SANDY L. ZABELL

THE RULE OF SUCCESSION

1. INTRODUCTION

Laplace's rule of succession states, in brief, that if an event has occurred m times in succession, then the probability that it will occur again is $(m+1)/(m+2)$. The rule of succession was the classical attempt to reduce certain forms of inductive inference – "pure inductions" (De Morgan) or "eductions" (W. E. Johnson) – to purely probabilistic terms. Subjected to varying forms of ridicule by Venn, Keynes, and many others, it often served as a touchstone for much broader issues about the nature and role of probability.

This paper will trace the evolution of the rule, from its original formulation at the hands of Bayes, Price, and Laplace, to its generalizations by the English philosopher W. E. Johnson, and its perfection at the hands of Bruno de Finetti. By following the debate over the rule, the criticisms of it that were raised and the defenses of it that were mounted, it is hoped that some insight will be gained into the achievements and limitations of the probabilistic attempt to explain induction. Our aim is thus not purely – or even primarily – historical in nature.

As usually formulated, however, the rule of succession involves some element of the infinite in its statement or derivation. That element is not only unnecessary, it can obscure and mislead. We begin therefore by discussing the finite version of the rule, its statement, history, and derivation (sections 2–3), and then use it as a background against which to study the probabilistic analysis of induction from Bayes to de Finetti (sections 4–9). Sections 4–6 deal largely with historical issues; sections 7–9 matters mathematical and foundational.

2. THE FINITE RULE OF SUCCESSION

One form of enumerative induction involves performing an experiment that can, at least in principle, be repeated an indefinitely large number of times ("trials"), with one of two possible outcomes ("suc-

Erkenntnis **31**: 283–321, 1989.

cess" vs. "failure"). In this case it makes sense to refer to the (unknown) probability p of success, i.e., the limiting frequency, propensity, or objective chance of success. Under the classical Laplacean analysis, if the trials are independent, and all possible values of p are assumed equally likely, then given r successes in m trials, the probability of a success on the next trial is

$$\int_0^1 p^{r+1}(1-p)^{m-r}dp \Big/ \int_0^1 p^r(1-p)^{m-r}dp = (r+1)/(m+2).$$

This is Laplace's *rule of succession.*[1]

For certain types of enumerative induction the Laplacean schema is unsatisfactory. If one is noting the color of ravens, tagging each one after its color is recorded, then the universe being sampled is finite, and the sampling is being done without replacement (i.e., each raven is observed at most once). For this reason, in 1918 the English philosopher C. D. Broad repeated the Laplacean analysis, but for sampling from a finite urn without replacement (the Laplacean picture can be thought of, in a way that can be made mathematically precise, as sampling from an urn with an infinite number of balls). Of course, there are questions about the extent to which observing the color of ravens corresponds to sampling balls from an urn (realistically, one only sees ravens in one's neighborhood) – important questions, and ones also considered by Broad – but let us set these aside for the moment and consider Broad's simple mathematical question:

Consider an urn with a finite but unknown number of balls n, each of which is either black or white. Suppose a sample of m balls is drawn at random without replacement from the urn. If nothing is known about the relative proportion of black and white balls, and all m of the balls drawn are black, what is the probability that the next ball drawn is black?

Of course, some assumption must be made about the prior probability for the proportion of blacks. The natural assumption, in analogy to the Laplacean treatment, is that all possible proportions j/n are equally likely, and this is the one that Broad made in 1918.[2] Broad discovered that, surprisingly, the answer does not depend on n, the population size, but only on m, the sample size, and that the answer is *identical* to Laplace's rule, i.e., $(m+1)/(m+2)$.

The proof is not difficult. A simple application of Bayes's theorem shows that the desired probability is

$$\frac{\sum_{j=m+1}^{n} j(j-1)(j-2)\dots(j-m)}{(n-m)\sum_{j=m}^{n} j(j-1)(j-2)\dots(j-m+1)}.$$

The problem thus reduces to the evaluation of two sums, and, as Broad notes, "it can easily be shown that" their ratio is $(m+1)(m+2)$. If the sum in the denominator is denoted $S_{m,n}$, then a simple inductive argument shows that

$$S_{m,n} = \frac{(n+1)!}{(m+1)(n-m)!},$$

and substitution then yields

$$\frac{(n-m)^{-1} S_{m+1,n}}{S_{m,n}} = \frac{(m+1)}{(m+2)}.$$

Broad did not give the mathematical details, and for completeness a proof is given in the appendix at the end of this paper.

The finite rule of succession has several important philosophical consequences:

(1) It eliminates a variety of possible concerns about the occurrence of the infinite in the Laplacean analysis (see, e.g., Kneale 1949, p. 205): attention is focused on a finite segment of trials, rather than a hypothetical infinite sequence.

(2) The frequency, propensity, or objective chance p that occurs in the integration is replaced by the fraction of successes; thus a purely personalist or subjective analysis becomes possible, and objections to "probabilities of probabilities" and "unknown probabilities" (see, e.g., Keynes 1921, pp. 372–75) are eliminated.

(3) It extends the domain of applicability of the rule to forms of enumerative induction not previously covered.

An important consequence of Broad's analysis was the remark that the probability of a universal generalization – i.e., that all n balls in the urn are black, given that the first m were – will be quite small unless m

is large relative to n (the exact probability is $(m+1)/(n+1)$). This was not a novel observation, but it was viewed at the time as a serious setback to the Laplacean program of justifying induction probabilistically, and was an important impetus for the early work of Jeffreys and Wrinch (1919). This question will be discussed in the final sections of the paper.

Historical Note. Although Broad is often credited with the finite rule of succession (e.g., by von Wright 1957; Jeffreys 1961; Good 1965, p. 18), he does not specifically claim priority in its derivation, and in fact it had been independently discovered several times prior to Broad's 1918 paper. The first of these was in 1799, in a paper by the Swiss mathematicians Pierre Prevost (1751–1839) and Simon L'Huilier (1750–1840). Both were interested in the philosophical implications of probability and wrote several papers on the subject in collaboration; see generally Todhunter (1865, pp. 453–463).[3]

As Prevost and L'Huilier state the problem,

> Soit une urne contenant un nombre n de billets; on a tiré $p+q$ billets, dont p sont blancs and q non-blancs (que j'appellerai noirs). On demande les probabilités que les billets blancs and les billets noirs de l'urne étoient des nombres donnés, dans la supposition qu'à chaque tirage on n'a pas remis dans l'urne le billet tiré.

Thus, Prevost and L'Huilier consider the more general case of p successes and q failures in $p+q=m$ trials, and derive the posterior probabilities for different constitutions of the urn. The law of succession is then derived as a consequence, with the result that the probability of a success on the next trial is $(p+1)/(m+2)$.

The result was later independently derived by Ostrogradskii (1848), as well as "a mere diocesan" (Keynes 1921, p. 179), Bishop Charles Terrot of Edinburgh, whose work (Terrot 1853) is mentioned by Boole in his *Investigation of the Laws of Thought* (1854). These early derivations are not without interest, and are discussed in the mathematical appendix at the end of this paper.

The result is also noted by the indefatigable Todhunter, who reports the work of Prevost and L'Huilier in his famous *History of the Mathematical Theory of Probability* (1865, pp. 454–57). Todhunter observes that the crucial sum may be readily evaluated by the use of the binomial theorem, remarks the identity of the answer with the Laplacean one, and comments that "the coincidence of the results obtained on the two different hypotheses is remarkable".[4]

3. FINITE EXCHANGEABILITY AND THE RULE OF SUCCESSION

Although the Prevost-L'Huilier and later proofs of the finite rule of succession are not difficult, they leave unexplained this "remarkable coincidence". It turns out that there is a very different approach, involving the concept of exchangeability, which clarifies why the finite and infinite rules agree.

Let $X_1, X_2, \ldots, X_n$ denote a sequence of exchangeable random variables taking on the values 0 and 1. By definition this means that the probability distribution of the random variables is invariant under permutations; i.e.,

$$P[X_1 = e_1, X_2 = e_2, \ldots, X_n = e_n]$$
$$= P[X_1 = e_{\sigma(1)}, X_2 = e_{\sigma(2)}, \ldots, X_n = e_{\sigma(n)}],$$

for all possible sequences $e_1, e_2, \ldots, e_n$ ($e_i = 0$ or 1), and permutations σ of $\{1, 2, \ldots, n\}$. There is a simple representation for such sequences. If $S_n = X_1 + X_2 + \cdots + X_n$, then the events $\{S_n = k\}$ form a partition, i.e., they are mutually exclusive and exhaustive (they are disjoint and one of them must occur). Thus, by the so-called theorem on total probability, one may write

$$P[X_1 = e_1, X_2 = e_2, \ldots, X_n = e_n]$$
$$= \sum_{k=0}^{n} P[X_1 = e_1, X_2 = e_2, \ldots, X_n = e_n | S_n = k] P[S_n = k].$$

By the definition of exchangeability, the conditional probabilities $P[X_1 = e_1, X_2 = e_2, \ldots, X_n = e_n | S_n = k]$ assign equal probabilities to the ${}_nC_k$ sequences of k 1s and $n-k$ 0s. This corresponds to drawing at random all n balls out of an urn containing k 1's and $n-k$ 0s, i.e., it is a *hypergeometric probability* which we will denote $H_{n,k}$. Let $p_k = P[S_n = k]$. The sequence $p_0, p_1, \ldots, p_n$ specifies the probabilities that the sequence $X_1, X_2, \ldots, X_n$ will contain $0, 1, \ldots$ or n 1s respectively. In this notation,

$$P = \sum_{k=0}^{n} p_k H_{n,k}.$$

That is, the exchangeable probability P may be viewed as a mixture of the hypergeometric probabilities $H_{n,k}$, using the p_k. If one were to arrange $n+1$ urns $U_0, U_1, \ldots, U_n$, with urn U_k containing k 1s and $n-k$ 0s, pick an urn U_k with probability p_k, and then draw all n balls at random out of the urn, the resulting probability distribution on sequences of length n would be identical with the original probability assignment P.

This simple, but very useful result, is the *finite de Finetti representation theorem*. Note the following basic properties of the representation:

FE1. The $H_{n,k}$ are independent of P; P only enters into the representation via the p_k.

FE2. The representation is unique: if $P = \sum p_k H_{n,k} = \sum q_k H_{n,k}$, then $p_k = q_k$, all k.

FE3. The probability distribution on sequences of length n, arising from any mixture $\sum p_k H_{n,k}$, is exchangeable. (The term mixture means that the p_k are arbitrary numbers satisfying $0 \leqslant p_k \leqslant 1$ and $p_0 + p_1 + \cdots + p_n = 1$.)

