QUADERNI

Università degli Studi di Siena
DIPARTIMENTO DI ECONOMIA POLITICA

ALBERTO BACCINI

F.Y. Edgeworth's *Treatise on Probabilities*

n. 494 - Gennaio 2007

Abstract: Probability theory has a central role in Edgeworth's thought; this paper examines the philosophical foundation of the theory. Starting from a frequentist position, Edgeworth introduced some innovations on the definition of primitive probabilities. He distinguished between primitive probabilities based on experience of statistical evidence, and primitive a priori probabilities based on a more general and less precise kind of experience, inherited by the human race through evolution. Given primitive probabilities, no other devices than the rules of calculus are necessary to infer complex probabilities, as the ones defined by Bayes's theorem –an enlargement of the frequentist tradition as defined by Venn.

The notion of probability is objective; the passage from this objective sphere to the epistemic one requires rules external to the theory of probability. Edgeworth distinguishes between two notions: credibility which is the direct translation of probability into the epistemic sphere, and obeys the same rules of the latter; and belief having a weak relation with probability, based as it it not only on experiential knowledge, but also on "instinct and sentiment". According to a Nineteenth century tradition, belief is the base of human action; Edgeworth concludes therefore that probability is not useful for the theory of decision.

We propose to classify Edgeworth's theory of probability as precursor of modern eclectic or pluralistic tradition on probability, and according to which probability has an irreducible dualistic nature.

Keywords: F.Y. Edgeworth, Philosophy of probability, Frequentist probability, Bayes's theorem
J.E.L. code: B13, B31, B40

16 maggio 2005 r.1.0
Revised version 22/01/2007 r.2.1

Alberto Baccini, Dipartimento di Economia Politica, Università di Siena

F.Y.Edgeworth's *Treatise on Probabilities*

John Maynard Keynes (1926: 261-262) wrote that "if [Francis Ysidro Edgeworth] had been of the kind that produces treatises, he would doubtless have published, some time between 1900 and 1914, a large volume in five books entitled *Mathematical Psychics*", dedicated "to the measurement of utility or ethical value, to the algebraic or diagrammatic determination of economic equilibriums, to the measurements of belief or probability, to the measurements of evidence or statistics, and to the measurements of economic value or index numbers." But when asked by Keynes why he had never ventured on a treatise, Edgeworth answered "with his characteristic smile and chuckle that large-scale enterprise, such as treatises and marriage, had never appealed to him." So, Edgeworth substituted these five books with hundreds of papers written in a style considered as a paradigm of complexity (Henderson 1993). Consequently he was condemned to be rediscovered years later from his contribution as a *precursor* of some important ideas in economics, statistics and –I hope- the philosophy of probability.

Among the five books, the least known is the third dedicated to the measurement of probability. This theme is central in Edgeworth's work not only for the number of pages dedicated to it in his bibliography (Baccini 2003), but also because the philosophical notion of probability is necessary to understand his approach to economics and statistics. According to Arthur L. Bowley "Edgeworth […] continued to devote himself to the conception of pure probability, while the problems he attacked appeared to be used rather as illustrations of theory as having practical importance, though this appearance was often illusory." (1928: 2). His approach "was from the side of the philosophy of probability", and this is a problem for the diffusion of his ideas because "neither philosophers nor mathematicians have been much interested in this inquiry, perhaps because the philosophers were not mathematicians, and the mathematicians who were philosophers were concerned rather with the foundations of mathematics, which did not involve questions of probability." (Bowley 1928: 3)

The most complete discussion of Edgeworth's theory of probability is again the one dispersed in Keynes's *Treatise on Probability*; but despite this, in the debate about Keynes's *Treatise* (e.g. Bateman 1987, Carabelli 1988, Runde-Mizuhara 2003) the place for Edgeworth is very small

(Baccini 2004); some remarks can be found also in works dedicated to Edgeworth's economics and statistics (McCann 1996; Mirowki 1994), and theory of choice under uncertainty (Baccini 2001).

The working hypothesis of this paper is that Edgeworth's *Treatise of Probability* exists (hereinafter TP); and that a coherent interpretation of probability is the starting point of this book.[1] The aim is to expose synthetically Edgeworth's theory of probability with particular attention to his empirical interpretation of probability statements, or to his semantics of probability. The basic idea is that only after having reconstructed the logical foundation of probability and its interpretation, it is possible to define correctly the use of probability for statistical induction and, more generally, its role for scientific reasoning (Chatterje 2003: 36).

The paper is organized as follows: in the first section the modality of definition of primitive probabilities is discussed; in the second Bayes's theorem, rule of succession and the nature of prior probability involved in them; the third discusses the epistemic notions of credibility and belief; in the fourth some conclusions are suggested.

1. Primitive probabilities

Frequency statements of probability. The landscape of Edgeworth's TP is characterized by the "epistemological flog" of the nineteenth century debate on the nature of probability, in which questions relating to calculus interwove with philosophical ones, and with problems of the application of statistical techniques to empirical data of natural and social sciences. *The Logic of Chance* by John Venn had a fundamental relevance in this debate.[2] It canonically systematised the frequentist theory of probability (Chatterjee 2003; Galavotti 2005; Gillies 2000), and was the principal reference text for Edgeworth's TP.

In this respect, an useful distinction must be made between *primitive* probabilities and *complex* probabilities. Primitive probabilities refer to simple (atomic) events, as for example the probability that "the 6 of diamonds comes out from a pack with 52 cards"; complex probabilities

[1] The TP is contained in the pages of three journals: *Mind*, *Philosophical Magazine* and *Journal of the Royal Statistical Society*. The first two contains the articles dedicated to the "methaphysical roots", the third the ones dedicated to the "the mathematical branches" (Edgeworth 1884: 258).

[2] The first edition is of 1866; the second, the one utilised by Edgeworth, is that of 1876. We cite from the third edition of 1888.

result from the application of the rules of calculus on primitive probabilities, as for example the probability that "a red card comes out" (Galavotti 2005: 39-40).

In the frequentist tradition, probability can be defined only in the context of a random experiment "whose outcome is unpredictable beforehand and which is capable of indefinite repetition under identical conditions, at least conceptually" (Chatterjee 2003: 37). The totality of possible outcomes is the outcome space, and probabilities are assigned to various sets – called the events- in this outcome space. Primitive probabilities are empirically measurable as the relative frequency of events, when the experiment is repeated a large number of times.

Venn in his *Logic of Chance* wrote that the probability of an event presupposes the existence of an adequately long series of events belonging to at least two distinct classes, in order to possess certain characteristics or not. A series is defined in reference to a large number of events/objects which are not necessarily temporally ordered. It combines the irregularity of the individuals with the regularity of the aggregate (Venn 1888: 4), and represents "the ultimate basis upon which all the rules of Probability must be based" (Venn 1888: 2). To determine the existence of the series and the numerical proportion of their characteristic properties, it is necessary to resort to experience: "Experience is our sole guide. If we want to discover what is in reality a series of *things*, not series of our own conceptions, we must appeal to the things themselves to obtain it, for we cannot find much help elsewhere. […] [T]he series we employ are ultimately obtained by experience only" (Venn 1888: 74-75).

Taking up a suggestion of Venn, Edgeworth referred probability to two different spheres, one *objective*, relative to the frequencies observed in a certain phenomenon, and one *subjective*, relative to the mental condition associated with those frequencies: "Probability may be described, agreeably to general usage, as importing partial incomplete belief […] as differing somehow in degree from perfect belief or rather credibility. […] Thus the object of the calculus is probability as estimated by statistical uniformity." (Edgeworth 1884: 223)

Let us consider the objective part for now, i.e. the reflections on the frequentist nature of probability. Edgeworth's aim was to remove the limitations on the theory imposed by Venn. Edgeworth, following Venn, considers as primitives of his probability theory the notions of event and series of events. Starting from events and series he defines primitive probabilities "as measured by the number of times that the event is found by experience to occur, in proportion to the number of times that it might possibly to occur." (Edgeworth 1911: 376; but see also 1884: 223; 1899: 208). These probabilities can be then manipulated according to the rules of calculus. According to Edgeworth, and following Venn, probability is based on empirical knowledge and never loses reference to facts.

To fix the idea, we shall write $P(R, A) = p$ where A indicates the outcome space –for example the repetition of a chemical experiment, or the English population of 1894- and R indicates the number of the elements of a set –an event- composed by individuals which share a same property relevant in the experiment –for example the result of the chemical experiment or the number of people age 65 or more in the English population. The proportion of events with result R is approximated – with a certain margin of error – to p; this proportion is called the (statistical) probability of R in A. That primitive probability is relative to a body of evidence which depends on the empirical recognizability of a set R of events, the *series* in Venn's and Edgeworth's terms, and of a subset A in R.

So defined, probability is relative not to the individual events, but to the frequency of a modality assumed by certain events in a long series. For example: it has been ascertained, by means of repeated extractions, that a certain urn contains white balls (W) and black balls (B). In particular, it has been ascertained that the proportion between the white and black balls is $7:3$, i.e. out of ten balls extracted, 7 are white and 3 are red. Therefore, we can write $P(W, (W + B)) = 0.7$ in a long series; which can be read: in a long series of extractions, the probability that white balls will be drawn is 0.7. On the basis of this definition, it is obvious that no rule of inference exists that allows us to arrive at the conclusion that "the probability that the ball that we are extracting is white is 0.7". This is *meaningless* within the frequentist theory: in fact, every single ball is white or black; the probability is a property attributed to the series of extractions, but not to the individual ball (Chatterje 2003: 40; Cohen, 1989: 48-49; Galavotti 2005: 83).

A priori probability. In Venn's theory the body of evidence utilizable in fixing initial probabilities is very small, based as it is only on experienced statistical evidence. Edgeworth's aim was to enlarge this body of evidence in respect to Venn's position. For this it is necessary to reply to the following question: on what evidential basis are the facts identified to measure probability

(Edgeworth 1911: 376)? According to Edgeworth's it is possible to modify the frequentist view starting from the idea that "the series which is regarded as the empirical basis of probability is not a simple matter of fact" (Edgeworth 1911: 377 §3). Humans only know from experience, but it is necessary to enlarge the domain of experience with respect to the position to which Venn had relegated it: "I only contend that Mr. Venn […] has not made the foundation wide enough, and that therefore he is unable to carry up the structure to the full height of generality. He is unable to rise an axiom of equal distribution of quantity in general, above the view that, in the absence of any such [specific] information, we are entirely in the dark."(Edgeworth 1884A: 160-161)

To correct this, beside statistical probabilities, Edgeworth introduces in his TP primitive *a priori* probability, that is probabilities "not determined by statistics", (Edgeworth 1884B: 204), or probabilities measured "when probability, founded upon statistical fact […] has reached the utmost degree of tenuity" (Edgeworth 1884: 229). *A priori* probability emerges for example when we consider the probability that the 6 will come out in a long series of throws of a dice that has never been previously thrown; in the games of chance, and more in general, in reference to the probabilities assigned in the *classical* or Laplacean domain of probability.

