. Pereira, C.A.B.; Stern, J.M. Special characterizations of standard discrete models. *REVSTAT Stat. J.* **2008b**, *6*, 199–230.
32. Peirce, Ch.S. A Theory of Probable Inference. *Studies in Logic* **1883**, 126–181.
33. Darwiche, A.Y. A Symbolic Generalization of Probability Theory. Ph.D. Thesis, Stanford University, Stanford, CA, USA, 1993.
34. Williams, D. *Weighing the Odds*; Cambridge University Press: Cambridge, UK, 2001.
35. Savage, L.J. *The Foundations of Statistics*; Dover: New York, NY, USA, 1972.
36. Pickett Inc. N525 Stat-Rule, A Multi-Purpose Sliderule for General and Statistical Use (Instruction manual). Santa Barbara, 1965.
37. Pawitan, Y. *In All Likelihood: Statistical Modelling and Inference Using Likelihood*; Oxford University Press: Oxford, UK, 2001.
38. Wechsler, S.; Pereira, C.A.B.; Marques, P.C. Birnbaum's Theorem Redux. In *Proceedings of the 28th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, Edmonton, Canada, 23–24 July 2008;
39. Rouanet, H.; Bernard, J.M.; Bert, M.C.; Lecoutre, B.; Lecoutre, M.P.; Le Roux, B. *New Ways in Statistical Methodology. From Significance Tests to Bayesian Inference*; Peter Lang: Berne, Switzerland, 1998.
40. Basu, D. *Statistical Information and Likelihood*; Ghosh, J.K., Ed.; Springer: Berlin, Heidelberg, Germany, 1988; Volume 45.

41. Pereira, C.A.B.; Stern, J.M. Evidence and credibility: Full bayesian significance test for precise hypotheses. *Entropy* **1999**, *1*, 69–80.
42. Eells, E.; Fitelson, B. Measuring confirmation and evidence. *J. Philos.* **2000**, *97*, 663–672.
43. Hawthorne, J. Confirmation Theory. *Handbook of the Philosophy of Science, Volume 7: Philosophy of Statistics*, Bandyopadhyay, P.S., Forster, M.R., Gabbay, D.M., Thagard, P., Woods, J., Eds.; Elsevier BV.: Amsterdam, The Netherlands, unpublished work, 2010.
44. Huber, F. Confirmation and Induction The Internet Encyclopedia of Philosophy. 2010. Available online: www.iep.utm.edu/conf-ind/ (accessed on 11 November 2010).
45. Kuipers, T.A.F. Inductive probability and the paradox of ideal confirmation. *Philosophica* **1971**, *17*, 197–205.
46. Maher, P. Confirmation Theory. In *The Encyclopedia of Philosophy*, 2nd ed.; Borchert, D.M., Ed.; Macmillan: London, UK, 2005.
47. Strevens, M. Notes on Bayesian Confirmation Theory. 2006. Available online: www.nyu.edu/gsas/dept/philo/user/strevens/Classes/Conf06/BCT.pdf (accessed on 11 November 2010).
48. Hilts, V. Aliis extendum, or the origins of the statistical society of London. *Isis* **1978**, *69*, 21–43.
49. Einstein, A. *Geometrie und Erfahrung*; Springer: Berlin, Heidelberg, Germany, 1921. Available online: http://www.alberteinstein.info/PDFs/CP7Doc52_pp382-388_403.pdf (accessed on 11 November 2010).
50. Burgin, M.; *On the Nature and Essence of Mathematics*; Ukrainian Academy of Information Sciences: Kiev, Ukrain, 1998; in Russian.
51. Szabó, A. *The Beginnings of Greek Mathematics*; Akadémiai Kiadó: Budapest, Hungary, 1978.
52. Kutrovátz, G. Philosophical Origins in Mathematics? Árpád Szabó Revisited. In *13th Novembertagung on the History of Mathematics*, Frankfurt, Germany, 2002.
53. Wigner, E. The unreasonable effectiveness of mathematics in the natural sciences. *Commun. Pure Appl. Math.* **1960**, *13*, 1–14.
54. Hamming, R.W. The unreasonable effectiveness of mathematics. *Am. Math. Mon.* **1980**, *87*, 2.