If we now let $j_n(\omega)$ stand for the number of heads in the first n trials of ω, the weak and strong form of the law of large numbers can be stated as follows:

WLLN The $\mathscr{P}\imath$ measure of the set of ω's for which $|(j_n(\omega)/n) - p| > \varepsilon$ approaches 0 as $n \to \infty$ for any $\varepsilon > 0$.

SLLN The $\mathscr{P}\imath$ measure of the set of all ω's such that $\lim_{n\to\infty}(j_n(\omega)/n) \neq p$ is 0.

To put (SLLN) in its positive form, the Pr probability is one that the limiting relative frequency of heads converges to p.

As indicated in section 7, a form of the weak law of large numbers can be formulated and proved without the help of countable additivity. Roughly, for any $\varepsilon > 0$, the probability (in the objective sense or in the degree-of-belief sense tempered by Lewis's principal principle) that the actually observed relative frequency of heads differs from p by more than ε goes to 0 as the number of flips goes to infinity. This form of the law of large numbers is to be found in the work of Bernoulli. The strong form of the law of large numbers, which requires countable additivity, was not proved until this century (see Billingsley 1979 for a proof).

3 Success Stories

The successes of the Bayesian approach to confirmation fall into two categories. First, there are the successes of Bayesianism in illuminating the virtues and pitfalls of various approaches to confirmation theory by providing a Bayesian rationale for what are regarded as sound methodological procedures and by revealing the infirmities of what are acknowledged as unsound procedures. The present chapter reviews some of these explanatory successes. Second, there are the successes in meeting a number of objections that have been hurled against Bayesianism. The following chapter discusses several of these successful defenses. Taken together, the combined success stories help to explain why many Bayesians display the confident complacency of true believers. Chapters 5 to 9 will challenge this complacency. But before turning to the challenges, let us give Bayesianism its due.

1 Qualitative Confirmation: The Hypotheticodeductive Method

When Carl Hempel published his seminal "Studies in the Logic of Confirmation" (1945), he saw his essay as a contribution to the logical empiricists' program of creating an inductive logic that would parallel and complement deductive logic. The program, he thought, was best carried out in three stages: the first stage would provide an explication of the qualitative concept of confirmation (as in 'E confirms H'); the second stage would tackle the comparative concept (as in 'E confirms H more than E' confirms H''); and the final stage would concern the quantitative concept (as in 'E confirms H to degree r'). In hindsight it seems clear (at least to Bayesians) that it is best to proceed the other way around: start with the quantitative concept and use it to analyze the comparative and qualitative notions. The difficulties inherent in Hempel's own account of qualitative confirmation will be studied in section 2. This section will be devoted to the more venerable hypotheticodeductive (HD) method.

The basic idea of HD methodology is deceptively simple. From the hypothesis H at issue and accepted background knowledge K, one deduces a consequence E that can be checked by observation or experiment. If Nature affirms that E is indeed the case, then H is said to be HD-confirmed, while if Nature affirms $\neg E$, H is said to be HD-disconfirmed. The critics of HD have so battered this account of theory testing that it would be unseemly to administer any further whipping to what is very nearly a dead horse.[1] Rather, I will review the results of the jolly Bayesian postmortem.

Suppose that (a) $\{H, K\} \models E$, (b) $0 < \Pr(H/K) < 1$, and (c) $0 < \Pr(E/K) < 1$.[2] Condition (a) is just the basic HD requirement for confirmation. Condition (b) says that on the basis of background knowledge K, H is not known to be almost surely true or to be almost surely false, and (c) says likewise for E. By Bayes's theorem and (a), it follows that

$$\Pr(H/E \,\&\, K) = \Pr(H/K)/\Pr(E/K). \tag{3.1}$$

By applying (b) and (c) to (3.1), we can conclude that $\Pr(H/E \,\&\, K) > \Pr(H/K)$, i.e., E *incrementally confirms* H relative to K. Thus Bayesianism is able to winnow a valid kernel of the HD method from its chaff.