A priori probability emerges when we have probability defined in reference to a sort of experiential knowledge which is handed down through heredity, and which belongs to the species before they belong to the individual. As the inherited memory of past experience of pleasure is at the basis of Edgeworth's *Mathematical Psychics*, the inherited experiential knowledge is at the basis of Edgeworth's TP (Baccini 2007).

The introduction of inherited or antenatal experience V makes it possible to establish primitive *a priori* probability values of the form: $P_{apriori}\left(R|V\right)=v$ which is read: the primitive *a priori* probability of R, conditioned to knowledge V derived from inherited experience, is v.

There are a lot of examples of primitive a priori probabilities in TP. The first and most recurrent one regards the measurement process for physical sizes: the *a priori* probability that a certain value of the measurement is correct is equal to that of any other value within a certain interval (Edgeworth 1884: 231; 1911: 377 §7). In the case of a barometric measurement, for example, certain pressure values do not have the same probability of being registered with respect to others: the postulate required is that the equi-probability among the possible values exists with respect to a small interval, not among very different or very extreme values (Edgeworth 1911: 377 §7). A second example concerns a universally-shared mathematical experiment relative to the distribution of figures in the numbers “which come under our notice (1884A: 159) or in the expansion of π (Edgeworth 1884: 229-230; 1911: 378 §10; 1921: 83; 1922A: 486-487). In both cases, it is expected, “with the conviction that greatly exceeds every specific verification” (Edgeworth 1911: 378), that a figure appears in the same proportion as every other one, or that any figure whatsoever can be found in every position: the rule is the randomness of the appearances (Edgeworth 1922: 261). In this sense, the experiments of R.A. Proctor (1874: 100) and of Edgeworth himself (Edgeworth 1884B: 159; cf. also the long note in 1922A: 487), on the basis of the calculations of the value of π by William Shanks (1853), only fortify a generalised conviction.

The integration of the experience as defined by Venn with the antenatal experience, permits a recover of the a priori primitive probability and the principle of sufficient reason as founding elements of Edgeworth’s TP. The recovery of the principle of sufficient reason was realized with a modification of his traditional-Laplacean epistemological interpretation. Its epistemological foundation is not justified as a practical rule in the absence of empirical knowledge, but as an ampliative induction. The principle is justified as a particular case of the procedures of empirical measurements. According to Edgeworth, the experiences of humans on measurement, point out that measurables can assume any value in a given interval. Analogously, if statistical evidence is lacking, the principle of sufficient reason suggests that each event of a set has the same probability. (Edgeworth 1884: 231 (1); 1884A: 158; 1884B: 204; 1911: 377 §§ 7, 8, 9.)

The principle is not utilized to draw knowledge from ignorance as in the Laplacean tradition, because it is justified from an experiential knowledge about the conditions of its applicability: “the equal distribution of *a priori* probability (in the absence of specific knowledge) rests on a rough but solid basis of experience.” (Edgeworth 1921: 82; 1884°: 160). Consider for example the problem of measuring the weight of an object; *a priori* we know that in a given interval the probability of a single point is the same of all other points of the same interval. This practical principle has its origin not from the ignorance of the observer, but from her knowledge; she *knows* that in the long run, repeating for many times the same measurement, in a given interval, every value has the same probability to be determined.

In Edgeworth's TP there are two kinds of primitive probabilities: probabilities derived by statistical knowledge, and a priori probabilities derived from antenatal knowledge. Complex probabilities are the result of the application of the calculus to primitive probabilities.

2. Inverse probability

"The most perfect types of probability" present "the two aspects: proportion of favourable cases given *a priori* and frequency of occurrence observed *a posteriori.*" (Edgeworth 1911: 377). This is the inverse probability "that is, crudely speaking, the arguing from observed events to the probability of the causes" (Dale 1991: xi). The principal instrument for this kind of inference is Bayes's theorem, the first definition of which appeared in 1763 in a memoir of Thomas Bayes, and in a work independently developed by Laplace and published in 1774.

Many modern discussions about the foundations of probability and statistics are derived from different interpretation of Bayes's theorem. A standard formulation of the theorem is the following. Give the complete set of incompatible and exhaustive events $H = (H_1, H_2, ..., H_n)$, the conditional probability of H_i given an event *A* with $P(A) > 0$ is $P(H_i|A) = \dfrac{P(H_i)P(A|H_i)}{\sum_{i=1}^{n} P(H_i)P(A|H_i)}$, i = 1,2,..., n.

$P(H_i)$ is the *a priori* (prior) probability of H_i and *A* is independent from H_i. The conditional probability $P(H_i|A)$ is the probability of H_i given *A* and it is called *a posteriori* (posterior) probability. Bayes's theorem states how it is possible to derive the posterior probability from the a priori probability $P(H_i)$, and the conditional probability $P(H_i|A)$ that is the probability of *A* given H_i. If we interpret H_i as a set of causes of *A*, Bayes's theorem permits to infer the probability of the hypothesis or causes of the event *A* from the a priori probability of H_i and the empirical observation of the occurrences of *A* and H_i together.

The modern formulation of Bayes's theorem is generally demonstrated by multiplicative rule, and, according to a reductionist view, it is only a special way of writing the conditional probability (Chatterjee 2003: 48-49; Feller 1968: 124-125). In fact given H_i and applying the multiplicative rule we have

$$P(AH_i) = P(A)P(H_i|A) = P(H_i)P(A|H_i), \text{ i} = 1,2,..., \text{n},$$

and

$$P(H_i|A) = \frac{P(H_i)P(A|H_i)}{P(A)},\ \mathrm{i} = 1,2,...,\mathrm{n}.$$

Considering that for

$$P(A) = \sum_{i=1}^{n} P(H_i)P(A|H_i),\ \mathrm{i} = 1,2,...,\mathrm{n};$$

with the necessary substitution we have Bayes's theorem.