In honor of those who were the first to study it, let us call the sequence generated by picking one of the $n+1$ urns U_k at random, and then drawing all n balls out of the urn at random, the *Prevost–L'Huilier process*, denoted for short as PL_n. The Prevost–L'Huilier process is a special example of a finite exchangeable sequence, with the $p_k = P[S_n = k] =: 1/(n+1)$ uniform. It is a consequence of FE1 that an exchangeable sequence is uniquely determined once the values $p_k = P[S_n = k]$ are specified.

Now we are ready to explain the strange coincidence of the rules of succession for the Prevost–L'Huilier process PL_n and the Bayes–Laplace process BL_∞, which is generated by picking p uniformly from the unit interval $[0, 1]$ and then tossing a p-coin infinitely often. The Bayes–Laplace process $X_1, X_2, X_3, \ldots$ is an infinitely exchangeable sequence; i.e., for any $n \geqslant 1$, the initial segment of the process $X_1, X_2, \ldots, X_n$ is exchangeable in the sense defined above. Thus it has some finite de Finetti representation $\sum p_k H_{n,k}$. But, the Bayes–Laplace process BL_∞ has the property that $p_k = P[S_n = k] = 1/(n+1)$, just as does the Prevost–L'Huilier process PL_n. Since they are both exchangeable, and since their mixing measures coincide, they are *identical*. That is,

> the initial segment $X_1, X_2, \ldots, X_n$ of the Bayes–Laplace process BL_∞ is stochastically identical to the Prevost–L'Huilier process PL_n.

Now it is clear why the rules of succession for the two processes coincide: they are actually the *same* process (up to stage n)! Not only do their rules of succession coincide, but every other probabilistic aspect as well. Although the two processes were generated by two distinct stochastic mechanisms, the resulting distributions are identical.

In retrospect, this is obvious: if we are given the initial probabilities $P[X_1 = e_1]$, and the rules of succession at each stage, it is possible to express the probabilities $P[X_1 = e_1, X_2 = e_2, \ldots, X_n = e_n]$ in terms of these quantities. For example, for both PL_4 and BL_∞,

$$\begin{aligned} P[X_1 = 1, X_2 = 0, X_3 = 1, X_4 = 1] &= (1/2)(1/3)(2/4)(3/5) \\ &= 1/20. \end{aligned}$$

Thus, if the initial probabilities and succession probabilities of two processes coincide, the processes are the same. For those allergic to exchangeability arguments of the type given above, this gives an alternative way of deriving the identity of PL_n and the initial nth segment of BL_∞ once their rules of succession have been shown to coincide.

Observing the identity of the two processes has the advantage that most properties of PL_n may be immediately and easily deduced. For example, consider the following question:

> Given a sequence of total length N, what is the probability that if the first n outcomes are all black, then the remaining $N - n$ outcomes will also be all black?

That is, how much evidence does the first n outcomes being black provide towards the universal generalization that all outcomes are black? Doing this directly (as Broad did) is elementary but messy, involving the usual sums. Far easier, however, is the observation that

$$\begin{aligned} P[S_N = N \mid S_n = n] &= P[S_N = N \text{ and } S_n = n]/P[S_n = n] \\ &= P[S_N = N]/P[S_n = n] \end{aligned}$$

$$= \frac{1}{(N+1)} \Big/ \frac{1}{(n+1)}$$

$$= \frac{(n+1)}{(N+1)},$$

which is the answer Broad derives, and which coincides (as it must) with the result for the Bayes–Laplace process (Laplace, *Théorie analytique*, p. 402; De Morgan 1838, p. 64).

How satisfactory an explanation of enumerative induction does the rule of succession provide? What are its limitations? Can these be eliminated? Broad's analysis came at the end of a century and a half of discussion and debate. It marks the end of an era, for in a few years the contributions of Keynes, Johnson, Ramsey, and de Finetti were to irretrievably change the way in which the problem was cast. The next three sections discuss some of the highlights of the preceding debate from Bayes to Broad. Those readers not interested in this previous history may turn directly to Section 7, where the emphasis shifts to the philosophical analysis and mathematical evolution of the rule.

4. WHEN AND WHY DID BAYES PROVE BAYES'S THEOREM?

Hume first stated the problem of induction; Bayes first advanced a solution to it. The chronological link between these two events is much closer than is usually recognized.

Like James Bernoulli before him, the Reverend Thomas Bayes perished before he was published. At some time prior to his death on 17 April 1761, Bayes wrote his famous 'Essay Towards Solving a Problem in the Doctrine of Chances', published posthumously by his friend Richard Price in 1764. Although Bayes's introduction to his essay has not survived, Price tells us that Bayes came to have doubts as to the validity of the postulate adopted by him in the solution of the problem. As Price puts it, Bayes "afterwards considered, that the postulate on which he had argued might not perhaps be looked upon by all as reasonable; and therefore he chose to lay down in another form the *proposition* in which he thought the solution of the problem is contained, and in a *scholium* to subjoin the reasons why he thought so, rather than to take into his mathematical reasoning any thing that might admit dispute".

For this reason some latter commentators have assumed that Bayes delayed publication of his results because of such doubts (e.g., Fisher 1973, pp. 9–10). How long did Bayes meditate on his solution? Surprisingly, there is evidence that suggests that Bayes may have arrived at at least the basic results in his essay some fifteen years prior to his death.

The first piece of evidence in question is a passage from David Hartley's *Observations on Man*, published in 1749. After discussing the law of large numbers for binomial trials given by De Moivre, Hartley states

> An ingenious Friend has communicated to me a Solution of the inverse Problem, in which he has shewn what the Expectation is, when an Event has happened p times, and failed q times, that the original Ratio of the Causes for the Happening or Failing of an Event should deviate in any given Degree from that of p to q. And it appears from this Solution, that where the Number of Trials is very great, the Deviation must be inconsiderable: Which shews that we may hope to determine the Proportions, and, by degrees, the whole Nature, of unknown Causes, by a sufficient Observation of their Effects. (Hartley 1749, p. 339)

If Hartley's ingenious friend were Bayes, this would mean that Bayes had arrived at his basic results no later than 1749, and probably somewhat earlier. The identity of the two is not only a natural conjecture, it is supported by the internal evidence of Hartley's own statement: the terminology used by Hartley is identical to that employed by Bayes, who refers in his essay to an "event... happening p times, and failing q times...". (Ingenious, moreover, was a word which came readily to mind when thinking of Bayes. Price, for example, calls Bayes "one of the most ingenious men I ever knew" (Price 1758, p. 248), and Laplace refers to Bayes's method as "très ingénieuse" (Laplace 1814, p. cxlviii).)

If Bayes did suppress his result for some 15 years, his diffidence in publication might well explain the anonymous nature of Hartley's reference. Since Bayes and Hartley were both members of the Royal Society and dissenters, they may well have known each other, although there is no direct evidence that they actually did. It is of course possible that Hartley's "ingenious friend" was someone other than Bayes, but absent Hartley's direct statement to this effect or clear evidence that Bayes's work had been independently duplicated, it is hard to credit this hypothesis.[5]

Very recently a new piece of evidence has come to light that seems decisive in favor of Hartley's reference to Bayes. Dr. A. I. Dale has discovered a passage in an early notebook of Bayes giving a proof of one of the rules in Bayes's essay (Dale 1986). Although the entry is undated, it is preceded by one dated July 4, 1746, and succeeded by one dated December 31, 1749. It is thus clear that at some point in the period 1746–1749 Bayes had derived at least some of his basic results, and the coincidence with the date of Hartley's book (1749) seems too striking to be coincidental.

What event in the period 1746 to 1749 led Bayes to investigate a problem that, in the words of his friend Richard Price, must be "considered by any one who would give a clear account of the strength of *analogical* or *inductive reasoning*"? Thus put, an obvious answer suggests itself. In 1748 David Hume had published his *Enquiries Concerning Human Understanding*, containing a clear and succinct statement of his famous problem of induction. Hume had laid down the challenge: "Let any one try to account for this operation of the mind upon any of the received systems of philosophy, and he will be sensible of the difficulty" (*Enquiry*, p. 59). Bayes may have answered it within a year.

Bayes's paper, however, had little immediate, direct influence and it is through Laplace that the techniques of inverse probability became widely known. A decade after the appearance of Bayes's essay, Laplace wrote the first of a series of papers in which he, apparently independently of Bayes, presented his solution to the problem of causes, in the form that was to gain widespread acceptance (Laplace 1774).[6] His older mentor Condorcet, recognizing the importance of Laplace's contribution to the inductive problem, rushed it into print. "The problem of finding the probability of the occurrence of an event, given only that it has occurred a number of times in the past, is the most fundamental in the calculus of probabilities, argued the assistant secretary [Condorcet], underlining the significance of Laplace's paper in the preface to the sixth volume of the *Mémoires des savants étrangers*" (Baker 1975, pp. 168–69). Hume's impact had been felt on the Continent as well.[7]

Laplace's own statement of the probabilistic solution of the problem of induction appears in the *Essai philosophique*. The example he provided is notorious:

Thus we find that an event having occurred successively any number of times, the probability that it will happen again the next time is equal to this number increased by unity divided by the same number, increased by two units. Placing the most ancient epoch of history at five thousand years ago, or at 1826213 days, and the sun having risen constantly in the interval at each revolution of twenty-four hours, it is a bet of 1826214 to one that it will rise again tomorrow. [Laplace, *Essai*, p. xvii]

5. THE RISING OF THE SUN

It is said that Laplace was ready to bet 1,826,214 to 1 in favor of regular habits of the sun, and we should be in a position to better the odds since regular service has followed for another century. A historical study would be necessary to appreciate what Laplace had in mind and to understand his intentions. (Feller 1968, p. 124)

Laplace has perhaps received more ridicule for this statement than for any other. Yet Feller, despite his general lack of sympathy for the Bayesian position, had too much respect for Laplace to dismiss his famous calculation unexamined. Let us attempt the study Feller suggests.

To begin with, it is important to realize that the example of the rising of the sun does not originate with Laplace. It goes back to Hume (at least), who in his *Treatise* of 1739 asserted: "One wou'd appear ridiculous, who wou'd say, that 'tis only probable the sun will rise to-morrow, or that all men myst dye; tho' 'tis plain we have no further assurance of these facts, than what experience affords us" (*Treatise*, p. 124). As we shall see, the example of the rising of the sun as a touchstone of inductive inference is a common thread through much of the latter literature on the subject.