(To digress, this alleged success story might be questioned on the grounds that HD testing typically satisfies not condition (a) but rather a condition Hempel calls the "prediction criterion" of confirmation; namely, (a′) E is logically equivalent to $E_1 \,\&\, E_2$, $\{H, K, E_1\} \models E_2$, but $\{H, K\} \not\models E_2$. That is, HD condition (a) is satisfied with respect to the conditional prediction $E_1 \rightarrow E_2$, but the total evidence consists of E_1 and E_2 together. Let us use Bayes's theorem to draw out the consequences of (a′). It follows that $\Pr(H/E_1 \,\&\, E_2 \,\&\, K) = \Pr(H/E_1 \,\&\, K)/\Pr(E_2/E_1 \,\&\, K)$. Thus if $\Pr(E_2/E_1 \,\&\, K) < 1$ and $\Pr(H/E_1 \,\&\, K) = \Pr(H/K)$, the total evidence $E_1 \,\&\, E_2$ incrementally confirms H. These latter two conditions are satisfied in typical cases of HD testing. For example, let H be Newton's theory of planetary motion, let E_1 be the statement that a telescope is pointed in such and such a direction tomorrow at 3:00 P.M., and let E_2 be the statement that Mars will be seen through the telescope. Presumably, E_1 is probabilistically irrelevant to the theory, and E_2 is uncertain on the basis of E_1 and K.)

Notice also that from (3.1) it follows that the smaller the value of the prior likelihood $\Pr(E/K)$, the greater the incremental difference $\Pr(H/E \,\&\, K) - \Pr(E/K)$, which seems to validate the saying that the more surprising the evidence is, the more confirmational value it has. This observation, however, is double-edged, as we will see in chapter 5.

The problem of irrelevant conjunction, one of the main irritants of the HD method, is also illuminated. If $\{H, K\} \models E$, then also $\{H \,\&\, I, K\} \models E$, where I is anything you like, including a statement to which E is, intuitively speaking, irrelevant. But according to the HD account, E confirms $H \,\&\, I$. In a sense, the Bayesian analysis concurs, since if $\Pr(H \,\&\, I/K) > 0$, it

follows from the reasoning above that E incrementally confirms $H \,\&\, I$. However, note that it follows from (3.1) that the amounts of incremental confirmation that H and $H \,\&\, I$ receive are proportional to their prior probabilities:

$$\Pr(H/E \,\&\, K) - \Pr(H/K) = \Pr(H/K)[(1/\Pr(E/K)) - 1]$$

$$\Pr(H \,\&\, I/E \,\&\, K) - \Pr(H \,\&\, I/K) = \Pr(H \,\&\, I/K)[(1/\Pr(E/K)) - 1].$$

Since in general $\Pr(H \,\&\, I/K) < \Pr(H/K)$, adding the irrelevant conjunct I to H lowers the incremental confirmation afforded by E.

Finally, it is worth considering in a bit more detail the case of HD disconfirmation. Thus, suppose that when Nature speaks, she pronounces $\neg E$. If $\{H, K\} \models E$ and if K is held to be knowledge, then H must be false, so HD disconfirmation would seem to be equivalent to falsification. But as Duhem and Quine have reminded us, the deduction of observationally decidable consequences from high-level scientific hypotheses often requires the help of one or more auxiliary assumptions A. It is not fair to ignore this problem by sweeping the A's under the rug of K, since the A's are often every bit as questionable as H itself. Thus from Nature's pronouncement of $\neg E$ all that can be concluded from deductive logic alone is that $\neg H \vee \neg A$. If HD methodology were all there is to inductive reasoning, then there would be no principled way to parcel out the blame for the false prediction, and we would be well on the way to Duhem and Quine holism (see section 4 below). In particular, H could be maintained come what may if the only constraints operating were those that followed from direct observation and deductive logic. But the fact that the majority of scientists sometimes regard the maintenance of a hypothesis as reasonable and sometimes not is a fact of actual scientific practice that cries out for explanation. The Bayesian attempt at an explanation will be examined in section 7 below.

2 Hempel's Instance Confirmation

Having rejected the HD or prediction criterion of confirmation, Hempel constructed his own analysis of qualitative confirmation on a very different basis. He started with a number of conditions that he felt that any adequate theory of confirmation should satisfy, among which are the following: