

# Picking Bayesian or Classical Statistics by Means of Mathematical Criteria

Paolo Rocchi<sup>1,2</sup>

<sup>1</sup>IBM, via Luigi Stipa 150, Roma, Italy

<sup>2</sup>LUISS University, via Romania 32, Roma, Italy  
procchi@luiss.it

**Abstract** – Classical and Bayesian statistics have own proper criteria and techniques. The scientific community debates the pros and cons of the two statistical schools since long, but this broad research driven by theoretical and practical criteria, has not provided definitive answers since a fundamental disagreement lies at the very heart of the problems. Conflicting probability theories underpin the two statistics and the numbers convey irreconcilable meanings in the two sectors.

Eminent authors have developed philosophical arguments to support their probability interpretations, by contrast we have conducted an inquiry which follows a purely mathematical approach. Two theorems prove that the frequentist and subjective models are not incompatible and lead to a precise rule for working statisticians who are called to select the most suitable methods in a project.

**Keywords:** Classical statistics; Bayesian statistics; frequency probability; subjective probability; philosophy and mathematics.

## 1. Introduction

Present day experts have the classical and Bayesian statistics at disposal. Both of the methods share the idea that the more information one gets, the more accurate are the calculated forecasts. Sometimes the two schools adopt symmetrical techniques but besides occasional convergence the two statistics show evident disparities. Usually professionals apply the first or the second methodology in a project on the basis of pragmatic criteria, and an investor who pays for a statistical study and requests for the best outcome, does not obtain univocal answers.

Classical statistics uses sampled data and thus data observed under similar circumstances may be different. The parameters are not random variables while randomness is associated with the variations in replicated observations. The input data are the premises, and the results of the statistical procedures yield the conclusion so classical statistics complies with the deductive or inferential logic.

For Bayesians a statistical parameter is a random variable and subject to probabilistic description. Analysis begins with the quantification of subjective priors deriving from the investigator's existing state of knowledge, beliefs, and assumptions which are taken just as they are. The analysis proceeds based on the Bayes' formula which computes the posterior using the prior and likelihood that are known for all hypotheses. The logical course starting with a prior distribution, getting data, and moving to the posterior distribution is associated with learning about the general from the particulars, hence scholars usually classify the Bayesian inference as inductive.

Using a rather paradoxical language one might conclude that the two statistics adopt opposed criteria and methods of work. This apparent disparity should entail a rigorous criterion to select the best statistics to employ in a project, but this criterion is still missing. Researchers have analyzed the strong and weak points of each statistics from various perspectives [1]-[4], but they have not got a shared criterion. The reason is apparent. Even when the classical and Bayesian procedures provide identical numerical results, those procedures assign irreconcilable meanings to the numbers.

The essential difference between the two schools is in how probability ( $P$ ) is conceived and used [5]. Frequentists see  $P$  as the long-run expected frequency of occurrence. The Bayesian relate probability to a degree of personal belief. In these terms the two views prove to be absolutely irreconcilable.

I have conducted an inquiry in order to look into this long-standing controversy. I have made the textual analysis of the most significant works written by eight masters of probability theory [6]. The textual analysis has made evident how frequentists and subjectivists have written verbose and redundant works. They share the style typical of philosophers and

humanists. They argue over a variety of non-mathematical topics, they discuss about linguistics, sociology, the scientific method etc. They devote quite a number of pages to demolish the rival proposals. In summary, the textual analysis brings evidence of how each author is aware of the multiple nature of probability but he is convinced that only one model of  $P$  is true. He creates a theory in consequence of his personal ideas and philosophical position and holds with all his strength that solely his construction establishes the authentic essence of  $P$ .

Facing this situation, I have developed a purely mathematical approach to the probability interpretation issue which is based on theorems and not on personal opinions.

## 2. The Mathematical Way

Let us recall a few basic assumptions. Probability is the measure of how likely the *random result*  $e$  will occur. Given a sample space  $\Omega$  and an associated sigma algebra  $\Sigma$ , the probability  $P$  is a function with domain  $\Sigma$  that satisfies:

1.  $P(e) \geq 0$  for all  $e \in \Sigma$ .
2.  $P(\Omega) = 1$ .
3. If  $e_{(1)}, e_{(2)}, e_{(3)}, \dots \in \Sigma$  are pairwise disjoint, then 
$$\left( \bigcup_{i=1}^{\infty} e_{(i)} \right) = \sum_{i=1}^{\infty} P(e_{(i)}) . \quad (1)$$

The probability established in abstract and used in applications, needs to be checked. It is necessary to discuss how and when the quantity  $P(e)$  can be tested. Falsifiability, according to the philosopher Karl Popper, endorses the inherent testability of any scientific hypothesis and in turn the genuine science. A scientific area may be checked in a more or less easy manner and the ensuing theorems – proved and discussed in [7] – determine the relation between the relative frequency and probability which emerges in the collective ( $e_{\infty}$ ) and the single trial ( $e_1$ ).