In denying that inferences such as the rising of the sun are merely probable, Hume was arguing that there are degrees of knowledge which, while not demonstratively certain, exceed all probability. This is a recurrent idea, which can also be found, for example, in Cardinal Newman's *Grammar of Assent*. Price, to contradict Hume, turns to this example in his appendix to Bayes's essay:

Let us imagine to ourselves the case of a person just brought forth into this world, and left to collect from his observation of the order and course of events what powers and causes take place in it. The Sun would, probably, be the first object that would engage his attention; but after losing it the first night he would be entirely ignorant whether he should ever see it again. He would therefore be in the condition of a person making a first experiment about an event entirely unknown to him. But let him see a second appearance or one return of the Sun, and an expectation would be raised in him of a second *return*, and he might know that there was an odds of 3 to 1 for *some* probability of this. This odds would increase, as before represented, with the number of returns to which he was witness. But no finite number of returns would be sufficient to produce absolute or physical certainty. For let it be supposed that he has seen it return at regular and stated intervals a million of times. The conclusions this would warrant would be such as follow. There would be the odds of the millioneth power of 2, to one, that it was likely that it would return again at the end of the usual interval.

This is *not* Laplace's rule of succession, but rather a calculation of the posterior probability that the unknown chance p of the sun's rising exceeds 1/2, i.e.,

$$P[p > 1/2] = \int_{1/2}^{1} p^{n-1}dp \Big/ \int_{0}^{1} p^{n-1}dp$$
$$= 1-(1/2)^n = (2^n-1)/2^n.$$

i.e., odds of 2^n to 1. (Note Price uses an exponent of $n-1$, since he considers the first trial to merely inform us that the event is possible; see Pearson 1978, pp. 368–69.)[8]

Although Price was a lifelong philosophical opponent of Hume, he read Hume carefully, and it is clear that his discussion of Hume's example was intended to rebut Hume's contention that "many arguments from causation exceed probability, and may be receiv'd as a superior kind of evidence. . . . which are entirely free from doubt and uncertainty." Indeed, not only does Price address Hume's example, but he goes on to stress that "instead of proving that events will always happen agreeably to [uniform experience], there will always be reason against this conclusion".[9]

But even if one concedes that our knowledge of future events such as the rising of the sun only admit of probability, there is a leap of faith in Price's argument. Price began his analysis by first considering "a solid or die of whose number of sides and constitution we know nothing; and that we are to judge of these from experiments made in throwing it", later explaining that he "made these observations chiefly because they are all strictly applicable to the events and appearances of nature". Condorcet, in his *Essai*, accepts this nexus without reservation:

Ainsi le motif de croire que sur dix millions de boules blanches mêlées avec une noire, ce ne sera point la noire que je tirerai du premier coup, est de la même nature que le motif de croire que le Soleil ne manquera pas de se lever demain, & ces deux opinions ne diffèrent entr'elles que par le plus ou le moins de probabilité. (Condorcet 1785, p. xi)

This was a sweeping claim, and it did not pass unchallenged. Prevost and L'Huilier, in a philosophical essay accompanying their paper on the finite rule of succession, soon took issue with Condorcet, arguing

La persuasion analogique qu'éprouve tout homme, de voir se répéter un événement naturel (tel que le lever du soleil), est d'un genre différent de la persuasion représentée par une fraction dans la théorie des probabilités. Celle-ci peut lui être ajoutée, mais l'une peut exister sans l'autre. Elles dépendent de deux orders de facultés différens. Un enfant, un animal éprouve la première, & ne forme aucun calcul explicite, ni même implicite: il n'y a aucune dépendance nécessaire entre ces deux persuasions. Celle que le calcul apprécié est raisonné, & même, jusqu'à un certain point, artificielle. L'autre est d'instinct & naturelle. Elle dépend de quelques facultés intellectuelles dont l'analyse n'est pas facile, & probablement in très-grande partie du principe de la liaison des idées. (Prevost and L'Huilier 1799a, p. 15)

This is one of the earliest arguments urging the distinction between induction ("la persuasion analogique") and probability ("une fraction dans la théorie des probabilités"), and it presages a debate that continued unabated through the next century. (For the possible influence of Prevost and L'Huilier on Mill, via Dugald Stewart, see Strong 1978, p. 35). Bertrand, for example, writing nearly a hundred years later in his distinctively acerbic French prose, singles out the same passage from Condorcet for criticism:

L'assimilation n'est pas permise: l'une des probabilités est objective, l'autre subjective. La probabilité de tirer la boule noire du premier coup est 1/10 000 000, ni plus ni moins. Quiconque l'évalue autrement se trompe. La probabilité pour que le Soleil se lève varie d'un esprit à l'autre. Un philosophe peut, sans être fou, annoncer sur la foi d'une fausse science que le Soleil va bientôt s'éteindre; il est dans son droit comme Condorcet dans le sien; tous deux l'excéderaient en accusant d'erreur ceux qui pensent autrement. (Bertrand 1907, p. xix.)

Many other examples could be adduced.[10] What is striking in many of these discussions is the virtual lack of serious argument. Positions are staked out, but there is often surprisingly little in the way of genuine analysis or critical discussion. (One exception is Bertrand 1907, pp. 173–74).

A common position was that such inductive inferences, even if "probable", could not be quantified – that what was in question was a species of *philosophical* probability, rather than *mathematical* probability. Strong (1978, p. 207, n. 5) cites an early example of this distinction in a rare work of K. H. Frömmichen of 1773. It will be apparent by now that the date is "no accident"; by this time the claims of probability in natural philosophy were beginning to provoke dissent.

Such considerations were not, however, foreign to Laplace. His rule of succession is an instrument for "pure inductions", or "eduction" as W. E. Johnson later termed them. That Laplace was not under the illusion that "hypothetical inductions" could also be so described is clear from the penultimate chapter of the *Essai philosophique*, "Concerning the various means of approaching certainty". At the end Laplace cautions.

It is almost always impossible to submit to calculus the probability of the results obtained by these various means; this is true likewise for historical facts. But the totality of the phenomena explained, or of the testimonies, is sometimes such that without being able to appreciate the probability we cannot reasonably permit ourselves any doubt in regard to them. In the other cases it is prudent to admit them only with great reserve.

De Morgan, too, later cautioned that "in the language of many, induction is used in a sense very different from its original and logical one .... What is now called induction, meaning the discovery of laws from instances, and higher laws from lower ones, is beyond the province of formal logic" (De Morgan 1847, p. 215). (Note from the title of his book that De Morgan includes probability within that province.)

Thus, when Laplace made his notorious remark in the *Essai philosophique*, he was writing against a background of 75 years of debate and discussion about inductive inference throughout which the example of the rising of the sun runs as a common and recurrent thread. In his dry style, Laplace omits virtually all reference to this previous debate.

How seriously did Laplace view the calculation itself? Certainly much less so than is usually implied. All too often it is cited out of context, for after the passage quoted, Laplace went on to immediately add:

But this number is incomparably greater for him who, recognizing in the totality of phenomena the regulatory principle of days and seasons ["connaissant par l'ensemble des phénomènes le principe régulateur des jours et des saisons"], sees that nothing at the present moment can arrest the course of it.

The point is clear: the calculation only establishes the probability that flows from the mere repetition of events.[11] And while Laplace did not belabor the point, he was far from the only one to make it. Price too had cautioned that "it should be carefully remembered that these deductions suppose a previous total ignorance of nature", and his fanciful narrative is clearly intended to stress the artificial nature of the assumption. When Quetelet gives a similar analysis for the rising of the tides, it is for someone who has never seen them before. The English logician and mathematician Augustus De Morgan, who played an important role in disseminating Laplace's work during the 19th century, also stressed the point, terming the rule of succession "the rule of probability of a *pure induction*", and adds that "the probabilities shown by the above rules are merely *minima* which may be augmented by other sources of knowledge" (De Morgan 1847, p. 215).

6. THE GREAT JEVONIAN CONTROVERSY

This then was the Laplacean contribution to the probabilistic analysis of enumerative induction. How did it fare during the 19th century?

The English logician William Stanley Jevons is often portrayed as the first important philosopher of science to systematically link probability and induction (Keynes 1921, p. 273; Madden 1960, p. 233; Heath 1967, p. 261; Lauden 1973). Indeed, in Laudan (1973), the history of the subject revolves around Jevons: why did inductive logicians and philosophers of science before Jevons spurn probability; why did another half-century have to pass after Jevons before the link between probability and induction was taken seriously? Laudan considers these issues, centering his discussion on Jevons's arguments in favor of the link, and its criticisms by the English logician John Venn.

Laudan's analysis is largely vitiated, however, by a surprising chronological error: he presents Venn's criticisms as – and apparently believes them to be – an attack on Jevons, despite the fact that the 1st edition of Venn's *Logic of Chance* appeared in 1866, eight years prior to the appearance of the 1st edition of Jevons's *Principles of Science* (1874). Although it is true that Venn made extensive revisions in the 2nd (1876) and 3rd (1888) editions of the *Logic*, the vital chapter on 'Induction and its Connection with Probability' goes back to the 1st, and while the 1888 edition of the *Logic* (which Laudan quotes) does refer on several occasions to Jevons's *Principles*, it does so only briefly: despite several passages where the wording has been recast, new material added, or the text shortened, the basic thrust and content of the chapter remains that of the 1st edition.

But if Venn was not, at least initially, directing his fire against

Jevons, who then? The answer is clearly the English mathematician and logician Augustus De Morgan. De Morgan was Laplace's most enthusiastic English advocate, the author of no fewer than three works during the decade 1838–1847 intended to popularize and spread the Laplacean approach to probability.[12] Indeed, De Morgan's *Formal Logic* of 1847 was the first English language textbook on logic to break with tradition by presenting probability as a branch of formal logic, a precedent followed by Boole several years later in the latter's *Investigation of the Laws of Thought* of 1854. Venn explicitly singles De Morgan out, saying that he would have felt no need to write *The Logic of Chance*, given De Morgan's writings on probability, save that he differed from De Morgan in too fundamental a way (Venn 1888, p. ix). (Jevons in fact was a student of De Morgan's, and it was from De Morgan that he learned probability theory.)

Jevons was thus not alone. The probabilistic basis of at least some forms of induction had been advocated prior to Jevons by Condorcet, Laplace, Lacroix, Quetelet, Herschel, and De Morgan, and after Jevons by Peirce, Pearson, and Poincaré. Jevons was neither the first to argue the connection, nor the first philosopher of science or inductive logician to do so, but among this latter tribe he was admittedly one of the few to do so. As Venn testifies,

> So much then for the opinion which tends to regard pure Induction as a subdivision of Probability. By the majority of philosophical and logical writers a widely different view has of course been entertained. They are mostly disposed to distinguish these sciences very sharply from, not to say to contrast them with, one another; the one being accepted as philosophical or logical, and the other rejected as mathematical. This may without offence be termed the popular prejudice against Probability. (Venn 1888, pp. 208–209)

"Why did we have to wait for Stanley Jevons, and C. S. Peirce, writing in the 1870s, rather than Hume in the 1740s or Mill in the 1840s, to find someone systematically arguing that inductive logic is based on probability theory?" (Laudan 1973, p. 429). For Hume, there is a simple answer: the necessary apparatus of inverse probability did not exist when he wrote his *Treatise* and *Enquiries*. As discussed earlier, both Bayes and Laplace were aware of the relevance of their contributions to the questions addressed by Hume.

But what of the period after Laplace? Even if one takes 1814, the year of publication of the *Essai philosophique* as a point of departure, what happened in the 60 years that elapsed before the publication of Jevons's *Principles*? That De Morgan should embrace the Laplacean position on induction is not surprising; as we have noted, De Morgan was Laplace's staunchest English advocate and his writings on probability were in large part a deliberate effort to bring Laplace's work to the attention of the English public.

But why were there so few others in the English philosophical community to embrace the Laplacean position? Here the answer is not complimentary to English philosophy: the mathematical prerequisites were such as to exclude most writers on the subject. On this we have the testimony of Venn himself, the archcritic of Laplacean probability:

> The opinion that Probability, instead of being a branch of the general science of evidence which happens to make much use of mathematics, *is* a portion of mathematics, erroneous as it is, has yet been very disadvantageous to the science in several ways. Students of Philosophy in general have thence conceived a prejudice against Probability, which has for the most part deterred them from examining it. As soon as a subject comes to be considered 'mathematical' its claims seem generally, by the mass of readers, to be either on the one hand scouted or at least courteously rejected, or on the other to be blindly accepted with all their assumed consequences. Of impartial and liberal criticism it obtains little or nothing. (Venn 1888, p. vii)

Interestingly, Venn sees as the most unfortunate result of this neglect the loss for probability rather than the loss for philosophy: "The consequences of this state of things have been, I think, disastrous to the students themselves of Probability. No science can safely be abandoned entirely to its own devotees." Probability is too important to be left to the mathematicians.

This then, was the background against which Jevons wrote. In truth, there is little new in Jevons, but despite his many weaknesses, he represents a clear and succinct statement of the Laplacean position. Nevertheless, nearly half a century was to pass before Jevons's program was to be pushed forward by philosophers such as Johnson, Broad, Keynes, and Ramsey.

This hiatus, however, is not surprising. During the decades immediately following the appearance of Jevons's book, epistemic probability was preoccupied. The two-pronged assault of Boole (on the logical front) and Venn (on the empirical front) had called into serious question the very foundations of the Laplacean edifice. Epistemic probability did not go under during this period (Zabell 1989), but it did have to put its foundational house in order before it could contemplate expanding its horizons. After the contributions of Johnson, Keynes, Ramsey, and de Finetti this became possible.

Although the old Laplacean approach to probability eventually died out, epistemic probability arose transfigured from its ashes. While some continued to defend the principle of indifference – indeed, some still do – the key step in this metamorphosis was the abandonment of uniform priors and, on the inductive front, any attempt at a unique quantitative explanation of inductive inference.

A complete account of this transformation has never been written, and would go far beyond the compass of the present study. But limiting our attention to charting the vicissitudes of the rule of succession throughout the following period provides the opportunity for a case study, highlighting in a microcosm many of the arguments and issues that arose in the broader debate.

7. DEATH AND TRANSFIGURATION

As the statistician R. A. Fisher once noted, the rule of succession is a mathematical consequence of certain assumptions, and its application to concrete examples can only be faulted when the examples fail to satisfy the presuppositions. Those presuppositions involve two distinct types of issues. At the level of balls in an urn, there is the assumption that the possible urn compositions are equally likely, i.e., the principle of indifference. And at the level of applying the resulting mathematics to the real world, there is the question of the relevance of the urn model. The attacks on the rule of succession involved both of these points.

7.1. *The principle of indifference*

The Achilles's heel of the rule of succession lies in its appeal to the principle of indifference. It assumes that all possible ratios are equally likely, and that in particular, on any single trial, the probability of an event "concerning the probability of which we absolutely know nothing antecedently to any trials concerning it" (Bayes 1764, p. 143), is 1/2. For example, in the analysis of the rising of the sun, it is assumed to be equally likely that the sun will or will not rise.

Apart from ridicule, this position was subjected to a number of telling criticisms, particularly by Boole (1854) and von Kries (1886), and a large number of now-standard paradoxes and counterexamples were adduced (for von Kries, see Kamlah 1983 and 1987). A common response to many of these examples is to point to the requirement of the absence of prior knowledge about the event in question, and argue that it is violated. The fatal flaw in all such defenses is that understanding the very words employed in describing an event necessarily implies some knowledge about that event. Thus, as Keynes notes, in Jevons's notorious example of the proposition "a platythliptic coefficient is positive", the force of the example derives from our entire ignorance of the meaning of the adjective "platythliptic" (Keynes 1921, p. 42, n. 2). Nevertheless, the example is defective, given we do possess considerable knowledge about the words "coefficient" and "positive". Keynes is not being sarcastic, but merely pursuing the argument to its logical conclusion when he asks whether Jevons would "maintain that there is any sense in saying that for those who know no Arabic the probability of every statement expressed in Arabic is even?" (Keynes 1921, p. 43).

Even at the syntactic level, it is easy to construct contradictory probability assignments using the principle of indifference whenever a complex proposition can be decomposed into simpler ones. If Wittgenstein's early program of logical atomism had been successful, then logical probability would be possible, but the failure of the former dooms all attempts to construct the latter. Lacking an ultimate language in one-to-one correspondence with reality, Carnapian programs retain an ultimate element of subjectivism, both in their choice of language and the assumption that a given partition consists of equiprobable elements.

For essentially such reasons, von Kries and others fell back on what was called the principle of *cogent* reason: alternatives are equally probable when we possess knowledge about them, but that knowledge is equally distributed or symmetrical among the alternatives. This was, in fact, the actual Laplacean position: "la probabilité est relative en partie à cette ignorance, en partie à nos connaissances" (Laplace, *Essai*, p. viii). The formulation of the principle of cogent reason, of course, is not without its own problems, and its most satisfactory statements verge on the tautologous. It was, however, a half-way house on the road to the only truly satisfactory formulation: alternatives are equally probable when we judge them to be so. Assignments of equiprobability can only originate as primitives of the system, inputs that are *given*, rather than logical consequences of the syntax of language. Ellis was entirely correct: *ex nihilo nihil*.

7.2. *The urn of nature*

The valid application of the rule of succession presupposes, as Boole notes, the aptness of the analogy between drawing balls from an urn – the *urn of nature*, as it was later called – and observing an event (Boole 1854, p. 369). As Jevons put it, "nature is to us like an infinite ballot-box, the contents of which are being continually drawn, ball after ball, and exhibited to us. Science is but the careful observation of the succession in which balls of various character present themselves . . ." (p. 150).

The origins of the urn of nature are perhaps to be found in James Bernoulli's *Ars conjectandi*. This was a key moment in the history of probability, when the domain of applicability of the theory was dramatically broadened to include physical, biological, and social phenomena far beyond the simple applications to games of chance originally envisaged. But lacking a suitable frequentist or epistemic foundation for probability, Bernoulli was forced to employ the Procrustean bed of equally likely cases: "the numbers of cases in which the same events, with similar circumstances prevailing, are able to happen and not to happen later on". In attempting to apply the doctrine of chances to questions of meteorology, human mortality, and competitive skill, Bernoulli saw the difficulty as one of enumerating these equipossible cases; for example, "the innumerable cases of mutations to which the air is daily exposed", or "the number of diseases". Who, Bernoulli asks, "has well enough examined the nature of the human mind or the amazing structure of our body so that in games which depend wholly or in part on the acumen of the former or the agility of the latter, he could dare to determine the cases in which this player or that can win or lose?" This is the origin of the urn of nature.

What is remarkable about these passages in the *Ars conjectandi* is the almost casual way in which Bernoulli passes from equally likely cases for games of chance to what is essentially a primitive form of propensity theory for physical, biological, and social phenomena. Price, too, began his analysis by first considering "a solid or die of whose number of sides and constitution we know nothing; and that we are to judge of these from experiments made in throwing it", later explaining that he "made these observations chiefly because they are all strictly applicable to the events and appearances of nature".

The aptness of this analogy between tossing a die, or drawing a ball from an urn, is one of the great points in the later debate. Some, like Comte, utterly rejected the application of probability theory outside its original narrow domain, referring contemptuously to Laplace's work as embodying a "philosophical aberration". Others might accept a probabilistic description of sex at birth, or suicide, or weather, but questioned the appropriateness of the analogy in cases such as the rising of the sun, or the movement of the tides.

Thus for enumerative induction the key question became: why and in what way can the relevant observations be viewed as drawings from an urn?

7.3. *W. E. Johnson's rule of succession*

In 1924 the English philosopher and logician William Ernest Johnson published the third and final volume of his *Logic*. In an appendix on "eduction" (i.e., inductive inference from particulars to particulars), Johnson derived a new rule of succession which met both of these basic objections. First, "instead of for two, my theorem holds for α alternatives, primarily postulated as equiprobable" (Johnson 1932, p. 418). Thus the principle of indifference *for alternatives* was exorcised, and the rule extended to cases of multinomial sampling. Although Johnson's form of the rule is sometimes viewed as a straightforward generalization of the original, it will now be appreciated why the generalization was crucial. (Although a proposition and its negation might not be judged equiprobable, the proposition might be one of a spectrum of possibilities which were.)

