Wittgenstein on Prior Probabilities

Michael Cuffaro

Abstract

Wittgenstein did not write very much on the topic of probability. The little we have comes from a few short pages of the *Tractatus*, some 'remarks' from the 1930's, and the informal conversations which went on during that decade with the Vienna Circle. Nevertheless, Wittgenstein's views were highly influential in the later development of the logical theory of probability. This paper will attempt to clarify and defend Wittgenstein's conception of probability against some oft-cited criticisms that stem from a misunderstanding of his views. Max Black, for instance, criticises Wittgenstein for formulating a theory of probability that is capable of being used only against the backdrop of the ideal language of the *Tractatus*. I argue that on the contrary, by appealing to the 'hypothetical laws of nature', Wittgenstein is able to make sense of probability statements involving propositions that have not been completely analysed. G.H. von Wright criticises Wittgenstein's characterisation of these very hypothetical laws. He argues that by introducing them Wittgenstein makes what is distinctive about his theory superfluous, for the hypothetical laws are directly inspired by statistical observations and hence these observations indirectly determine the mechanism by which the logical theory of probability operates. I argue that this is not the case at all, and that while statistical observations play a part in the formation of the hypothetical laws, these observations are only necessary, but not sufficient conditions for the introduction of these hypotheses.

Wittgenstein a peu écrit au sujet de la probabilité. Le peu que nous avons vient de quelques pages du *Tractatus*, de quelques remarques des années 1930 et de conversations de la décennie lorsqu'il était actif dans le Cercle de Vienne. Néanmoins, l'oeuvre de Wittgenstein a beaucoup contribué au développement de la théorie logique des probabilités. Cet article tente de clarifier et de défendre la conception

de Wittgenstein de la probabilité contre certaines critiques souvent citées qui découlent d'une mécompréhension de sa contribution. Max Black, par example, écrit que la théorie des probabilités, telle que Wittgenstein l'a formulée, n'est susceptible d'être utilisée que dans le contexte de la langue idéale du *Tractatus*. Je tiens qu'en faisant appel aux lois de la nature hypothétique, Wittgenstein est en mesure de faire sens d'énoncés de probabilité qui n'ont pas été complètement analysés. G.H. von Wright critique la caractérisation de Wittgenstein de ces mêmes lois hypothétiques. Il écrit qu'en introduisant de telles lois Wittgenstein rend le trait distinctif de sa théorie superflu, car les lois hypothétiques sont directement inspirées par les observations statistiques et donc déterminent indirectement le mécanisme par lequel la théorie logique des probabilités fonctionne. Je tiens que ce n'est pas du tout le cas; tandis que les observations statistiques jouent un rôle dans la formation des lois hypothétiques, ces observations ne sont que des conditions nécessaires, et pas suffisantes, pour l'introduction de ces hypothèses.

1   Introduction

Wittgenstein did not write very much on the topic of probability. The little we have comes from a few short pages of the *Tractatus*, some 'remarks' from the 1930's, and the informal conversations which went on during that decade with the Vienna Circle. Nevertheless, Wittgenstein's views were highly influential; he heavily influenced Waismann's and Carnap's logical interpretations of probability, and through them, ultimately, modern Bayesianism.

This paper will attempt to clarify Wittgenstein's conception of probability and defend it against two common objections that, I will argue, stem from a misunderstanding of his position and not from any actual faults that can be attributed to it. The first of these objections, due to Max Black, criticises Wittgenstein for formulating a theory of probability that is only capable of being used against the backdrop of the ideal, completely analysable, language of the *Tractatus*. I will argue, against Black, that by way of his notion of the hypothetical laws of nature, Wittgenstein is able to make sense of probability statements involving propositions that have not been completely analysed. The second objection, given by G.H. von Wright, criticises Wittgenstein's characterisation of these very hypothetical laws. Von Wright argues that by introducing them Wittgenstein makes his own logical theory of probability

superfluous, for the hypothetical laws are directly inspired by statistical observations and hence these observations indirectly determine the mechanism by which the logical theory of probability operates. I will argue that this is not the case at all, and that while statistical observations play a part in the formation of the hypothetical laws, these observations are only necessary, but not sufficient conditions for the introduction of these hypotheses.