### 2.1 Theorems #1 and #2

Suppose  $e$  is an i.i.d variable, the *theorem of large numbers* (TLN) states that as the number of trials grows, the relative frequency almost surely gets closer to the probability  $P(e_{\infty})$

$$F(e_n) \xrightarrow{a.s.} P(e_{\infty}), \quad \text{as } n \rightarrow \infty. \quad (2)$$

The *theorem of a single number or lower bound theorem* (TSN) demonstrates that the relative frequency unfits with probability in a single trial

$$F(e_1) \neq P(e_1), \quad n = 1. \quad (3)$$

### 2.2 Remarks about TLN

The theorem of large numbers has a conceptual link with the law of large numbers (LLN), but a neat difference divides them.

LLN expresses the convergence of empirical random data toward an expected value. Émile Borel developed a special strong version of LLN where  $q$  is the number of empirical occurrences and  $P$  is a purely numerical value

$$\frac{q}{n} \xrightarrow{a.s.} P, \quad \text{as } n \rightarrow \infty. \quad (4)$$

It is necessary to put beforehand that an experiment dealing with the parameter  $x$  cannot verify  $x$  in all its extension. For example, one checks  $x$  in a finite range and cannot control  $x$  from  $-\infty$  to  $+\infty$ . Therefore, the hypothesis  $n \rightarrow \infty$  implies that TLN deals with a peculiar testable event and a particular kind of probability whereas LLN deals with probability in general. The statement (2) describes the property of  $P(e_{\infty})$  and not of  $P$  because TLN is an applied theorem and not a general law. Therefore, TLN proves that – at least in principle – the probability of repeated events (called *collective* by von

Mises and *series* by Venn) can be tested,  $P(e_\infty)$  is a parameter that exists in the real world, the *frequency probability*  $P(e_\infty)$  is an *authentic physical quantity*.

### 2.3 Remarks about TSN

TSN shows how one cannot test the probability of a single random event. It is not a question of instruments or experimental setting; the theorem proves that never ever one can corroborate  $P(e_1)$  and therefore it *is not a physical quantity*. This conclusion perfectly matches with the famous aphorism by de Finetti: “Probability does not exist”.

TSN implies that scientists should reject  $P(e_1)$  because it is out of control; instead statisticians and even common people are concerned with the probability of a single result. Subjectivists and Bayesians exploit the *semantic value* of  $P(e_1)$  and circumvent this obstacle caused by (3) [7]. What does ‘semantic value’ exactly mean?

*Semiotics* is the science of signs and teaches us that numbers, words, signs etc. are items of information [8-9] and as such they can convey meanings. The number  $P(e_1)$  does not have any physical significance nonetheless it does not lose the capability of conveying significance. Subjectivists and Bayesians use it to express a *personal credence about the occurrence of*  $(e_1)$ , namely  $P(e_1)$  is a *subjective probability*.

Some writers have raised criticism against this model [10] which I recall in concise terms:

1. Subjective probability means to quantify the personal belief which in principle can be affected by the variety of personal convictions held by an individual. The suspicions of arbitrariness have been raised since the early beginnings.
2. The betting scheme appears strange in scientific and engineering sectors.
3. Testing is a key criterion for the scientific method but in principle subjective probability is alien to experimental validation. This turns out to be repellent to the science which investigates objective situations and strives for results independent of the observer.

TLN and TSN negate the existence of a sole probability model and the critiques against the subjective probability lose weight (symmetrically also the negative appraisals against the frequentism become paler). In the light of results (2) and (3), the disapprovals have completely different import and tone. In fact, TSN proves that  $P(e_1)$  has no physical meaning and remark **3** can but recognize this essential property. TSN leads us to grasp how the subjectivists and Bayesians masters conducted an intelligent plan of action to recycle  $P[E_1^{(6)}]$  as epistemic probability. Keynes teaches us how epistemic probabilities are not so easy to assess and to check [11], hence this manoeuvre presents limitations **1** and **2** of necessity. The landscape emerging beyond TSN shows how the most critical annotations against the subjective model decrease in importance. It may be said that the remarks listed above and even other negative commentaries give details about the price paid to reuse  $P(e_1)$  that otherwise should be rejected from the scientific domain.

### 2.4 Theorem #3

TSN establishes the lower constraints of the impracticable tests, it is necessary to complete this analysis and specify the upper limit for the impossibility of testing  $P$ .

*The upper-bound theorem* (UBT) holds that if  $n$  and  $z$  positive integers ( $n < z$ ) and probability verifies

$$P(e_n) = 1/z. \tag{5}$$

Then the relative frequency does not fit with the probability in  $n$  trials

$$F(e_n) \neq P(e_n). \tag{6}$$

In substance the number of trials is to be greater than  $z$  in order to employ the relative frequency.

### 3. What is the Best?

The results achieved by means of the previous theorems can be summarized in the following way

$$\text{Frequentist Probability: } P(e_n) \quad n \rightarrow \infty \tag{7a}$$

$$\text{Subjective Probability: } P(e_n) \quad 1 \leq n < z \quad (7b)$$

### 3.1 Theorem #4

The *theorem of compatibility* (TC) holds that the frequentist and subjective models do not contradict

$$P(e) = [P(e_n) \text{ with } n \rightarrow \infty] \text{ OR } [P(e_n) \text{ with } 1 \leq n < z]. \quad (8)$$

*Proof:* The frequentist and subjective probabilities derive from the assumptions of TLN, TSN and UBT summarized in (7a) and (7b). The hypotheses deal with two disjointed intervals for  $n$ , in consequence of this separation (8) is true.