The mere multinomial generalization, however, had already been discussed by Laplace and De Morgan.[13] But in its derivation Johnson introduced a new and important concept: *exchangeability*. Johnson assumed that "each of the different orders in which a given proportion $m_1 : m_2 : \cdots : m_\alpha$ for M instances may be presented is as likely as any other, what ever may have been the previously known orders". Johnson termed this the "Permutation-Postulate". Its importance is that *it is no longer necessary to refer to the urn of nature*. To what extent is observing instances like drawing balls from an urn? Answer: to the extent that the instances are judged exchangeable. Venn and others had adduced examples where the rule of succession was clearly inappropriate and rightly argued that some additional assumption, other than mere repetition of instances, was necessary for valid inductive inference. From time to time various names for such a principle have been advanced: Mill's Uniformity of Nature; Keynes's Principle of Limited Variety; Goodman's "projectibility". It was Johnson's achievement to have realized both that "the calculus of probability does not enable us to infer any probability-value unless we have some probabilities or probability relations given" (Johnson 1924, p. 182); and that the vague, verbal formulations of his predecessors could be captured in the mathematically precise formulation of exchangeability.[14]

But the rule of succession does not follow from the assumption of exchangeability alone. As we have already seen in Section 3, an assumption must be made about the probability of the different urn-compositions. Johnson called the assumption he employed the *combination postulate*: In a total of M instances, any proportion, say $m_1 : m_2 : \cdots : m_\alpha$, where $m_1 + m_2 + \cdots + m_\alpha = M$, is as likely as any other, prior to any knowledge of the occurrences in question (Johnson 1924, p. 183). This is the multinomial generalization of the Bayes–Laplace assumption that all proportions k/n are equally likely in the binomial case.

Given the permutation and combination postulates, Johnson was able by simple algebra to deduce the multinomial generalization of the rule of succession: $(m_i + 1)/(M + \alpha)$. Because of the setting, infinite trials never came into consideration, and thus this provided a multinomial generalization of the Prevost–L'Huilier/Broad result (although by a clever argument Johnson was able to avoid the problem of explicitly summing the relevant series).

Johnson's result thus coped with two of the three major defects in the rule of succession. If it went largely unappreciated, it was because it was soon superceded by other, more basic and fundamental advances.

7.4. *W. E. Johnson's sufficientness postulate*

The one remaining defect in the rule of succession, as derived by Johnson, was its assumption of the combination postulate. Although Johnson made no appeal to the principle of indifference, the justification for the combination postulate seemed problematical. Johnson himself recognized this, for he soon proposed another, more general postulate, the "sufficientness postulate": the probability of a given type, conditional on n previous outcomes, only depends on how many instances of the type in question occurred, and not on how the other instances distributed themselves amongst the other types (Johnson 1932). Johnson was then able to show that the rule of succession in this case was $(m_i + k)/(M + k\alpha)$, where k can be any positive number. That is, assuming only the sufficientness postulate, a unique answer is no longer determined. This new rule is, of course, none other than Carnap's later "continuum of inductive methods".[15]

7.5. *De Finetti and the rule of succession*

The one final step that remained to be taken was the realization that it was unneccessary to make any further assumption beyond exchangeability. As de Finetti noted in his famous 1937 article, there is a general form of the rule of succession which holds true for an arbitrary finite exchangeable sequence. Namely, if $\omega_r^{(n)}$ denotes the probability of r successes in n trials, then the succession probability given r successes and s failures in $r + s = n$ trials is

$$\frac{r+1}{n+2+(s+1)(\omega_r^{(n+1)}/\omega_{r+1}^{(n+1)}-1)},$$

(de Finetti 1937, p. 144). If $\omega_r^{(n+1)} = \omega_{r+1}^{(n+1)}$, then this reduces to the classical rule of succession. The condition is satisfied exactly if the classical uniformity assumption is made, or approximately in many cases for large n. Venn, in earlier editions of the *Logic of Chance*, had objected to the rule, adducing anti-inductive examples where past successes make future successes less likely rather than more; the de Finetti version of the rule of succession encompasses such situations.

De Finetti's analysis, appearing nearly two centuries after the appearance of Hume's *Treatise* in 1739, represents a watershed in the probabilistic analysis of induction. It abolishes all reference to the infinite, all reference to the principle of indifference, all reference to probabilities of probabilities, all reference to causation, all reference to principles of limited independent variety and other extraneous assumptions. In order to attack it, one must attack the formidable edifice of epistemic probability itself. Modern philosophy continues to ignore it at its own peril.

8. UNIVERSAL GENERALIZATIONS

In 1918 Broad had noted that if there are N balls in an urn, and all n in a random sample are black, then (under the usual equiprobability assumption), the probability that all the balls in the urn are black is $(n+1)/(N+2)$. If n is considerably smaller than N, this probability will also be small; i.e., unless a considerable segment of the sequence $X_1, X_2, \ldots, X_N$ has been observed, or a considerable number of balls drawn from the urn, or most ravens observed, the probability of the universal generalization will be small. This observation has been persistently viewed as an essentially fatal defect of this form of reducing induction to probability, going back at least to the neo-Kantian J. J. Fries in 1842.[16]

The assumption on which the calculation is based, that $P[S_N = k] = (N+1)^{-1}$, $0 \leq k \leq N$, is a *symmetry* assumption and, like many symmetry assumptions, contains "hidden baggage" not always initially apparent. In this case the hidden baggage lies surprisingly close to the surface; adopting the uniformity assumption $P[S_N = k] = 1/(N+1)$ means assuming in particular that

$$P[S_N = 0] = P[S_N = N] = \frac{1}{(N+1)},$$

i.e., that *the universal generalizations* $\{S_N = 0\}$ *and* $\{S_N = N\}$ *are a priori highly improbable* (if N is large). It is hardly surprising that hypotheses initially thought highly improbable should remain improbable unless a considerable fraction of the sequence has been observed.

If one deals with an infinitely exchangeable sequence, the problem becomes even more acute: taking the limit as $N \to \infty$ shows that for the Bayes–Laplace process the conditional probability of a universal generalization given n successes is always zero:

$$P[X_{n+1} = 1, X_{n+2} = 1, X_{n+3} = 1, \ldots, \mid S_n = n] = 0.$$

Once again, this is only surprising if one fails to understand what is being assumed. In the Bayes–Laplace process, the prior probability is zero that the unknown chance p equals one (i.e., $P[p = 1] = 0$), and thus the probability that the limiting frequency of 1s is one must necessarily also be zero.

There are two ways out of this dilemma for those who wish to conclude inductively that repeatedly confirmed universal hypotheses are a posteriori probable:

(1) *Stonewalling*. That is, argue that the initial intuition was in fact false; no matter how many successes have been observed does not warrant expecting arbitrarily long further strings of success. This is a very old defense, and it appears both in Price's appendix to Bayes's essay (p. 151), and Laplace's first paper on inverse probability (Laplace 1774).

De Morgan gives a good statement of the position:

> [E]xperience can never, on sound principles, be held as foretelling all that is to come. The order of things, the laws of nature, and all those phrases by which we try to make the past command the future, will be understood by a person who admits the principles of which I treat as of limited application, not giving the highest degree of probability to more than a definite and limited continuance of those things which appear to us most stable. *No finite experience whatsoever can justify us in saying that the future shall coincide with the past in all time to come, or that there is any probability for such a conclusion.* (De Morgan, 1838, p. 128; emphasis De Morgan's)

Obviously such a position is open to the objection that some inductive inferences *are* of universal character, and it has been subjected to varying forms of ridicule. Keynes (1921, p. 383) notes that the German philosopher Bobeck calculated that the probability of the sun's rising every day for the next 4,000 years, given Laplace's datum that it has risen for the last 5,000 years or 1,826,213 days, is no more than 2/3. (Actually, using the formula above, the probability may be readily calculated to be $(1{,}826{,}213 + 1)/(3{,}287{,}183 + 1) = 0.56$.)

(2) *Different priors*. Edgeworth, in his review of Keynes's *Treatise*, remarks that "pure induction avails not without some finite initial probability in favour of the generalisation, obtained from some other source than the instances examined" (Edgeworth 1922, p. 267). This is the nub of the matter: in order for the posterior probability of any event to be positive, its prior probability must be positive (cf. Keynes 1921, p. 238). Within a year of Broad's 1918 paper, Jeffreys and Wrinch (1919) noted that the difficulty could be averted by using priors which place point masses at the endpoints of the unit interval.[17]

In 1919 this may have seemed to some to beg the question; after the ascendancy of the Ramsey/de Finetti approach to epistemic probability, it seems quite natural. Permitting passage from a unique uniform prior to a much wider class was crucial if this objection was to be successfully met.

What is the justification for assigning positive probability to the end points? The argument is in fact quite simple: *not* assuming some positive prior probability for the universal generalization is not an assumption of neutrality or absence of knowledge; it means that the universal generalization is assumed to have probability *zero*, i.e., we are *certain* it will not occur. Thus this classical objection is in fact a non-objection: unless one is *certain* that the universal generalization is false, its posterior probability increases with the number of confirming instances.

9. BRUNO DE FINETTI AND THE RIDDLE OF INDUCTION

Despite its virtues, the Jeffreys–Wrinch formulation suffers both from its continuing appeal to the "unknown chance" p, and the apparently ad hoc nature of the priors it suggests. Both of these defects were removed in 1931, when Bruno de Finetti proved his justly famous infinite representation theorem.[18] De Finetti showed that if an infinite sequence of 0–1 valued random variables $X_1, X_2, \ldots, X_n, \ldots$ is exchangeable for every n, then the limiting frequency of ones, $Z = \lim(X_1 + X_2 + \cdots X_n)/n$, will exist with probability 1, and the probability distribution of the sequence can be represented as a mixture of binomial probabilities having this (random) limiting frequency Z as success parameter. Thus de Finetti provided a subjectivist account of objectivist chance and the role of parameters in statistics.

De Finetti also employed his infinite representation theorem, moreover, to provide a qualitative explanation for induction similar to and directly inspired by the French mathematician Henri Poincaré's *method of arbitrary functions* (Poincaré 1902, chap. 11). Poincaré had also sought to explain the existence of objective chance: to account for the apparently paradoxical fact that the outcome of tossing a coin, rolling a die, or spinning a roulette wheel results in equiprobable outcomes, despite our ignorance of and inability to control the complex physical conditions under which these occur. Poincaré was able to show that essentially independent of the distribution of physical inputs – in tossing a die, for example, the imparted linear velocity and angular momentum – the outcomes will occur with *approximately* equal probabilities for reasons that are purely mathematical. Detailed quantitative knowledge of input is thus unnecessary for approximate knowledge of outcome. Likewise, de Finetti was able to show for exchangeable sequences that, essentially independent of the initial distribution for the limiting frequency Z, after observing a sufficiently long initial segment of the sequence the posterior distribution of Z will always be highly peaked about the observed frequency p^*, and future trials expected to occur with a frequency very close to that of p^*.