If it is true that Bayes's theorem is a result of conditional probability, it is neverthless true that in the practical application of it we encounter problems relating to the modality in which prior probability are defined. In fact the passage from conditional probability $P(A|H_i)$ to posterior probability $P(H_i|A)$ is possible only if at least[3] the a priori probability $P(H_i)$ is known or determinable by a conventional procedure (Chateerjee 2003: 104-115; Cohen 1989: 24).

2.1. Edgeworth on Bayes's theorem

For Edgeworth the problem is clear: "The distinction [between inverse and direct probability] consists chiefly in the «arbitrary assumptions» which are required in inverse probability with respect to the *a priori* probability that the causes which are the subject of inquiry exist and act." (Edgeworth 1908: 652; but see also 1884: 228; 1911: 377 §13)

This is the crucial point regarding the acceptability of Bayes's theorem in a frequentist theory of probability. Venn denied that it is possible to define by experience a priori probability and then excluded Bayes's theorem and inverse probability from his theory (Dale 1991: 324-326); in his frequentist approach, a priori probability can be defined only as an epistemic or subjective measure, depending on the status of non experiential knowledge of the observer; for this posterior probability also looses its objective characterization, and must be banned from probability theory.

The enlarged notion of experience allowed Edgeworth to define cogently and experientially primitive a priori probability and then to save Bayes theorem and inverse probability in a frequentist framework. In Edgeworth's TP the derivation of the posteriori distribution is simply an exercise in inductive inference; the prior distribution being objective, the posterior probability has the same objective interpretation.

---

[3] Given that by multiplicative rule, the a priori probability *P(A)* can be eliminated

In an entire booklet of his TP, entitled *Metretike* (1887), Edgeworth dedicated a deep discussion about the nature of probabilities involved in Bayes's theorem. The booklet has a very strange construction for modern readers, based as it is on variations on the analogy between the measurement of probability in probability theory and of utility in economics. It is based on a simple formalization of Bayes's Theorem presented in reference to a John Stuart Mill's example (Mill 1843: 543). The case is the one in which is known that the given effect *E* must have been produced by one or the other of the two causes *A* and *B*, that is $A \cap B = 0$ and $A \cup B = \Omega$; α and β are the a priori probability of respectively *A* and *B*; $P(E|A) = p$ and $P(E|B) = q$ the conditional probability that a cause "if it [is] existed, would have produced the given effect" (Edgeworth 1887: 2). The probability that the given effect has been produced by *A* is then $\frac{\alpha p}{\alpha p + \beta q}$ where $\alpha + \beta = 1$.

According to Edgeworth, it is necessary to consider two different problems: how to measure *α* and *β*, and the degree of precision of this measurement. As we know, a priori probability is attained by antenatal experience; but antenatal experience does not allow us to reach the degree of precision ordinarily attained by statistical data. Only in the case of games of chance is it always possible to have a precise measure of *a priori* probability; normally we have a priori probability with a lesser degree of precision. The point is clear in the Millian example discussed by Edgeworth: after *n* throws of a coin, heads have turned up *n* times. What is the probability that the coin is loaded? The set of possible causes is composed by (i) *A*: the coin is symmetrical, hence the outcome of the throws is the result of chance; and (ii) *B*: the coin is loaded, hence the outcome is not the result of chance only. The a priori probability of *A* and *B*, respectively *α* and *β*, depend by the state knowledge about the coin. If the coin is manufactured by the English mint, we have reason to think that *α* is near to 1; if the coin is antique, *β* is considerable. We have a similar problem also with *q*, that is the conditional probability of *n* heads in succession with a loaded coin. Provided that we have no antecedent experience in throwing this coin, "in any case, our knowledge of α and β [and q] is not numerical" (Edgeworth 1887: 4). We know with accuracy only the conditional probability p of the succession of heads if the coin is symmetrical: $p = 2^{-n}$. Then it is possible to write the probability that the outcome is the result of throws of a symmetrical coin as $\frac{\alpha 2^{-n}}{\alpha 2^{-n} + \beta q}$; and analogously $\frac{\beta q}{\alpha 2^{-n} + \beta q}$ is the probability that the result is caused by a loaded coin. If it is known that the coin is antique, β is considerable and there are no reasons to think that q is very small; so it is possible to infer that the probability of the outcome was the result of the loaded coin is greater than that of a symmetrical coin. "In general, unless we have antecedent certainty that the coin is symmetrical, the stage is soon reached […] at which the chance of irregularity in the die is greater than that of the orther alternative" (Edgeworth 1887: 4).

In this problem there is a "compound" of probabilities, defined in reference to different knowledges, and with different degrees of precision. The fact that it is impossible to measure with the same degree of precision all the probabilities involved, does not impede the process of drawing a solid inference from the application of Bayes's theorem.

A third problem, relevant for our reasoning is the following: we have seen that the set A and B is exhaustive; the choice of an exhaustive set can be interpreted differently. If the exhaustive condition is interpreted as the complete enumerability of all the possible causes –in the example, A and B are all the possible causes of the succession of heads -, then for the application of Bayes's theorem it is necessary to know in advance all the possible causes and their prior probabilities. If, otherwise, the exhaustive condition is interpreted as exhaustive in relation to a given set of knowledge, then the conditions for applying Bayes's theorem are much looser: it is necessary to know only a subset, at least one, of the possible causes, and to attribute to a fictitious cause (all the causes explicitly considered are not real causes of the event) all other unknown causes. In other words: it is possible to apply Bayes's theorem without the complete knowledge of all the causes of the event considered.

In the first case, it is impossible to improve the status of knowledge, given that the researcher knows in advance all the possible causes, and it is unclear how new information must change the prior; in the second case there exists an informative dynamic. The Edgeworthian position in this respect is ambiguous; he chooses in every case examples in which the condition of exhaustivity in the first sense is fulfilled; but he recovers the dynamic nature of knowledge in a more general context, considering the informal, gradual, cumulative and auto corrective nature of statistical knowledge: an exhaustive set of causes, given a certain status of knowledge, become incomplete if the status of knowledge changes: "a course [that] violates established Bayesian credo" (Chatterjee 2003: 123).

As we have seen, according to Edgeworth the passage from *direct* to *inverse* probability requires no new principles than the fundamental ones of the theory of probability. The same reasoning is true when it is necessary to calculate the probability of future effects from the probability of known causes. In fact Bayes's theorem can be utilized "to deduce from the frequently experienced occurrence of a phenomenon the large probability of its recurrence." (Edgeworth 1884: 228-229).

To illustrate the preceding points, it is useful to follow the example of the double extraction of random digits (Edgeworth 1911: 382 § 44).[4]

A digit is taken at random from a mathematical table (logarithm or from expansion of π); after having taken a second digit, we are interested in the event that the sum of the two extraction is greater than 9. In this problem there is a "compound" of probabilities: *a priori* in the first extraction; conditional for the result of the double extraction. Indicating, in modern notation, with $x_i = i-1$ for i=1, 2,...,10 the first digit, $P(x_i)=\frac{1}{10}$ is the *a priori* probability of x_i. Indicating with $y_i = i-1$ for i=1, 2, ...,10 the digit of the second extraction, and with V the event $x_i+y_i>9$; the conditional probabilities are $P(V|y_i)=\frac{i-1}{10}$ (as in the second row of the table 1). In this context it is possibile to state a problem of inverse probability as: given V, calculate the probability of the digit y_i in the second extraction. Applying the Bayes's theorem, it is possible to calculate the inverse or *a posteriori* probability that the event V is the result of the *cause* "the digit y_i has been extracted":

$$P(y_i|V)=\frac{P(y_i)P(V|y_i)}{\sum_{i=1}^{10}P(y_i)P(V|y_i)}=\frac{\frac{1}{10}P(V|y_i)}{0+\left(\frac{1}{10}\right)\left(\frac{1}{10}\right)+\left(\frac{1}{10}\right)\left(\frac{2}{10}\right)+\ldots+\left(\frac{1}{10}\right)\left(\frac{9}{10}\right)}=\frac{\frac{1}{10}\left(\frac{i-1}{10}\right)}{\frac{45}{100}}=\frac{i-1}{45}$$

These probabilities are reported in the third row of the table 1.

But if Bayes's theorem can be cogently utilized in the theory of probability in reasoning about the *probability of the causes*, it is possible also to consider it in reference to the problem of the *future effects* deduced by known causes, or more generally "to deduce from the frequently experienced occurrence of a phenomenon the large probability of its recurrence." (Edgeworth 1884:

[4] In the TP there are a lot of examples of Bayes's theorem. The loaded coin in Edgeworth 1887; the Englishman in Madeira, drawn from Venn 1888: 222-228; the urn with balls in unknown proportion in Edgeworth 1911: 382 § 45 drawn from Laplace 1886: 185-189; the probability of testimony Edgeworth 1911: 383, §§48-53).

228-229). Also in this case, the aim of Edgeworth's reasoning is to demonstrate that it is not necessary to introduce new principles in the theory of probability.

To illustrate this point, it is useful to recur to a variation of the preceding example: given the event *V*, calculate the probability that (in the future) taking at random a third digit, the sum of the three is greater than 10. Indicating the result of the third extraction with $z_i = i - 1$ for i=1,2, ..., 10; with $P(z_i) = \frac{1}{10}$ the *a priori* probability of each z_i, the problem is to calculate $P(W|V)$ where W is the event $y_i^* + z_i > 10$ and $y_i^* = x_i + y_i > 9$.

The problem can be solved in three steps. The first step consists in calculating $P(W|y_i)$, i.e. the probability of *W* conditional to any result of the second extraction y_i and independently to the result of the first extraction (fourth row of the table 1); the second step: given that $P(y_i|V)$ is known, it is possible to calculate for each $y_i^* = x_i + y_i > 9$, the probability that $y_i^* + z_i > 10$ for each z_i (fifth row of the table 1). This probability is a compound of the *a posteriori* probability $P(y_i|V)$, and the conditional probability $P(W|y_i)$, that is:

$$P(W|V) = \sum_{i=1}^{10} P(W|y_i)P(y_i|V) = \frac{240}{450} = \frac{8}{15}.$$

Edgeworth concludes this example stating: "it may be expected that actual trial would verify this result." (Edgeworth 1911: 384 §53)

Table 1. Probabilities in the double extraction of random digits.

x_{i+1}	0	1	2	3	4	5	6	7	8	9
$P(v\|y_i)$	0	1/10	2/10	3/10	4/10	5/10	6/10	7/10	8/10	9/10
$P(y_i\|V)$	0	1/45	2/45	3/45	4/45	5/45	6/45	7/45	8/45	9/45
$P(W\|y_i)$	0	0	1/10	2/10	3/10	4/10	5/10	6/10	7/10	8/10
$P(W\|y_i)P(y_i\|V)$	0	0	2/450	6/450	12/450	20/450	30/450	42/450	56/450	72/450

As it can be seen, no new principles are necessary for drawing predictive inferences starting from probabilistic statements, it is sufficient to iterate in the calculus the application of inverse probability and conditional probability. Bayes's theorem justifies syntactically and deductively prediction stated in probabilistic form.