2   The Classical and Frequentist Interpretations of Probability

The so-called 'probability calculus' consists of a number of axioms, definitions, and theorems derived from them from which one can formulate rules for computing complex probabilities from prior probabilities. The calculus provides us, for instance, with the addition rule for the probability of occurrence of A or B, when we already know the probability of the individual events A and B:

$$(\forall AB)[P(A \cup B) = P(A) + P(B)] \qquad (\text{where } A \cap B = \emptyset)$$

But the probability calculus says nothing about how the prior probabilities themselves (i.e., how '$P(A)$' and '$P(B)$', above) should be determined. In other words, the question of the *meaning* of primitive probability statements is left open. It is precisely with this question of the meaning of prior probability statements that the various 'interpretations' of probability are concerned.

What we now call the classical interpretation of probability was formulated by Pierre Simon Laplace in the early part of the 19th century. The basic principle behind Laplace's theory is the 'Principle of Insufficient Reason',[1] according to which one must consider two outcomes as equally possible if one has no reason to prefer one outcome over the other. For example, consider a six-sided die. Assuming the die is not biased, I cannot say that rolling a 3 is any more likely than rolling another number, for as far as I can tell, all of these outcomes are equally possible. But now once we have divided the possible outcomes of an event into a number of equally possible cases, the probability of a particular outcome becomes a simple ratio: of the cases that are 'favourable' to an outcome, to the total possible cases (Laplace, 1825, p. 7). Thus if I want to calculate the probability of rolling a 2, the number of

---

[1]This is sometimes called the 'Principle of Indifference'.

cases favourable to this outcome is one. If, on the other hand, I would like to calculate the probability of rolling an even number, the number of cases favourable to this outcome is three (i.e., 2, 4, and 6). To get the probability, I divide the cases favourable to the outcome by the total number of possible cases. The probability of rolling a 2, therefore, is 1/6, and the probability of rolling an even number is $3/6 = 1/2$.

The classical interpretation was the dominant interpretation of probability for many decades, and it accords very well with our intuitions concerning games of chance (e.g., craps, roulette wheels, lotteries, etc.). However applying it more generally, as we will see shortly, proves to be problematic. Thus by the mid-19th century, an alternative interpretation began to emerge, largely through the work of Leslie Ellis and John Venn. This came to be called the *frequency* interpretation, and its first rigorous formulation was given by Richard von Mises early in the 20th century.

To understand the motivation behind the frequency interpretation, consider the case of a biased die. For Laplace, statements of probability must be defined with respect to equally probable cases. But suppose I shift the centre of gravity of this die, or file away one of its corners. If I do this, then it will no longer be the case that each outcome is equally possible. But regardless, we would still like to say that there is some specific probability of throwing an even number with this die. On the classical conception, however, this probability appears to be impossible to calculate.

A more telling example has to do with the probability of death (e.g., as used by insurance companies). Suppose we say that the probability of death for a forty year-old non-smoking male is 0.011. It is not clear at all how the classical interpretation can conceive of a case like this one. Von Mises asks: "Are there 1000 different probabilities, eleven of which are 'favourable' to the occurrence of death, or are there 3000 possibilities and thirty-three 'favourable' ones? It would be useless to search the textbooks for an answer, for no discussion on how to define equally likely cases in questions of this kind are given" (von Mises, 1957 [1928], pp. 69-70).

Probability gets defined, for von Mises, as the relative frequency of an outcome with respect to a sequence of repeatable events. For example, given a sequence of 1000 rolls of a single die, we might observe that 6 turns up 300 times. We can then say that the relative frequency of 6 in this sequence is $300/1000 = 0.3$. The longer the sequence, the more closely the observed relative frequency approaches the true probability of rolling a 6 (which is considered to be a 'physical' attribute of the sequence). The probability,

then, of getting a 6 with this die, is a kind of idealisation: it is defined as the limiting value of the relative frequency of 6 with respect to an *infinite* sequence of rolls. To illustrate how this works: suppose that after each roll of the die I calculate the relative frequency of 6, rounding off to the first decimal place. I eventually find, after n rolls, that the relative frequency ceases to change; it remains constant at, say, 0.3. At this point, I increase the number of decimal places to two. The value begins to fluctuate again, but after m more rolls, it again ceases to change; it stays constant at, say .32. I then increase the precision to three places. Again, the relative frequency eventually stabilises; this time at .324. If I continue the process *infinitely*, the relative frequency will be accurate to an infinite number of decimal points (von Mises, 1957 [1928], pp. 14-15).