### 3.2 Remarks about TC

The founders of the frequentist and subjective schools sustain diverging models and share the same philosophy; they assume that probability has only one meaning, they believe there is a sole ‘authentic’ probability model. This philosophical opinion is unproven and constitutes a sort of hidden axiom raising endless debates and preventing researchers from attacking the core of the interpretation issue. TC disproves the dogmatic assumptions of von Mises, Savage, De Finetti and others; TC demonstrates that probability is not a monolith but has a multifold nature.

TC is consistent with the mind of dualist writers such as Ramsey, Popper, Carnap, Lewis, Constantini and others who accept two model of  $P$ .

### 3.3 How to Select the Method

The four theorems regulate the exercise of probability and leads to the criterion for the management of statistical methods. Specifically, from (8) one can reasonably infer the following rules:

1. *If one means to investigate the long-term event (7a), the one must resort to use the classical statistics.*
2. *If one means to investigate a single event (7b), he must adopt the Bayesian statistics.*

There is no middle way and rules **1** and **2** do not allow exceptions.

Assumption (7a) is consistent with the classical statistical inference that makes propositions about a population, using data drawn from that population with some form of sampling. Bayesian statistics teaches us to update our beliefs in the evidence of new data when we have to forecast a future single event. Note how the Bayesian procedures are not confined to a lone observation. When a Bayesian applies to a sequel of repeated events, he addresses situations one by one, his conclusions regard each individual case.

## 4. Comments and Conclusion

1 - Gauss, Hilbert, Einstein and other eminent mathematicians consider simplicity as the virtue of theoretical works, a sign of elegance and not a defect. Experts often perceive simplicity to be a characteristic of ‘beautiful’ mathematics because it is easier to gain enlightenment from a simple proof compared to a complex proof. The theorems presented here sound very simple from the mathematical viewpoint, the illustration very concise and all this should not be deemed as a fault.

2 - In terms of methodology, the present research suggests innovative pathways since it examines the nature of probability by means of the mathematical method and rejects the usual philosophical approach which the book [7] shows to be dogmatic and arbitrary.

3 - The subject contents of the theorems provide new answers to knotty controversies that are open since decades. Specifically:

- a. TLN and TSN tackle the problem of testability that is a typical of science and go beyond the qualitative discussion of the probability ‘meanings’.
- b. TLN and TSN demonstrate that there are two different models of probability:  $P(e_\infty)$  and  $P(e_1)$ ; and not only one.
- c. TSN enables us to grasp how  $P(e_1)$  is reused as subjective probability instead of being scrapped.
- d. TC proves that  $P(e_\infty)$  and  $P(e_1)$  are not irreconcilable.

- e. The impossibility of testing  $P$  involves a small set of events whose boundaries are fixed by TSN and UBT.
- f. TC yields a precise guideline to statisticians who are called for selecting the most appropriate methods in a project.

Experts are examining the pros and cons of classical and Bayesian methods using empirical and sometimes personal criteria, instead rules **1** and **2** indicate a way that descends from theorems and not from individuals' ideas.

## References

- [1] J. Huber, K.E. Train - On the similarity of classical and Bayesian estimates of individual mean partworths – *Marketing Letters*, 2011, vol. 12, no. 3, pp. 259-269.
- [2] B. Efron - Bayesians, frequentists, and scientists - *J. of the American Statistical Association*, 2005, vol. 10, pp. 1-5.
- [3] R. Kass - Statistical inference: The big picture - *Statistical Science*, 2011, vol. 26, no. 1, pp. 1-9.
- [4] J. Sprenger (2016) - Bayesianism vs. Frequentism in Statistical Inference - in A. Hajek and Hitchcock C. (eds.), *Oxford Handbook of Probability and Philosophy*, Oxford University Press, pp. 382-405.
- [5] J. Vallverdú - *Bayesians Versus Frequentists: A Philosophical Debate on Statistical Reasoning* - Springer, 2016.
- [6] P. Rocchi - Four foundational questions in probability theory and statistics - *Physics Essays*, 2017, vol. 30, no. 3, pp.314-321 <https://arxiv.org/abs/1901.03876>
- [7] P. Rocchi - *Janus-faced Probability* - Springer, 2014.
- [8] D. Chandler - *Semiotics: The Basics* - Routledge, 2017.
- [9] P. Rocchi - *Logic of Analog and Digital Machines* - 2nd revised ed., Nova Science Publishers 2010.
- [10] H. Kyburg - Subjective probability: Criticisms, reflections, and problems - *Journal of Philosophical Logic*, 7(1), pp. 157-180, 1978.
- [11] B.W. Bateman - Keynes's changing conception of probability - *Economics and Philosophy* – 1987, 3(1), pp. 97-119