3. *Epistemic interpretation of probability*

Edgeworth's antenatal experience permits to open the frequentist tradition to the a priori probability, and then to introduce the Bayes's theorem in the *more objective* domain of probability, as results from the rules of the calculus. The problem is, now, the evaluation of probability statements in terms of belief: all correct probability statements –strictly frequentist, a priori, a posteriori, or inverse- are syntactically or logically true; but an agent does not necessarily have a belief in them. The fact that a probability statement is correct, not necessarily allows people to construct a belief in it; it is in fact necessary an evaluation or translation mechanism of the probability statements in terms of belief, and this mechanism is external to the theory of probability.

In more abstract terms: according to Edgeworth, it is necessary to distinguish between a *more objective* level of probability -in which correct probability statements can be manipulated according to correct rules to infer new probability statements-, and a *more subjective* one, in which probability statements are evaluated in epistemic terms.

This is a strong innovation in respect to the standard Vennian frequentist approach. According to Venn, inverse probability is based on belief of the agent, and then, belonging to the subjective sphere, must be banned from the proper domain of probability. The theory of probability cannot be utilized to make ampliative inductions starting from experiential data; it is only a correct system of devices used to systematize them. Bayes's theorem based as it is on non experiential knowledge, cannot be utilized to make inferences on the domain of probability, it is a *wrong* device. According to Edgeworth, instead, Bayes's theorem is based on experiential knowledge, so it belongs to the objective part of probability and it is only necessary to specify the correct syntactical domain of its application.

More precisely, Edgeworth flanked probability with two measures of the epistemic sphere: the notions of credibility and belief, both belonging to the subjective sphere, and differing from each other "somehow in degree" (Edgeworth, 1884: 223). In accordance with Venn's lesson, Edgeworth did not intend to investigate their origin, or the concrete – physiological or psychological – modalities by means of which individuals construct judgements of belief and credibility. However, he wanted to furnish a minimal formulation for them.

Credibility is the result of the transposition, on the epistemic level, of a probability value. Edgeworth suggested that an automatic correspondence exists between probability values and degrees of credibility, or, more precisely, that the latter depend on the former: “the credibility of an event […] is measured by the frequency with which in the long run such an event occurs. The correspondence between the external measure and the «inner feeling» measured is indeed rough and loose, like the correspondence between the degrees of a thermometer and the sensation of heat: accepted not *semper*, not *ab omnibus*, not by every fevered patient, yet with sufficient generality to be of common use, for instance in the warming of a public library.”(Edgeworth, 1908: 384)[5]

The notion of credibility can be written as $Cr(R|P(R,A)=p)=c=p$ which signifies: the degree of credibility of *R*, given its statistical frequency, is *c*, and this value coincides with probability *p*. The credibility function in fact operates a transformation of a relative frequency in a random experiment or trial in an epistemic measure indicating the degree of credibility associated with that relative frequency.

Once it is established that the values of probability and credibility coincide, the problem of defining the measure of the credibility can be reduced to defining the measure of the probability. The domain over which it is possible to define probability values is more extensive than the one Venn established, as it includes – as we have said – direct probabilities, *a priori* probabilities, and probabilities derived by Bayes’s theorem. Therefore, while they are situated on different cognitive levels, probability and credibility should enjoy the same mathematical characteristics and obey the same axioms.

A very different treatment is reserved to belief. Like Venn, Edgeworth picked up the notion of belief developed by Alexander Bain, according to which belief is *preparedness to act* (Edgeworth, 1883: 434; 1884: 223; 1887: 53). That is, belief represents the basis – not merely cognitive - with reference to which an individual decides to act; it is the device that permits an individual to choose one action out of a set of possible actions. In this Bainian framework belief is identical with volition; the test for belief is action. The construction of belief is a process in which

[5] Automatism – it, too, perhaps intuitionist – is entirely similar to the one we find for belief in Keynes. Precisely in commenting on *A Treatise on Probability*, Edgeworth noted, “[…] I have suggested [Edgeworth, 1884: 225] that the first principles of credibility are like those of conduct which according to Mill do not admit of proof in the ordinary sense.” (Edgeworth, 1922: 260). The last reference is to Mill, 1864: 6 (1861) and, more generally, to all of Chapter 4.

psychological determinants are of fundamental importance; it is based on complex sets of information, comprehending all the knowledge accumulated and inherited by individuals, *and* on convictions derived from spheres of the intellective activity not relative to cognitive processes. Statistical evidence is but an element of all that; and for evaluating statements of probability in terms of belief, probability theory is absolutely useless.

The end of the story is clear: if probability is not identical with belief, probability is not linked with volition of agents; and probability is not a guide for action. Imagine an agent betting on the next throw of a loaded die; according to frequentist theory, it is possible to say only that "the next throw is a member of a long sequence in which the relative frequency of 6 approaches $\frac{1}{4}$". Is this information sufficient to guide her bet? Or is it better that she follows her night dream in which 3 will come out in the next throw? Probability theory is not a criterion to choose between the two actions (Baccini 1997, 2001; Chatterjee 2003: 40). It is the belief of the agent that will guide his choice; not probability nor credibility.