One problem with von Mises' version of frequentism is that it involves an inference from actually observed sequences of events, which are finite, to infinite sequences of those same events. This was enough for Wittgenstein to reject the frequentist interpretation of von Mises, for as he pointed out, it is possible to infer infinitely many infinite sequences from a finite sequence (which we take as the infinite sequence's initial segment).

> If we infer from the relative frequency of an event its relative frequency in the future, we can of course only do that from the frequency which has in fact been so far observed. And not from one we have derived from observation by some process or other for calculating probabilities. For the probability we calculate is compatible with *any* frequency whatever that we actually observe, since it leaves the time open (Wittgenstein, 1975, §234).

For example, consider a physical process described by a function that maps to 0 for values of x less than n (where n is some very large value), but maps to $2x$ for values of x greater than or equal to n. If we consider only an initial segment of the sequence, we will be led to infer, using the limiting value of the relative frequency as our guide, that the probability of $f(x)$ resulting in 0 is 1. If n is so large that it is humanly impossible for anyone to observe n actual instances of this process, then there is no way to infer the true relative frequency of 0 with respect to the process. Further, infinitely many hypotheses about f are compatible with the relative frequency that we observe. We might hypothesize, for instance, that $(\forall x \geq n)f(x) = 3x$, or that $(\forall x \geq n)f(x) = 4x$, or that $(\forall x \geq n)f(x) = 323467x$ or even that $(n \leq \forall x < 2n)f(x) = 3x$ & $(\forall x \geq 2n)f(x) = 30x$.

3   Wittgenstein's Logical Interpretation

The main sources for Wittgenstein's views on probability are propositions 5.1 - 5.156 of the *Tractatus* (his early period), and also §§225 - 237 of the *Philosophical Remarks* (his middle period). Like Laplace, for Wittgenstein the probability of a proposition depends on our knowledge situation at the time the statement is made. However, for Wittgenstein, probability represents a relation, not between events, but propositions: between the propositions representative of our knowledge situation and the propositions for which we are seeking to fix a probability value to. It is easiest to understand Wittgenstein's interpretation if we consider a truth table. For example:

	A	B	C	$\mathscr{P}:$	$(\sim A \supset (B \vee C))$	$\mathscr{Q}:$	$(\sim B)$
1	T	T	T	T		F	
2	T	T	F	T		F	
3	T	F	T	T		T	
4	T	F	F	T		T	
5	F	T	T	T		F	
6	F	T	F	T		F	
7	F	F	T	T		T	
8	F	F	F	F		T	

Consider the proposition $\mathscr{P}$. We can view it as a truth function, and like any other function, it accepts a number of arguments as input (here given in the first three columns) and produces a determinate output. Now call the *truth grounds* of a proposition the truth values of its truth arguments which make it true (5.101).[2] The rows on which $\mathscr{P}$ is true are the rows 1-7. Therefore the set of its truth grounds is: { $\{T,T,T\}$, $\{T,T,F\}$, $\{T,F,T\}$, $\{T,F,F\}$, $\{F,T,T\}$, $\{F,T,F\}$, $\{F,F,T\}$ }. Similarly, the truth grounds of $\mathscr{Q}$ are on lines 3-4,7-8: { $\{T,F,T\}$, $\{T,F,F\}$, $\{F,F,T\}$, $\{F,F,F\}$ }.[3] Now if we compare the truth grounds of these two propositions, we find that the truth grounds of $\mathscr{Q}$ and $\mathscr{P}$ overlap. But we can ask the question, 'given $\mathscr{P}$, when is $\mathscr{Q}$ true?' $\mathscr{P}$ is true on lines 1-7 for a total of seven instances. In the context of these seven lines, $\mathscr{Q}$ is true three times (on lines 3, 4, and 7). So

---

[2]Here, and in the sequel, references to the *Tractatus* will simply note the proposition number.

