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

Alberto Baccini

"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" (Keynes [1926] 1972, 262). But Edgeworth was not of the kind that writes treatises because, as he answered "with his characteristic smile and chuckle" to a question posed by John Maynard Keynes, "large-scale enterprise[s], such as treatises and marriage, had never appealed to him" (261). So he left only a slim volume titled *Mathematical Psychics*, along with hundreds of articles dispersed in various journals.

The main thesis of this article is that Edgeworth's actual writings on probability can be treated as, in effect, constituting one volume of the five-book treatise suggested by Keynes. Then we will refer to the dozens of papers dedicated by Edgeworth to probability as his *Treatise on Probabilities*.[1]

Correspondence may be addressed to Alberto Baccini, Università di Siena, Dipartimento di Economia Politica, Via P.A. Mattioli, 10 I53100 Siena, Italy; e-mail: baccini@unisi.it

1. The *Treatise on Probabilities* is contained in the pages of three journals: *Mind*, *Philosophical Magazine*, and the *Journal of the Royal Statistical Society*. The first two contain the articles dedicated to the "metaphysical roots" of probability; the third, the ones dedicated to its

History of Political Economy 41:1 DOI 10.1215/00182702-2008-039

For modern readers, accustomed to rediscovering Edgeworth as a precursor of some important ideas in economics or statistics, his contribution to the theory of probability is the least known part of his work. In the copious literature on the history and philosophy of probability, Edgeworth is hardly mentioned (e.g., Howson 1995; Hald 1998; Gillies 2000; Chatterjee 2003; Galavotti 2005). Some remarks on Edgeworth's theory of probability can be found in works dedicated to his economics and statistics (McCann 1996; Mirowski 1994) and theory of choice under uncertainty (Baccini 2001). But the most complete discussion of Edgeworth's theory of probability is again the one found in Keynes's *Treatise on Probability.* Still, although Keynes's *Treatise* discusses at length Edgeworth's theory of probability, the secondary literature on Keynes's 1921 work (e.g., Bateman 1987; Carabelli 1988; Runde-Mizuhara 2003) hardly mentions Edgeworth at all (Baccini 2004).

The theory of probability is central in Edgeworth's work not only for the number of pages dedicated to it in his bibliography, but also because the philosophical notion of probability is necessary for understanding his approach to economics and statistics. According to Arthur L. Bowley (1928, 2), "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." 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" (3).

The working hypothesis of this essay is that a coherent interpretation of probability is the starting point of Edgeworth's contribution to the theory of probability. My 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 is it possible to define correctly the use of probability for statistical induction and, more generally, its role in scientific reasoning (Chatterjee 2003, 36).

"mathematical branches" (Edgeworth 1884c, 223). The complete list of papers published by Edgeworth in the three journals can be found in Baccini 2003.

Primitive Probabilities

Frequency Statements of Probability

Edgeworth's intellectual landscape is characterized by the epistemological fog of the nineteenth-century debate on the nature of probability, in which questions relating to calculus of probabilities were interwoven with philosophical ones, and with problems of the application of statistical techniques to empirical data of the natural and social sciences. *The Logic of Chance* (1888) by John Venn had a fundamental relevance in this debate.[2] It canonically systematized the frequentist theory of probability (Chatterjee 2003, 282–83; Galavotti 2005, 74–81) and was the principal reference text for Edgeworth's *Treatise on Probabilities.*

In this respect, a 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 result from the application of the rules of calculus of probabilities 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 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 that are not necessarily temporally ordered. It combines the irregularity of the individual events or objects with the regularity of the aggregate (Venn 1888, 4) and represents "the ultimate basis upon which all the rules of Probability must be based" (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

2. The first edition of Venn's text was published in 1866; the second, the one used by Edgeworth, is that of 1876. We cite from the third edition of 1888.

Bayes's theorem can be used "to deduce from the frequently experienced occurrence of a phenomenon the large probability of its recurrence" (Edgeworth 1884c, 228–29).

To illustrate the preceding points, it is useful to follow the example of the double extractions of random digits (Edgeworth 1911, 382).[3] A digit is taken at random from a mathematical table (either from a logarithm table or from a table showing computations of π); after having taken a second digit, we are interested in the event that the sum of the two extractions 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. In modern notation, let $x = i - 1$ for $i = 1, 2, \ldots, 10$, be the first digit, $P(x_i) = (i - 1)/10$, is the a priori probability of x_i. Let $y_i = i - 1$ for $i = 1, 2, \ldots, 10$ be the digit of the second extraction, and V the event $xi + yi > 9$; the conditional probabilities are $P(V \mid y_i) = (i - 1)/10$ (as in the second row of the table 1). In this context it is possible to state a problem of inverse probability as follows: given V, calculate the probability of the digit yi in the second extraction. Applying 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":

$$P(y_i \mid V) = \frac{(y_i)\, P(V \mid y_i)}{\sum_{i=1}^{n} P(y_i)\, P(V \mid y_i)}$$

$$= \frac{\frac{1}{10} P(V \mid y_i)}{0 + \left(\frac{1}{10}\right)\left(\frac{1}{10}\right) + \left(\frac{1}{10}\right)\left(\frac{2}{10}\right) + \ldots + \left(\frac{1}{10}\right)\left(\frac{9}{10}\right)}$$

$$= \frac{\frac{1}{10}\left(\frac{i-1}{10}\right)}{\frac{45}{10}} = \frac{i-1}{45}.$$

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

But if Bayes's theorem can be cogently used in the theory of probability in reasoning about the *probability of the causes*, it is possible also to con-

3. In the *Treatise on Probabilities* there are a lot of examples of Bayes's theorem, among them the loaded coin in Edgeworth 1887; the Englishman in Madeira, drawn from Venn 1888, 222–28; the urn with balls in unknown proportion in Edgeworth 1911, 382, drawn from Laplace 1886, 185–89; and the probability of testimony in Edgeworth 1911, 383.