De Finetti's is thus a *coherence* explanation of induction: if past and future are judged exchangeable, then – if we are to be consistent – we must expect future frequencies to resemble those of the past. But unlike the unique quantitative answer of Bayes, or the continuum of quantitative answers provided by Jeffreys, Wrinch, and Johnson, de Finetti's solution to Hume's problem of induction is a qualitative one. Whatever our prior beliefs, our inferences about the future based on our knowledge of the past will indeed be inductive in nature, but in ways that do not admit of unique numerical expression and may vary from person to person.

De Finetti's solution of Hume's problem of induction is a profound achievement, one of the few occasions when one of the deep problems of philosophy has yielded to direct attack. De Finetti's solution *can* be criticized, but such criticisms must go either to the nature of probable belief itself (can it be quantified? how does it change when new information is received?), or the illposed nature of Hume's problem (*how* is the future supposed to resemble the past?).

What is remarkable about de Finetti's essay *Probabilismo* is the clarity with which de Finetti saw these issues from the very beginning, and how closely they fit there into a unified view of science and philosophy. To many it may seem a strange and unfamiliar landscape; perhaps it "will only be understood by someone who has himself already had the thoughts that are expressed in it".[19]

APPENDIX: DERIVATION OF THE FINITE RULE OF SUCCESSION

Although the derivation of the finite rule of succession via exchangeability is both simple and elegant, the direct combinatorial proofs of the rule have a mathematical attractiveness of their own. We begin with the simplest case, considered by Broad, when all outcomes observed are successes.

Fix $m \geqslant 1$, and for $n \geqslant m$, let $a_n =: n!/(n-m)!$ and $S_n =: a_m + a_{m+1} + \cdots + a_n$.

LEMMA 1. For all $n \geq m$, $S_n = (n+1)!/\{(m+1)(n-m)!\}$.

Proof. By induction. For $n = m$, $S_m = a_m = m! = (m+1)!/\{(m+1)(m-m)!\}$. Suppose next that the lemma is true for some $n \geq m$, and let

$$A_n = \frac{(n+1)!}{(m+1)(n-m)!}.$$

The induction hypothesis then states that $S_n = A_n$, and we wish to prove that $S_{n+1} = A_{n+1}$. But

$$\begin{aligned}
S_{n+1} &= S_n + a_{n+1} \\
&= A_n + a_{n+1} \\
&= \frac{(n+1)!}{(m+1)(n-m)!} + \frac{(n+1)!}{(n-m+1)!} \\
&= (n+1)! \frac{(n-m+1)+(m+1)}{(m+1)(n-m+1)!} \\
&= (n+1)! \frac{(n+2)}{(m+1)(n-m+1)!} \\
&= \frac{(n+2)!}{(m+1)(n-m+1)!} \\
&= A_{n+1}. \qquad \square
\end{aligned}$$

The integer m was fixed in the above argument. If we now denote the dependence of S_n on m by $S_{n,m}$, then Broad's result follows by noting that the succession probability is the quotient $(n-m)^{-1}S_{n,m+1}/S_{n,m}$ and cancelling terms. Thus

$$\begin{aligned}
P[X_{m+1} = 1 \mid X_1 = X_2 = \cdots = X_m = 1] &= \frac{\sum_{j=m+1}^{n} j(j-1)(j-2)\cdots(j-m)}{(n-m)\sum_{j=m}^{n} j(j-1)(j-2)\cdots(j-m+1)} \\
&= \frac{1}{(n-m)} \frac{S_{n,m+1}}{S_{n,m}}
\end{aligned}$$

$$= \frac{1}{(n-m)} \frac{(n+1)!/\{(m+2)(n-m-1)!\}}{(n+1)!/\{(m+1)(n-m)!\}}$$

$$= \frac{(m+1)}{(m+2)}.$$

If instead of all m trials being successes, it is assumed that there are p successes and q failures, then the necessary sum becomes

$$\sum_j \frac{(n-q-j)!}{(n-m-j)!} \frac{(q+j)!}{j!},$$

and it may be again evaluated to be

$$\frac{(p!q!)}{(m+1)!} \frac{(n+1)!}{(n-m)!}.$$

Namely, if there are p successes and q failures, the possible urn compositions are

$\mathbf{H}_j : (q+j)$ black and $(n-q-j)$ white,
$0 \leqslant j \leqslant n-m.$

Under hypothesis $\mathbf{H}_j$ the probability that p whites and q blacks will be observed in the first m trials is

$$\mathbf{P}_j = \binom{m}{q} \frac{P_j Q_j}{(n)(n-1)\cdots(n-m+1)},$$

where

$$P_j = (n-q-j)(n-q-j-1)\cdots(n-q-j-p+1)$$

$$= \frac{(n-q-j)!}{(n-m-j)!}$$

and

$$Q_j = (q+j)(q+j-1)(q+j-2)\cdots(j+1)$$

$$= \frac{(q+j)!}{j!}.$$

By Bayes's theorem, it follows that the posterior probability of $\mathbf{H}_i$, given that p whites and q blacks have been observed, is

$$\begin{aligned} P[\mathbf{H}_i \mid p, q] &= \mathbf{P}_i(n+1)^{-1} \Big/ \sum_j \mathbf{P}_j(n+1)^{-1} \\ &= \mathbf{P}_i \Big/ \sum_j \mathbf{P}_j \\ &= P_i Q_i \Big/ \sum_j P_j Q_j, \end{aligned}$$

where

$$P_i Q_i = \frac{(n-q-i)!}{(n-m-i)!} \frac{(q+i)!}{i!},$$

and

$$\sum_j P_j Q_j = \sum_j \frac{(n-q-j)!}{(n-m-j)!} \frac{(q+j)!}{j!}.$$

LEMMA 2.

$$\sum_{j=0}^{n-m} \frac{(n-q-j)!}{(n-m-j)!} \frac{(q+j)!}{j!} = p!q! \frac{(n+1)!}{(m+1)!(n-m)!}.$$

Proof. Dividing by $p!q!$ and denoting $n-m$ by k yields

$$\sum_{j=0}^{k} \binom{k+p-j}{k-j} \binom{q+j}{j} = \binom{k+p+q+1}{k},$$

which is a well-known combinatorial identity (see, e.g., Whitworth 1897, p. xiii; Feller 1968, p. 65, problem 14 and p. 238, problem 15). □

All of the classical derivations of the finite rule of succession reduce to the evaluation of this sum.

A.1. *Prevost and L'Huillier's proof*

The Prevost–L'Huillier proof draws on the machinery of the "figurate numbers."

Let

$$f_0^j = 1, \quad j = 1, 2, 3, \ldots$$

and

$$f_i^j = \sum_{k=1}^{j} f_{i-1}^k, \quad i, j \geq 1;$$

the f_i^j are called the *figurate numbers* and are related to the binomial coefficients by the formula $f_i^j = \binom{j-i+1}{i}$; see generally Edwards (1987).

Prevost and L'Huilier prove the figurate identity

$$f_{p+q+1}^n = \sum_{i=1}^{n} f_q^i f_p^{n-i+1}$$

by induction on q, from which the corresponding result for binomial coefficients immediately follows. (Edwards 1987, p. 107, attributes this result in the special case $q = 1$ to Leibniz in 1673 but does not cite a source for the general result, his equation (8.24). Edwards's k, l, and s are our $p + q + 1$, n, and q respectively.)

Proof: If $q = 0$, then the identity is valid for all $p \geq 0$, since $f_0^i = 1$, all i, and the identity follows from the definition of the figurate numbers. Suppose the identity has been proved for some $q \geq 0$ and all $p \geq 0$. Then it holds for $q + 1$ and all $p \geq 0$ because

$$\begin{aligned} f_{p+q+2}^n &= \sum_{j=1}^{n} f_q^j \, f_{p+1}^{n-j+1} && \text{(induction step)} \\ &= \sum_{j=1}^{n} f_q^j \left\{ \sum_{i=j}^{n} f_p^{n-i+1} \right\} && \text{(definition of } f_i^j\text{)} \\ &= \sum_{i=1}^{n} \left\{ \sum_{j=1}^{i} f_q^j \right\} f_p^{n-i+1} && \text{(rearrangement)} \\ &= \sum_{i=1}^{n} f_{q+1}^i f_p^{n-i+1} && \text{(definition of } f_i^j\text{);} \end{aligned}$$

the first step uses the induction hypothesis for q and $p + 1$.

Alternatively,

$$\sum_{i=1}^{n} f_{q+1}^i f_p^{n-i+1} = \sum_{i=1}^{n} \left\{ \sum_{j=1}^{i} f_q^j \right\} f_p^{n-i+1} \qquad \text{(definition of } f_i^j\text{)}$$

$$= \sum_{i=1}^{n} \left\{ \sum_{j=1}^{i} f_q^{i-j+1} \right\} f_p^{n-i+1} \qquad \text{(reverse summation)}$$

$$= \sum_{j=1}^{n} \sum_{i=j}^{n} f_q^{i-j+1} f_p^{n-i+1} \qquad \text{(rearrangement)}$$

$$= \sum_{j=1}^{n} \sum_{k=1}^{n-j+1} f_q^k f_p^{n-j-k+2} \qquad (k = i - j + 1)$$

$$= \sum_{j=1}^{n} f_{p+q+1}^{n-j+1} \qquad \text{(induction)}$$

$$= f_{p+q+2}^{n} \qquad \square$$

This second proof is in fact the one given by Prevost and L'Huilier (although they only prove the specific cases $q = 0, 1, 2$, and 3, "la marche de la démonstration générale étant entièrement semblable à celle des exemples précédents, et ne présentant aucune difficulté" (p. 122)).

The fundamental binomial identity is also a special case of the more general result

$$\sum_{k=0}^{n} A_k(a, b) A_{n-k}(c, b) = A_n(a + c, b),$$

where $A_k(a, b) = (a/(a + bk))\binom{a+bk}{k}$ and a, b, c are arbitrary real numbers. As Gould and Kaucky (1966, p. 234) note, this identity "has been widely used, rediscovered repeatedly, and generalized extensively". Gould and Kaucky attribute it to H. A. Rothe, who proved it in his Leipzig dissertation of 1793, and thus appears to have priority over Prevost and L'Huilier.

A.2. *Bishop Terrot's Proof*

Bishop Terrot sums the series in Lemma 2 by using Abel partial summation (see, e.g., Knopp 1947, p. 313) and the identity

$$\sum_{j=0}^{a} \frac{(b+j)!}{j!} = \frac{(a+b+1)(a+b)\cdots(a+1)}{(b+1)}.$$

If one denotes the sum in Lemma 2 by $A_{n,p,q}$, then it follows that

$$A_{n,p,q} = \left\{ \frac{q}{(p+1)} \right\} A_{n,p+1,q-1},$$

and repeating the process a total of q times yields Lemma 2.

A.3. *Todhunter's proof*

Todhunter remarks that the sum is readily obtained "by the aid of the binomial theorem" (Todhunter 1865, p. 454). By this he means comparing the coefficients in $(x+y)^{a+b} = (x+y)^a(x+y)^b$, with negative exponents permitted; see Feller (1968, p. 65, problem 14).

A.4. *Ostrogradskii*

The Russian mathematician M. V. Ostrogradskii also appears to have analyzed this problem in 1846; see Maistrov (1974, pp. 182–84).

A.5. *Whipple's Proof*

Jeffreys (1973, Appendix II) reports a simple counting argument proof. Combinatorial identities can often be profitably thought of as counting a certain quantity in two different ways, and Whipple discovered such an interpretation for Lemma 2.

NOTES

[1] Laplace (1774; 1812, p. 402; 1814, p. xvii). The terminology is due to Venn, who gave this name to Chapter VIII of his *Logic of Chance*, adding in a footnote: "A word of apology may be offered here for the introduction of a new name. The only other alternative would have been to entitle the rule one of Induction. But such a title I cannot admit, for reasons which will be almost immediately explained" (Venn 1888, p. 190).

[2] Broad later abandoned this assumption; see Broad (1927).

[3] For further information about Prevost and L'Huilier, see the entries on both in the *Dictionary of Scientific Biography*.

[4] The reader should be cautioned the literature abounds with confusions between the finite and infinite cases, and the cases involving sampling with and without replacement. For example, Keynes (1921, p. 378) states that "the rule of succession does not apply, as it is easy to demonstrate, even to the case of balls drawn from an urn, if the number of balls is finite" (Keynes 1921, p. 378). Likewise Strong (1976, p. 203) confuses the answers for sampling with and without replacement from a finite urn, a confusion that may stem in part from an unfortunate typographical error in Boole.

[5] Stigler's nomination of Nicholas Saunderson, Lucasian Professor of Mathematics at Cambridge from 1711 to 1739, as a plausible alternative candidate (Stigler 1983), cannot be seriously credited. There is no evidence that Saunderson ever wrote on any topic in probability. Hartley's wording, moreover ("an ingenious friend has communicated to me..."), suggests information recently received from a living person, rather than a friend long dead (Saunderson had died a decade earlier, in 1739). The anonymity employed would scarcely make sense otherwise. Stigler's concluding "Bayesian calculation" (summarized in his Table 1), purporting to show that the odds are 3 to 1 in favor of Saunderson over Bayes, curiously omits the single most important item of evidence in favor of Bayes – that Bayes is known to have written a manuscript on the subject, while Saunderson is not.

[6] In his original memoir on inverse probability, where the rule of succession is first stated, Laplace begins by considering an urn "supposed to contain a given number of white and black tickets in an unknown ratio", but in his solution he assumes that the urn contains "an infinity of white and black tickets in an unknown ratio". Strictly construed, of course, this latter statement makes no sense, and is clearly intended as an abbreviated way of saying something else. De Morgan considers that the contents of the urn are assumed infinite "only that the withdrawal of a definite number may not alter the ratio" (1847, p. 214), and he goes on to note that if the contents of the urn are finite, but the sampling is *with* replacement, then as the number of balls increases, the resulting rule of succession will approximate the Laplacean answer of $(m+1)/(m+2)$. Bishop Terrot, on the other hand, uses the expression "infinite or indefinite" in describing this case (Terrot 1853, p. 543), and clearly takes the reference to infinite contents to be a circumlocution for the asymptotic, large-sample result. The paper by Prevost and L'Huilier was an attempt to clarify matters by determining the exact, small-sample answer, and it must have come as a considerable surprise to find that there was no difference between the two. The philosophical importance of the result is that, whatever its other real or alleged defects, Laplace's analysis cannot be faulted on the grounds of its appeal to the infinite. This point is sometimes overlooked.

[7] "... it was Hume who furnished the Laplacean school with its philosophy of science," as Stove (1973, p. 102) notes. Hume's influence, especially on Condorcet, is ably discussed by Baker (1975, chap. 3, passim), and appears most clearly in the work of Condorcet's disciple, Sylvestre-Francois Lacroix (Stove 1973, p. 103; Baker 1975, pp. 186–87).

[8] Strictly speaking, Price misstates the rule in two ways: (a) the odds are $2^n - 1$ to 1, not 2^n to 1; (b) the exponent of 2 should be the number of risings, not the number of returns. (Thus the true odds are $2^{1,000,001} - 1$ to 1.) Price gives the correct formula, however, earlier in his appendix.

[9] The French naturalist Buffon gives the 2^n to 1 rule in his *Essai d'arithmétique morale* of 1777; taking the age of the earth to be 6,000 years, he concludes the probability that the sun will rise the following day is 22,189,999 to 1. Although Buffon does not cite the rule's source, it is clearly taken from Bayes's essay: Price's fanciful narrative of a man who observes the rising of the sun for the first time has been lifted, without attribution, virtually word-for-word! (Zabell 1988).

[10] Likewise, the English mathematician Waring (1794, p. 35) dissented from the identification:

I know that some mathematicians of the first class have endeavoured to demonstrate the degree of probability of an event's happening *n* times from its having happened *m* preceding times; and consequently that such an event will probably take place; but, alas, the problem far exceeds the extent of human understanding; who can determine the time when the sun will probably cease to run its present course?

[11] The Truscott and Emory translation of the *Essai* renders "principe régulateur" as "principal regulator". This is not only incorrect, it introduces a deistic note entirely foreign to Laplace, and obscures the essential point of the passage.

[12] The level of English mathematics was at ebbtide at the beginning of the 19th century, and De Morgan was one of a group of English mathematicians and scientists (including Babbage, Herschel, and Galloway) who attempted to remedy the situation during the first half of the century through a series of popular and technical tracts.

[13] The mathematical machinery for the generalization is provided by Laplace (*Théorie*, p. 376), although it is not stressed. The uniform prior on the unit interval is replaced by the uniform prior on the simplex $\Delta_t =: \{(p_1, p_2, \ldots, p_t): p_1, p_2, \ldots, p_t \geq 0, p_1 + p_2 + \cdots + p_t = 1\}$, and the rule of succession becomes $(m_i + 1)/(M + t)$, where m_i is the number of outcomes in the ith category. (This is, of course, nothing other than Carnap's c^*.) De Morgan discusses the rule of succession in this more general context (De Morgan 1838, pp. 66–69; 1845, pp. 413–14), including the so-called "sampling of species problem", where one does not know the total number of categories.

[14] Now the inappropriateness of the application of the rule of succession to the rising of the sun becomes manifest: successive risings are not exchangeable. (For example, although for most people the probability that the sun will rise tomorrow, but not rise the day after is quite small, the probability that it will not rise tomorrow but will rise the day after is much smaller still.)

[15] See generally Zabell (1982).

[16] For discussion of Fries's analysis, see Krüger (1987, pp. 67–68).

[17] The objection that simply because the limiting frequency of 1s in the sequence can equal 1, it does not follow that all elements of the sequence will equal 1, is easily met by working at the level of sequence space and giving the sequences $\{1, 1, 1, \ldots\}$ and $\{0, 0, 0, \ldots\}$ positive probability.

[18] The following briefly summarizes an argument given in much greater detail in Zabell (1988).

[19] I thank Persi Diaconis for a number of helpful discussions over the years regarding finite exchangeability (in particular, for pointing out the identity of the Prevost–L'Huilier and the Bayes–Laplace processes).

REFERENCES

Baker, Keith Michael: 1975, *Condorcet: From Natural Philosophy to Social Mathematics*, University of Chicago Press, Chicago.

Bayes, Thomas: 1764, 'An Essay Towards Solving a Problem in the Doctrine of Chances', *Philosophical Transactions of the Royal Society of London* **53**, 370–418, reprinted in E. S. Pearson and M. G. Kendall (eds.), *Studies in the History of Statistics* *and Probability*, Vol. 1, Charles Griffin, London, 1970, pp. 134–53, page references are to this edition.

Bernoulli, Jakob: 1713, *Ars conjectandi*, Thurnisiorum, Basel, reprinted in *Die Werke von Jakob Bernoulli*, Vol. 3, Birkhauser, Basel. 1975, pp. 107–286.

Bertrand, J.: 1907, *Calcul des probabilités*, Gauthier-Villars, Paris, (1st ed., 1889).

Bolzano, Bernard: 1837, *Wissenschaftslehre*, translated 1972 under the title *Theory of Science*, R. George (ed. and trans.), University of California Press, Berkeley.

Boole, George: 1854, *An Investigation of the Laws of Thought*, Macmillan, London, reprinted 1958, Dover, New York.

Broad, C. D.: 1918, 'The Relation Between Induction and Probability', *Mind* **27**, 389–404; **29**, 11–45.

Broad, C. D.: 1922, 'Critical Notice on J. M. Keynes', *A Treatise on Probability, Mind* **31**, 72–85.

Broad, C. D.: 1924, 'Mr. Johnson on the Logical Foundations of Science', *Mind* **33**, 242–61, and 369–84.

Broad, C. D.: 1927, 'The Principles of Problematic Induction', *Proceedings of the Aristotelian Society* **28**, 1–46.

Carnap, Rudolph: 1950, *Logical Foundations of Probability*, 2nd ed. 1962, The University of Chicago Press, Chicago.

Carnap, Rudolph: 1952, *The Continuum of Inductive Methods*, The University of Chicago Press, Chicago.

Condorcet, Le Marquis de: 1785, *Essai sur l'application de l'analyse à la probabilité dés décisions rendues à la pluralité des voix*, Imprimerie Royale, Paris.

Cournot, Antoine Augustin: 1843, *Exposition de la théorie des chances et des probabilités*, Libraire de L. Hachette, Paris.

Dale, A. I.: 1986, 'A Newly-Discovered Result of Thomas Bayes', *Archive for History of Exact Sciences* **35**, 101–13.

De Finetti, Bruno: 1937, 'La prevision: ses lois logiques, ses sources subjectives', *Annales de l'Institut Henri Poincaré* **7**, 1–68, translated in H. E. Kyburg, Jr. and H. E. Smokler, (eds.), *Studies in Subjective Probability*, Wiley, New York, 1964, pp. 93–158, page references are to this edition.

De Morgan, Augustus: 1838, *An Essay on Probabilities, and their Application to Life Contingencies and Insurance Offices*, Longman, Orme, Brown, Green, and Longmans, London.