[3]Strictly speaking, $\mathscr{Q}$, takes only one argument. Since the truth-values of A and C are irrelevant, however, including them has no effect.

we can say that $\mathcal{Q}$ is true $3/7$ of the time whenever $\mathcal{P}$ is true; i.e., we can calculate the conditional probability: $P(\mathcal{Q}|\mathcal{P}) = 3/7$.

In general, we can say that: "If T_r is the number of the truth-grounds of a proposition 'r', and if T_{rs} is the number of the truth-grounds of a proposition 's' that are at the same time truth-grounds of 'r', then we call the ratio $T_{rs} : T_r$ the degree of *probability* that the proposition 'r' gives to the proposition 's' " (5.15). When two propositions have no truth arguments in common, the probability of one given the other is simply $1/2$. For example, A, B, and C above have no truth arguments in common, and we can see from the truth table that $P(A|B) = P(B|A) = P(A|C) = P(C|A) = P(B|C) = P(C|B) = 0.5$.

The truth-grounds of a proposition can be said to define a proposition's range.[4] Wittgenstein writes, "The truth-conditions of a proposition determine the range that it leaves open to the facts. ... " (4.463). Consider the truth table above once again, and imagine that the totality of our knowledge consists of the elementary propositions A, B, and C. We can say, then, that each line of the truth table represents a possible state description of the universe. For example, line 6 represents a possible state of the universe such that $(\sim A \& B \& \sim C)$. We can now define the range of the proposition $\mathcal{Q}$ as the state descriptions that are compatible with it. Further, we can say that $\mathcal{Q}$ is logically equivalent to the disjunction of the state descriptions making up its range, i.e.: $(A \& \sim B \& C) \vee (A \& \sim B \& \sim C) \vee (\sim A \& \sim B \& C) \vee (\sim A \& \sim B \& \sim C)$

Now for the Wittgenstein of the *Tractatus*, all non-elementary propositions are logically analysable in principle into truth-functions of elementary propositions. Thus if we could perform such a complete analysis, then we could compute the probability of some proposition of unknown truth value simply by comparing its range with the total number of state descriptions that are left open by the other propositions which make up our knowledge situation. Thus imagine, for instance, that we know independently that the state description described by line 1 can be ruled out. Then in the case where we do not know the truth value of $\mathcal{P}$, we can say that its probability given our knowledge situation is $6/7$: it is true six out of seven times (i.e., true on

---

[4]The German word is *Spielraum*. The notion of the *Spielraum* of an hypothesis dates to the 19th century and is first found in von Kries' work on probability. Whether or not Wittgenstein was actually acquainted with von Kries' work is debatable; however his notion of a *Spielraum* is similar to von Kries' notion in its essentials. See Heidelberger (2001) for a discussion.

lines 2-7 but false on line 8).

Now while Black sees this as the chief virtue of Wittgenstein's account, he considers it to be at the same time its chief *defect*, for while it may be possible to completely analyse the propositions of the idealised language of the *Tractatus*, we know of no way of doing this for real languages. Thus, he writes: "Ironically, its one claim to novelty ensures its lack of relevance to concrete examples. ... we have no way of analysing the propositions of ordinary life or of science, and so no way of calculating the degree of probability between propositions. ... 'complete analysis' is a metaphysical mirage, whose hypothetical existence leaves the problem of handling the propositions with which we are acquainted, in all their remoteness from the canons of an 'ideal language', stubbornly insoluble" (Black, 1966, p. 256).

But note, first, that Wittgenstein is not concerned with giving us a practical method for actually calculating prior probabilities, but rather with giving us a theory of the meaning of probability statements. Thus *if* we were to perform a complete analysis of the propositions of our language, any probability that we calculated in that case would be taken as representing the 'true' probability of the proposition in question. In our everyday scientific practice, then, when we make a probability statement, what we mean to say is that the result of the above procedure of complete analysis would agree with the probability that we have asserted. But it is not to say that we actually attempt to perform such a procedure, for second, and more importantly, with respect to the 'everyday' calculation of probabilities, Wittgenstein does not require the sort of analysis that Black envisions at all.