Table 1 Probabilities in the Double Extractions of Random Digits

	Result of the First Extraction i									
	$i = 1$	$i = 2$	$i = 3$	$i = 4$	$i = 5$	$i = 6$	$i = 7$	$i = 8$	$i = 9$	$i = 10$
x_{i-1}	0	1	2	3	4	5	6	7	8	9
$P(V\|y_i)$	0	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{3}{10}$	$\frac{4}{10}$	$\frac{5}{10}$	$\frac{6}{10}$	$\frac{7}{10}$	$\frac{8}{10}$	$\frac{9}{10}$
$P(y_i\|V)$	0	$\frac{1}{45}$	$\frac{2}{45}$	$\frac{3}{45}$	$\frac{4}{45}$	$\frac{5}{45}$	$\frac{6}{45}$	$\frac{7}{45}$	$\frac{8}{45}$	$\frac{9}{45}$
$P(W\|y_i)$	0	0	$\frac{1}{10}$	$\frac{2}{10}$	$\frac{3}{10}$	$\frac{4}{10}$	$\frac{5}{10}$	$\frac{6}{10}$	$\frac{7}{10}$	$\frac{8}{10}$
$P(W\|y_i)P(y_i\|V)$	0	0	$\frac{2}{450}$	$\frac{6}{450}$	$\frac{12}{450}$	$\frac{20}{450}$	$\frac{30}{450}$	$\frac{42}{450}$	$\frac{56}{450}$	$\frac{72}{450}$

sider it in reference to the problem of the *future effects* deduced by known causes. A variation of the preceding example can be used to illustrate the point. Given the event V, that is, a double extraction with a total result greater than 9, calculate the probability that (in the future) taking at random a third digit, the sum of the three is greater than 10. Let the result of the third extraction be $z_i = i - 1$ for $i = 1, 2, \ldots, 10$, the prior probability of z_i is $P(z_i) = 1 / 10$, and the problem is to calculate $P(W \mid V)$ where W is the event $x_i + y_i + z_i > 10$.

The problem can be solved calculating first $P(W \mid y_i)$, that is, the probability of W conditional on any result of the second extraction y_i, and independently of the result of the first extraction (fourth row of table 1). Now it is possible to calculate for each double extraction with a result greater than 9, the probability that a third extraction gives a result greater than 10. This can be done multiplying the a posteriori probability $P(y_i \mid V)$ and the conditional $P(W \mid y_i) : P(W \mid y_i)P(y_i \mid V)$ (fifth row of table 1). The solution of the problem is then $P(W \mid V) = \sum_{i=1}^{10} P(W \mid y_i)P(y_i \mid V) = (240 / 450) = (8 / 5)$. According to Edgeworth (1911, 384), "It may be expected that actual trial would verify this result."

As can be seen, no new principles are necessary for drawing predictive inferences starting from probabilistic statements; it is sufficient to

According to our reconstruction, Edgeworth can be numbered among the antidogmatics, together with Antoine A. Cournot, Frank P. Ramsey, and Rudolf Carnap. The passage from the dogmatic Vennian position to Edgeworth's eclecticism 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 on 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 slight indeed. As documented by Stephen M. Stigler (1978, 1986), Edgeworth played a central role in the historical development of the theory and the applications of statistical inference techniques to the social sciences. In fact Edgeworth's statistics was no longer the deterministic one of William Stanley Jevons or Alfred Marshall (Baccini 2006). The central question—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 1884d, 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 on it by Venn. Edgeworth constructed this new theory, and applied it extensively.

References

Baccini, A. 1997. Edgeworth on the Fundamentals of Choice under Uncertainty. *History of Economic Ideas* 2:7–31.

———. 2001. Frequentist Probability and Choice under Uncertainty. *HOPE* 33.4:743–72.

———. 2003. Bibliography of Edgeworth's Writings. In *F. Y. Edgeworth's "Mathematical Psychics" and Further Papers on Political Economy*, edited by P. Newman, 621–47. Oxford: Oxford University Press.

———. 2004. High Pressure and Black Clouds: Keynes and the Frequentist Theory of Probability. *Cambridge Journal of Economics* 28.5:653–66.

———. 2006. Francis Ysidro Edgeworth. In *The Elgar Companion to Alfred Marshall*, edited by T. Raffaelli, G. Becattini, and M. Dardi, 578–82. Cheltenham: Edward Elgar.

———. 2007. Edgeworth on the Foundations of Ethics and Probability. *European Journal of the History of Economic Thought* 14.1:79–96.

Bateman, B. W. 1987. Keynes's Changing Conception of Probability. *Economics and Philosophy* 3.1:97–119.

Bowley, A. L. 1928. *F. Y. Edgeworth's Contributions to Mathematical Statistics*. London: Royal Statistical Society.