De Morgan, Augustus: 1845, 'Theory of Probabilities', *Encyclopedia Metropolitana, Vol. 2: Pure Mathematics*, pp. 393–490, B. Fellowes et al., London.

De Morgan, Augustus: 1847, *Formal Logic: Or the Calculus of Inference Necessary and Probable*, Taylor and Watton, London.

Diaconis, Persi: 1977, 'Finite Forms of de Finetti's Theorem on Exchangeability', *Synthese* **36**, 271–81.

Edgeworth, Francis Ysidro: 1884, '*A Priori* Probabilities', *Philosophical Magazine* (Series 5) **18**, 205–10.

Edgeworth, Francis Ysidro: 1922, 'The Philosophy of Chance', *Mind* **31**, 257–83.

Edwards, A. W. F.: 1978, 'Commentary on the Arguments of Thomas Bayes', *Scandinavian Journal of Statistics* **5**, 116–18.

Edwards, A. W. F.: 1987, *Pascal's Arithmetic Triangle*, Oxford University Press, New York.

Feller, William: 1968, *An Introduction to Probability Theory and its Applications*, Vol. 1, 3rd ed., Wiley, New York.

Fisher, Ronald A.: 1973, *Statistical Methods and Scientific Inference*, 3rd ed., (1st ed., 1956; 2nd ed., 1959), Hafner Press, New York.

Good, Irving John: 1950, *Probability and the Weighing of Evidence*, Hafner Press, New York.

Good, Irving John: 1965, *The Estimation of Probabilities: An Essay on Modern Bayesian Methods*, MIT Press, Cambridge, Massachusetts.

Goodman, Nelson: 1979, *Fact, Fiction, and Forecast*, 3rd ed., Hackett, Indianapolis.

Gould, H. W. and J. Kaucky: 1966, 'Evaluation of a Class of Binomial Coefficient Summations', *Journal of Combinatorial Theory* **1**, 233–47.

Hacking, Ian: 1975, *The Emergence of Probability*, Cambridge University Press, New York.

Hartley, David: 1749, *Observations on Man, his Frame, his Duty, and his Expectations*, S. Richardson, London.

Heath, P. L.: 1967, 'Jevons, William Stanley', in P. Edwards (ed.), *The Encyclopedia of Philosophy*, Vol. 4, Macmillan, New York, pp. 260–61.

Hume, David: 1739, *A Treatise of Human Nature*, London. Page references are to the 2nd edition of the L. A. Selbe-Bigge text, revised by P. H. Nidditch, Clarendon Press, Oxford, 1978.

Jeffreys, Harold: 1939, *The Theory of Probability*, Clarendon Press, Oxford, (2nd ed. 1958; 3rd ed. 1961).

Jeffreys, Harold: 1973, *Scientific Inference*, 3rd ed., Cambridge University Press, New York.

Jeffreys, Harold and Dorothy Wrinch: 1919, 'On Certain Aspects of the Theory of Probability', *Philophical Magazine* **38**, 715–31.

Jevons, William Stanley: 1874, *The Principles of Science: A Treatise on Logic and Scientific Method*, 2 vols., Macmillan, London, (2nd ed., 1877), reprinted 1958, Dover, New York.

Johnson, William Ernest: 1924, *Logic, Part III: The Logical Foundations of Science*, Cambridge University Press, Cambridge.

Johnson, William Ernest: 1932, 'Probability: The Deductive and Inductive Problems', *Mind* **49**, 409–23.

Kamlah, Andreas: 1983, 'Probability as a Quasi-Theoretical Concept: J. v. Kries' Sophisticated Account After a Century', *Erkenntnis* **19**, 239–51.

Kamlah, Andreas: 1987, 'The Decline of the Laplacian Theory of Probability: A Study of Stumpf, von Kries, and Meinong', in L. Kruger, L. J. Daston, and M. Heidelberger (eds.), *The Probabilistic Revolution, Volume 1: Ideas in History*, MIT Press, Cambridge, Massachusetts, pp. 91–110.

Keynes, John Maynard: 1921, *A Treatise on Probability*, Macmillan, London.

Kneale, William: 1949, *Probability and Induction*, The Clarendon Press, Oxford.

Knopp, Konrad: 1947, *Theory and Application of Infinite Series*, Hafner Press, New York.

Kries, Johann von: 1886, *Die Principien der Wahrscheinlichkeitsrechnung*, Mohr, Tübingen, (2nd ed., 1927).

Krüger, Lorenz: 1987, 'The Slow Rise of Probabilism', in L. Krüger, L. J. Daston, and M. Heidelberger (eds.), *The Probabilistic Revolution, Volume 1: Ideas in History*, MIT Press, Cambridge, Massachusetts, pp. 59–89.

Laplace, Pierre Simon Marquis de: 1774, 'Mémoire sur la probabilité des causes par les évenements', *Mémoires de l'Académie royale des sciences presentés par divers savans* **6**, 621–56, reprinted in *Oeuvres complètes de Laplace* (1878–1912), Vol. 8, pp. 27–65, Gauthier-Villars, Paris, translated in Stigler (1986).

Laplace, Pierre Simon Marquis de: 1812, *Théorie analytique des probabilités*, Courcier, Paris, 2nd ed., 1814; 3rd ed., 1820, page references are to *Oeuvres complètes de Laplace*, Vol. 7, Gauthier-Villars, Paris, 1886.

Laplace, Pierre Simon Marquis de: 1814, *Essai philosophique sur les probabilités*, Courcier, Paris, page references are to *Oeuvres complètes de Laplace*, vol. 7, Gauthier-Villars, Paris, 1886.

Laudan, L.: 1973, 'Induction and Probability in the Nineteenth Century', in P. Suppes et al. (eds.), *Logic, Methodology and Philosophy of Science IV*, North-Holland, Amsterdam, pp. 429–38.

Madden, Edward H.: 1960, 'W. S. Jevons on Induction and Probability', in E. H. Madden (ed.), *Theories of Scientific Method: The Renaissance Through the Nineteenth Century*, University of Washington Press, Seattle, pp. 233–47.

Maistrov, L. E.: 1974, *Probability Theory: A Historical Sketch*, Academic Press, New York.

Mill, John Stuart: 1843, *A System of Logic*, 2 Vols., John W. Parker, London.

Ostrogradskii, M. V.: 1848, 'On a Problem Concerning Probabilities', *St. Petersburg Academic of Sciences* **6**, 321–46, reprinted in M. V. Ostrogradskii, *Polnoe sobranie trudov*, Vol. 3, Academy of Sciences URRSSR, Kiev, 1961.

Pearson, Karl: 1978, *The History of Statistics in the 17th and 18th Centuries*, E. S. Pearson (ed.), Macmillan, New York.

Poincaré, Henri: 1902, *La science et l'hypothèse*, Paris, translated 1905 as *Science and Hypothesis*, J. W. Greenstreet (trans.), Walter Scott, London; translation reprinted 1952, Dover, New York.

Prevost, Pierre and S. A. L'Huilier: 1799, 'Sur les probabilités', *Mémoires de l'Academie Royale de Berlin* 1796, pp. 117–42.

Prevost, Pierre and S. A. L'Huilier: 1799a, 'Mémoire sur l'art d'estimer la probabilité des causes par les effets', *Mémoires de l'Academic Royale de Berlin* 1796, pp. 3–24.

Price, Richard: 1758, *A Review of the Principal Questions in Morals*, London (2nd ed., 1769; 3rd ed., 1787), reprinted 1974, D. D. Raphael (ed.), Clarendon Press, Oxford, page references are to the 1974 edition.

Ramsey, Frank Plumpton: 1926, 'Truth and Probability', in R. B. Braithwaite (ed.), *The Foundations of Mathematics and other Logical Essays*, Routledge and Kegan, Paul, London, 1931, pp. 156–98.

Stigler, Stephen M.: 1982, 'Thomas Bayes's Bayesian Inference', *Journal of the Royal Statistical Society Series A* **145**, 250–58.

Stigler, Stephen M.: 1983, 'Who Discovered Bayes's Theorem?', *American Statistician* **37**, 290–96.

Stigler, Stephen M.: 1986, 'Laplace's 1774 Memoir on Inverse Probability', *Statistical Science* **1**, 359–78.
Stove, D. C.: 1973, *Probability and Hume's Inductive Scepticism*, Clarendon Press, Oxford.
Strong, John V.: 1976, 'The Infinite Ballot Box of Nature: De Morgan, Boole, and Jevons on Probability and the Logic of Induction', *PSA 1976: Proceedings of the Philosophy of Science Association* **1**, 197–211.
Strong, John V.: 1978, 'John Stuart Mill, John Herschel, and the 'Probability of Causes', *PSA 1978: Proceedings of the Philosophy of Science Association* **1**, 31–41.
Terrot, Bishop Charles: 1853, 'Summation of a Compound Series, and its Application to a Problem in Probabilities', *Transactions of the Edinburgh Philosophical Society* **20**, 541–45.
Todhunter, Isaac: 1865, *A History of the Mathematical Theory of Probability from the Time of Pascal to that of Laplace*, Macmillan, London, reprinted 1965, Chelsea, New York.
Venn, John: 1866, *The Logic of Chance*, Macmillan, London, (2nd ed., 1876; 3rd ed., 1888), reprinted 1962, Chelsea, New York.
von Wright, Georg Henrik: 1957, *The Logical Problem of Induction*, 2nd revised edition, Macmillan, New York.
Waring, E.: 1794, *An Essay on the Principles of Human Knowledge*, Deighton Bell, Cambridge.
Whitworth, William Allen: 1901, *Choice and Chance*, 5th ed., Deighton Bell, Cambridge.
Zabell, Sandy L.: 1982, 'W. E. Johnson's "Sufficientness" Postulate', *Annals of Statistics* **10**, 1091–99.
Zabell, Sandy L.: 1988, 'Symmetry and its Discontents', in Harper, W. and Skyrms, B. (eds.), *Probability, Chance, and Causation*, Vol. I, Kluwer, Dordrecht, pp. 155–190.
Zabell, Sandy L.: 1988a, 'Buffon, Price, and Laplace: Scientific Attribution in the 18th Century', *Archive for History of Exact Sciences* **39**, 173–181.
Zabell, Sandy L.: 1989, 'R. A. Fisher on the History of Inverse Probability', *Statistical Science*, to appear.

Manuscript submitted May 1988

Departments of Mathematics and Statistics
Northwestern University
2006 Sheridan Road
Evanston, IL 60208
U.S.A.