In this regard, it will be instructive if we turn to one of Wittgenstein's own examples. If I place an equal number of white and black balls into an urn and then ask "what is the probability of drawing a white ball from this urn?", the answer will invariably be 1/2. Why? Wittgenstein's answer is the following "... if I say, 'The probability of my drawing a white ball is equal to the probability of my drawing a black one', this means that all the circumstances that I know of (including the laws of nature assumed as hypotheses) give no *more* probability to the occurrence of the one event than to that of the other. ..." (5.154).

The key phrase here, ironically, is the one in parentheses. It is important to understand what Wittgenstein means by these laws of nature assumed as hypotheses.[5] Thus imagine the set of all (not necessarily elementary) propo-

---

[5]In my interpretation of this passage I am indebted to von Wright (1969). However, as

sitions whose truth value is known. Call this set the bulk of our knowledge, $\mathscr{K}$. Now, for Wittgenstein, the proposition $(P \supset (Q \vee R))$ can belong to $\mathscr{K}$ even though neither P nor Q nor R belong to $\mathscr{K}$. This may seem strange at first, but that this is possible follows from Wittgenstein's views on quantification. For Wittgenstein, quantified sentences are to be thought of as schemas or prototypes for constructing propositions. Note that he held this view both in the *Tractatus* as well as later. Thus in the *Tractatus*, he writes: "What is peculiar to the generality-sign is first, that it indicates a logical prototype ..." (5.522); In the *Philosophical Remarks*, he writes: "An hypothesis is a law for forming propositions" (Wittgenstein, 1975, §228).

For example, the proposition: 'If Socrates is a man, then Socrates is mortal' is an instance of the rule specified by the hypothesis: 'All men are mortal'. The latter is a schema for constructing propositions such as the former. Thus if we accept a certain law of nature, e.g., $(\forall_x)(A_x \supset B_x)$, then all propositions capable of being constructed from it, e.g.: $(A_a \supset B_a)$, $(A_b \supset B_b)$, $(A_c \supset B_c)$, ..., will belong to $\mathscr{K}$; and this will be the case even if the truth-values of $A_a, A_b, A_c, B_a, B_b, B_c, ...$ are unknown.

Thus it is possible to have knowledge of a non-elementary proposition without having knowledge of its constituents; for the knowledge of such propositions is *derived* from our general laws and not *built up* from elementary propositions. The hypothetical general laws provide a means to add propositions expressing relations between propositions (elementary or not) to $\mathscr{K}$ without actually adding knowledge of these propositions themselves to $\mathscr{K}$. But if we can 'know' a proposition without knowing the truth or falsity of its constituents, then we can compute probabilities based on these unanalysed propositions as well. This is Wittgenstein's meaning when he writes: "We use probability only in default of certainty—if our knowledge of a fact is not indeed complete, but we do know something about its *form*. ..." (5.156, emphasis mine).

Coming back to the urn example, say that it is a 'law of nature' that 'for all urns containing an equal number of white and black balls, all balls drawn will be either white or a black' (where 'or' is, of course, taken in the exclusive sense), or symbolically: $(\forall_{x \in U})(W_x \oplus B_x)$. This implies that for this particular draw, that $(W_a \oplus B_a)$. Now consider the proposition expressing the fact that I draw a white ball, (W_a). The probability conferred on this

---

I explain below, I disagree with von Wright's negative evaluation of Wittgenstein's views on the probability of hypotheses.

proposition by $(W_a \oplus B_a)$ is 1/2 (all else being equal). Note that we have calculated this probability in spite of the fact that we do not know either W_a or B_a.

But was it not through an analysis of the proposition $(W_a \oplus B_a)$ that we inferred the probability of W_a? Thus even if we do not know the *truth or falsity* of W_a or B_a, we still need to know what the constituents of $(W_a \oplus B_a)$ are in order to determine a probability for W_a, do we not? Yes, however, it need not be a *complete* analysis, for W_a need not be elementary. W_a may itself be a complex proposition, representing a conjunction, perhaps, of further elementary propositions. The laws of nature (which are simply assumed, a priori), tell us something about the relationships between propositions, and in describing these relationships, they of course must describe the constituents of these relationships—but these need not be described in every detail for we do not *need* to know what the constituents of these constituents are; these constituents can represent 'abstractions', in a sense. Thus, these laws of nature express *relationships* between (possibly complex) propositions. These laws constitute the framework by which we describe the world (6.342); they form the a priori assumptions that we hold in the background when we make any statement of probability. Of course, as our knowledge grows, we may learn more about the constituents of these propositions, and this will, in turn, affect our calculated probabilities. Thus assume that $\mathscr{K}$ consists of the single proposition, $(P \supset (Q \vee R))$. Given that I know this proposition, but not P, Q, or R, the probability that this proposition confers on the proposition R is 4/7. Now assume that as our knowledge grows we eventually add Q to $\mathscr{K}$. In this new situation, the probability of R becomes 1/2. And if, one day, we discover that Q is capable of further analysis, and we come to know the truth or falsity of one of its constituents, say, q_1, this will affect the probability of R even further.

4   The Inspiration of Hypotheses

Recall that Wittgenstein was critical of the frequentist interpretation of probability. As we discussed above, for Wittgenstein, observed relative frequencies can have no direct bearing on the calculation of probabilities. By themselves, they show nothing, for, as he was wont to say, they 'leave the time open'. Relative frequencies can have a role to play with regard to probabilities, however, in an indirect way; i.e., they can inspire us to formulate new, or revise

our old, hypothetical laws of nature, or even to postulate that there is some *unknown* law of nature responsible for an observed phenomenon. Whether or not we do this depends on the cost to us: "The probability of an hypothesis has its measure in how much evidence is needed to make it profitable to throw it out. It's only in this sense that we can say that repeated uniform experience in the past renders the continuation of this uniformity in the future probable" (Wittgenstein, 1975, §229).

Now for von Wright, the fact that observed frequencies can 'inspire' new hypotheses in this way makes Wittgenstein's logical definition of probability (i.e., the definition in terms of the measures of ranges or truth grounds) "superfluous as a method for computing probability-values. ... The licensed appeal to "unknown circumstances", i.e. to the operations of unknown laws ... knocks the bottom out of the original definition. ... It does not perform the function or rôle of feeding the calculus with actual values of probabilities. There is no need for this mediating rôle of a definition of probability. Statistical experiences "inspire" *directly* hypothetical assignments of probabilities" (von Wright, 1969, pp. 275-276).

Von Wright overestimates the role that observed relative frequencies have, for Wittgenstein, in 'inspiring' hypotheses, and I think that it is simply incorrect to say that they *directly* inspire hypothetical laws of nature. In one of his conversations with the Vienna Circle in the early thirties, Wittgenstein makes the observation that "If the relative frequency deviates systematically from the probability that was calculated, we as it were lay down the *postulate* that there must be further causes to be found ... The other circumstances that we introduce must not have the character of assumptions contrived *ad hoc*" (McGuinness, 1979, p. 95).

Von Wright makes it appear as though observations of relative frequency alone are enough to prompt a Wittgenstinian to postulate arbitrary unknown laws of nature in order to explain them. But surely there is more to it than this. Take the case of a die (Wittgenstein's example). Imagine that the relative frequency of 1 for a long series of trials converges to 1. We might hypothesise that this die is loaded; if so, we will not stop there, but we will crack it open and inspect it. We will search for a cause—perhaps a shifted centre of gravity, perhaps faulty manufacturing processes or shoddy materials. We will not be satisfied, and will not be willing to accept it as a law of nature that the die can give nothing but ones, until we have somehow connected this fact with our physical theories of the world.