Venn had affirmed that probability is not a measurement of belief. Edgeworth built a more complex mechanism: by distinguishing between credibility and belief, he left to the former the strictly epistemic sphere, and to the second the complexity of the motivations for human actions.

4. Conclusion

At this point, we must dwell on the usual exercise of classifying Edgeworth's probability according to a some classification. The choice is not simple because we have a lot of different classifications, for example the ones of Jonathan Cohen (1989), Gillies (2000), Chatterjee (2003), or Galavotti (2005), which complicate substantially the simplest and traditional one between objective and subjective probabilities.

In the existing literature, Edgeworth is classified differently by different authors. Classic interpreters as Keynes (1921) and Bowley (1928) considered him a frequentist with some special features. The modern ones developed alternative interpretations: according to Mirowski Edgeworth "never really believed the frequentist doctrine"; he "was one of the last of the classical probabilists", given the "deep dependence" of his doctrine on the principle of sufficient reason

(Mirowski 1994: 97 n. 26). McCann, probably on the basis of a hurried reading of Bowley (1928: 7-8), considered Edgeworth's theory of probability as a subjectivist one, "defined with respect to a logic of partial or incomplete belief, in opposition to the then-dominant frequency interpretation identified with […] Venn" (1996: xix).

The picture drawn in this paper credits Edgeworth's theory of probability of the 'virtue of eclectism', according to Chatterjee's expression (2003: 138). In fact, in our reconstruction, Edgeworth defines a physical-experiential basis on which to found the theory of probability. With respect to the 19th-century frequentist tradition and mainly to Venn's *Logic*, in TP we face a weaker foundation: the ascertainment of a plurality of forms of experience justifies the utilisation of a plurality of primitive probabilities, defined directly in reference to experiential data collected in statistical tables, or defined *a priori* in reference to antenatal experience. Starting from these probabilities, and applying the rules of calculus, it is possible to infer complex probabilities. In particular, as we have seen, if it is possible to define prior probabilities, it is also possible to use Bayes's theorem to infer posterior probability, remaining in a strict experiential framework.

The second point noted in this paper is that, according to Edgeworth probability values can be interpreted in epistemic terms: when an agent knows a probability statements, she has also a measure of its credibility; and this credibility obeys the same axioms of probability. But credibility is not belief, it is not the reason which enables the agent to decide her action; it is not a measure of volition of the agent, but only of her probable knowledge. Probability, therefore, is not useful for the theory of choice under uncertainty (Baccini 1997, 2001).

Edgeworth's TP is based, at the end, on the idea that probability is "Janus-faced. On the one side it is statistical […], on the other is epistemological" (Hacking 1975: 12). In the philosophy of science, at least starting from Carnap (1950; Galavotti 2005: 164-169), this kind of view of probability is labelled as *pluralism*, and is considered antithetical to *dogmatism* (Costantini-Geymonat 1982; Cohen 1989: 81-114), that is to the theories which asserts that probability has a single and well-defined meaning, while all the other interpretations must be considered false. In the dogmatic framework, after an explicit definition of probability has been introduced, all the principles and theorems were obtained from it. In this sense, both Laplace, Venn, Keynes, and

Bruno De Finetti -respectively from classical-objective, frequentist, logicist, and personal subjectivist perspectives- were dogmatic.

According to our reconstruction, Edgeworth can be numbered among the anti-dogmatics, together with Antoine A. Cournot, Frank P. Ramsey, and Carnap. The passage from the dogmatic Vennian position to Edgeworth's eclectism is the crucial point for the development of statistical inference. Because it is possible to develop the theory of statistical inference as Edgeworth did, the underlying probability cannot suffer from the limitation imposed to it by Venn. The diffusion of statistical methods based on probability was the necessary condition for its application to the social sciences where the possibility of controlled experiments is very reduced (Stigler 1986). Edgeworth made a central contribution to this diffusion. Edgeworth's statistics was no longer the deterministic one of William Stanley Jevons or that of Alfred Marshall (Baccini 2006). The central point – which is new in the panorama of the social sciences – becomes, "under what circumstances does a difference of figures correspond to a difference of fact?" (Edgeworth 1884C: 38). To adventure in this new direction, Edgeworth needed a theory of probability that was, so to speak, completely free of the narrow limits imposed to it by Venn. Edgeworth constructed this new theory, and applied it extensively.

References

BACCINI 1997
Baccini, A., "Edgeworth on the fundamentals of choice under uncertainty", *History of Economic Ideas,* 2: 7-31

BACCINI 2001
Baccini, A., "Frequentist probability and choice under uncertainty", *History of Political Economy*, 33.4: 743-772

BACCINI 2003
Baccini, A., "Bibliography of Edgeworth's writings", in Newman, P. (a cura di), *F.Y. Edgeworth's* Mathematical Psychics *and Further Papers on Political Economy*, Oxford, Oxford University Press: 621-647.

BACCINI 2004
Baccini, A., "High pressure and black clouds: Keynes and the frequentist theory of probability", *Cambridge Journal of Economics,* 28, 5: 653-666.

BACCINI 2006
Baccini, A., “Francis Ysidro Edgeworth” in T. Raffaelli, G. Becattini and M. Dardi (eds.), *The Elgar Companion to Alfred Marshall*, Edward Elgar.

BACCINI 2007
Baccini, A., “Edgeworth on the foundations of ethics and probability”, *European Journal for the History of Economic Thought*, forthcoming

BATEMAN 1987
Bateman, B.W., "Keynes's changing conception of probability", *Economics and Philosophy*, 3 (1): 97-119.