If it turns out, miraculously, that there is no reason that we can give—even

after all of our best scientists and theoreticians have investigated this peculiar phenomenon—why this die gives rise to such a frequency of ones, the scientific community will be in crisis. A photograph of the die will be on the front page of every magazine and of every newspaper; millions of dollars will be offered to the first who can come up with a successful explanation of this phenomenon—all because we will refuse to simply hypothesise, without sufficient reason, that there is some special but unknown law of nature such that this particular die can give nothing but ones.

And even in the case where we throw up our hands and say that there is some unknown law of nature at work, we will still attempt to constrain the unknown law in such a way that, whatever it is, it accords with our other hypothetical laws of nature; we will say that it must be a law of such and such a kind, or that the general characteristics of its operation must be by way of such and such a mechanism. But we will not simply say that it is a law of nature that this die gives nothing but ones without further ado. In short, relative frequencies alone do not inspire hypothetical laws of nature; they do so *only when they can be reconciled* with our network of already established hypotheses.

> In actual fact, he will refuse to accept it as a natural law that he can throw nothing but ones. At least, it will have to go on for a long time before he will entertain this possibility. But why? I believe, because so much of his previous experience in life speaks against there being a law of nature of such a sort, and we have—so to speak—to surmount all that experience, before embracing a totally new way of looking at things (Wittgenstein, 1975, §234).

One might respond, of course, that even if statistical observations do not directly inspire hypotheses, nevertheless they do amount to necessary conditions for these hypotheses; for surely all our hypothetical laws of nature must at least begin from statistical observation. Obviously Wittgenstein would not deny this. However this is, in my view, a trivial point; for the question here is one of relative stability. Our hypotheses presuppose an enormous system of experience. We do not simply alter them with every new statistical observation, and we only alter them if they can fit coherently into the grand "mesh" which we use to describe the world. In other words, while statistical observation may indeed 'drive' or 'inspire' us to hypothesise new laws of nature, by themselves they do not constitute a sufficient basis

for them; for the journey from 'inspiration' to hypothesis is typically a long and arduous one—not a quick ad hoc procedure, as von Wright would have us believe.

5   Conclusion

It is unfortunate that Wittgenstein did not write more on the subject of probability. He gives us nothing like an objective measure of the degree of confirmation for a hypothetical law of nature (i.e., *just how much* evidence is required to 'give it up'), nor does he give us anything but vague ideas as to the amount of evidence required in order to establish these hypotheses in the first place. Then there are the 'pure statistical' cases; e.g., the probability of death of a forty year old non-smoking male. With respect to *these* cases, von Wright's account of 'inspiration' may be correct. But aside from an obscure comment that "This has nothing at all to do with probability" (McGuinness, 1979, p. 94), Wittgenstein leaves us in the dark as to how these examples are related to a logical theory of probability. This all said, Wittgenstein's ideas are at least not self-defeating. His definition of logical probability is not superfluous as von Wright has asserted, nor is it irrelevant as Black has tried to show.

References

Black, M. (1966). *A Companion to Wittgenstein's Tractatus*. Ithaca, NY: Cornell University Press.

Heidelberger, M. (2001). Origins of the logical theory of probability: Von Kries, Wittgenstein, Waismann. *International Studies in the Philosophy of Science*, *15*, 177–188.

Laplace, P. S. (1825). *Essai philosophique sur les probabilités*. Paris: Courcier.

McGuinness, B. (Ed.) (1979). *Wittgenstein and the Vienna Circle: conversations recorded by Friedrich Waismann*. Trans. B. McGuinness, & J. Schulte. Oxford: Basil Blackwell.

von Mises, R. (1957 [1928]). *Probability, Statistics and Truth*. Trans. J. Neymann, D. Scholl, & E. Rabinowitsch. London: George Allen and Unwin Ltd.

von Wright, G. H. (1969). Wittgenstein's views on probability. *Revue Internationale de Philosophie*, *23*, 259–283.

Wittgenstein, L. (1975). *Philosophical Remarks*. Trans. R. Hargreaves, & R. White. Oxford: Basil Blackwell.

Wittgenstein, L. (2005 [1921]). *Tractatus Logico-Philosophicus*. Trans. D. Pears, & B. McGuinness. London: Routledge.