BOWLEY 1928
Bowley, A.L., *F.Y. Edgeworth's Contributions to Mathematical Statistics,* London, Royal Statistical Society. (Clifton, N.J., A.M. Kelley, 1972)

CARABELLI 1988
Carabelli, A.M., *On Keynes's Method*, London, Macmillan.

CARNAP 1950
Carnap, R., *Logical Foundations of Probability*, Chicago, Chicago University Press, 1950.

CHATTERJEE 2003
Chatterjee, S.K., *Statistical Thought: a Perspective and History*, Oxford, Oxford University Press.

COSTANTINI-GEYMONAT 1982
Costantini, D., Geymonat, L., *Filosofia della probabilità*, Milano, Feltrinelli.

DALE 1991
Dale, A.I., *A History of Inverse Probability*, New York, Springer Verlag.

EDGEWORTH 1883
Edgeworth, F.,Y., "The physical basis of probability", *Philosophical Magazine,* XVI: 433-435

EDGEWORTH 1884
Edgeworth, F.,Y., "The philosophy of chance", *Mind*: 223-235.

EDGEWORTH 1884A
Edgeworth, F.,Y., "Chance and law", *Hermathena:* 154-163.

EDGEWORTH 1884B
Edgeworth, F.,Y., "*A priori* probabilities", *Philosophical Magazine*: 204-210.

EDGEWORTH 1884C
Edgeworth, F.,Y., Review of Jevons, W.S., *Investigations in Currency and Finance*, Academy, June 19, 637: 38-39

EDGEWORTH 1887
Edgeworth, F.,Y., *Metretike or the Method of Measuring Probabilities and Utilities*, London, Temple Company

EDGEWORTH 1899
Edgeworth, F.,Y., "Probability" in Palgrave, R.H.I. (ed.), *Dictionary of Political Economy*, London, Macmillan and Co., vol. III: 208.

EDGEWORTH 1908
Edgeworth, F.,Y., "On the probable errors of frequency constants", *Journal of the Royal Statistical Society*, 1908: 381-397, 499-512, 651-661, 662-678; 1909: 81-90.

EDGEWORTH 1911
Edgeworth, F.,Y., "Probability" in *Encyclopaedia Britannica*, XIth ed., 1911: 376-403.

EDGEWORTH 1921
Edgeworth, F.,Y., “Molecular statistics”, *Journal of the Royal Statistical Society*, LXXXIV: 71-89.

EDGEWORTH 1922
Edgeworth, F.,Y., "The philosophy of chance", *Mind,* 1922: 257-283.

EDGEWORTH 1922A
Edgeworth, F.,Y., “Molecular statistics”, *Journal of the Royal Statistical Society*, LXXXV: 479-487

FELLER 1968
Feller, W., *An Introduction to Probability Theory and Its Applications*, New York, London, Sydney*, John Wiley & Sons*.

GALAVOTTI 2005
Galavotti, M.C., *Philosophical Introduction to Probability*, Stanford, CSLI.

GILLIES 2000
Gillies, D., *Philosophical Theories of Probability*, London, Routledge.

HACKING 1975
Hacking, I., *The Emergence of Probability*, Cambridge, Cambridge University Press

HALD 1998
Hald, A., *A History of Mathematical Statistics from 1750 to 1930*, New York, John Wiley

HENDERSON 1993
Henderson, W., "The problem of Edgeworth 's style" in Backhouse, R., Dudley-Evans, T., Henderson, W. (eds.), *Economics and Language*, London, Routledge: 200-222

JONATHAN COHEN 1989
Jonathan Cohen, L., *An Introduction to the Philosophy of Induction and Probability*, Oxford, Clarendon Press.

KEYNES 1926
Keynes, J.M., *A Treatise on Probability*, London, Macmillan, 1921

KEYNES 1926
Keynes, J.M., "Obituary: Francis Ysidro Edgeworth, 1845-1926", *Economic Journal,* 1926 (36): 140-153. [*Essays in Biography*, vol X. dei *Collected Writings*, London, Macmillan, 1972: 251-266]

LAPLACE 1814
Laplace, P.S., *Essai Philosophique sur les Probabilités*, Paris, 1814 (citation from the sixth edition of 1840)

LAPLACE 1886
Laplace, P.S., *Théorie analytique des Probabilités*, in *Œvres complètes de Laplace*, Paris, Gauthiers-Villars, vol. 7, 1886.

MCCANN 1996
McCann, C.R. (ed.), *F.Y. Edgeworth: Writings in Probability, Statistics and Economics*, Cheltenham, Edward Elgar.

MILL 1843
Mill, J.S., *A System of Logic Ratiocinative and Inductive*, London, J.W. Parker (1846[2]) [in *Collected Works*, voll. VII-VIII, edited by J.M. Robson, University of Toronto Press, Routledge & Kegan Paul, 1973]

MILL 1864
Mill, J.S., *Utilitarianism,* London, 1864[2]

MIROWSKI 1994
Mirowski, P. (ed.), *Edgeworth on Chance, Economic Hazard and Statistics*, Lanham, Rowman & Littlefield Publishers.

PROCTOR 1874
Proctor, R.A., *The Universe and the Coming Transits*, London.

SHANKS 1853
Shanks, W., *Contributions to Mathematics, Comprising Chiefly the Rectification of the Circle to 607 Place of Decimals*, London.

STIGLER 1986
Stigler, S.M., *The History of Statistics*, Cambridge (Mass.) and London, Belknap Press.

VENN 1888
Venn, J., The *Logic of Chance*, London, Macmillan.