In one form or another, this confrontation has held for some time. Our fundamental political document comes directly out of it. The Founding Fathers were caught up in the problem: they wound up adding a liberal Bill of Rights to their welfarist Constitution. This was an act of statesmanship, a commitment to both sides at once. There is no way of harmonizing conflicting ideologies. But there are ways of living in tension between them.

FREDERIC SCHICK

Rutgers University

COMMENTS AND CRITICISM

CONDITIONALIZATION*

ISAAC LEVI and Teddy Seidenfeld have taken me to task for failing to give conditionalization its due.† The issues they have raised are important and bear not merely on the analysis of inductive logic and scientific inference, but on such down-to-earth matters as the role of randomization in epidemiological studies and on such broad philosophical questions as the morbidity of empiricism. While it is unrealistic to expect to resolve fundamental philosophical issues in these pages, it is worth while getting clear about what they are; a number of the other issues concern matters of fact and matters of emphasis about which it might be hoped that agreement could be reached.

I

In order to set the stage, we must first of all briefly explore the relation between Levi's approach and vocabulary and mine. In *The Logical Foundations of Statistical Inference*[1] (henceforth LFSI) I

define a logical notion of probability, roughly as follows: Given a language L (a first-order language containing set theory), a set of statements K in that language representing a body of knowledge, the *probability of a statement* S is the interval (p,q) just in case: there are terms of the language, x, y, and z, such that '$S \equiv x \epsilon z$' is in K; '$x \epsilon y$' is in K; there is a statement in K to the effect that the proportion of y's that are z's lies between p and q; and, relative to K, x is a *random member* of y with respect to z. The metalinguistic relation "is a random member of" (discussed in DI, and in RRRC) is the relation that obtains when y is the appropriate "reference class" for assessing the probability of '$x \epsilon z$' and therefore of S.[2]

I do not stipulate that K be consistent or deductively closed, but for purposes of comparison with Levi's approach we shall here suppose that it is.[3] If it is, then, under some very mild constraints, it follows that: Every statement in the language has a probability—though in some cases it may be the uninformative probability $(0, 1)$; for any finite set of sentences there exists a classical probability function Q_K such that the domain of Q_K is the set of sentences of L, and for any sentence S in the set, $Q_K(S)$ lies in the epistemological probability interval for S, $EP_K(S) = (p, q)$.

Levi characterizes probability in a quite different way. He takes the credal state of an agent to consist of a set of Q_K-functions; this set of Q_K-functions is to be coherent (each Q_K-function satisfies the axioms of the probability calculus) and convex (the weighted mixture of a number of Q_K functions is to be a Q_K function in the set). He also imposes a condition of *consistency:* The set of Q_K-functions shall be nonempty if and only if K is consistent.

Note that, at this point, both Levi and I emerge with interval-valued probabilities. But there are already differences. For me, the intervals are determined, once and for all, by the set of statements K: given a body of knowledge, there is only one credal state permissible for an agent. Levi regards this as implausible, and labels me a "necessitarian." On the other hand, each of the Q_K functions that make up the credal state of an agent for Levi is a function that

[2] For the latest and clearest version of those rules, see my "Randomness," forthcoming.

[3] I reject consistency and closure as general requirements. This is not to say that I think people should be inconsistent, but merely that I think there may be circumstances under which it would be philosophically advantageous to consider rational corpora whose deductive closures are inconsistent. My original description of what has come to be called "the lottery paradox" was intended to illustrate one such possible circumstance. Anyone who wants to sacrifice generality for conventionality in this regard is free to do so—but then he must provide more constraints on his rule of acceptance than I have; and the symmetries of the lottery suggests to me that these constraints must be ad hoc.

could be used to determine coherent betting odds simultaneously on all the sentences of a language. This strikes me as implausible, and I hereby label Levi, with a mild degree of opprobrium, a "universalist."

Nevertheless, there is a close connection between Levi's intervals and mine. Consider a finite set of sentences. Let Q_K be called "fitting" if, for every S in that set, $Q_K(S) \epsilon EP_K(S)$. This set of Q_K-functions is coherent and convex. Question: is every point in $EP_K(S)$ representable as a value of $Q_K(S)$ for some Q_K in the set of fitting Q_K-functions? This turns out to be a nontrivial question, but I would not find it disturbing if the answer should turn out to be negative: I do not *require* my agent to be a universal bookmaker, and in any specific circumstance, I know that he will not allow a book to be made against him. Leaving this question to one side, however, for present purposes, let us take the answer to the question to be affirmative. We shall suppose that $EP_K(S)$ can be identified with the set of numbers $Q_K(S)$ for fitting Q_K-functions.

Conditional epistemological probability (*CEP*) is defined very straightforwardly for me: $CEP_K(S, T) = EP_{K'}(S)$, where K' is the set of sentences obtained by forming the deductive closure of K and T. As I pointed out in Chapter 11 of LFSI, there need exist *no* fitting Q_K function such that $Q_K(S \wedge T)/Q_K(T) \epsilon CEP_K(S, T)$. "Conditionalization fails" in the sense that the multiplication theorem, applied to fitting Q_K functions, may not yield a fitting $Q_{K'}$-function. (In the following section I shall consider the circumstances under which conditionalization does *not* fail in this sense.)

Levi takes the confirmational commitment of an agent to be the criterion he endorses for evaluating revisions of his credal state due to changes in his corpus of knowledge (DI 18). Where the change consists of adding the statement T, consistent with K, to K and then forming the deductive closure, the new credal state should consist of the set of $Q_{K'}$-functions satisfying $Q_{K'}(S) = Q_K(S \wedge T)/Q_K(T)$ for some Q_K in his initial credal state. Now this will not always happen; it may be that when the agent adds T to his corpus, he also changes his confirmational commitment. But as long as his confirmational commitment remains unchanged, he should derive his new credal state from the old one via conditionalization. This is the principle of confirmational conditionalization, and it is this principle that Levi shows to conflict with my rules for direct inference.[4]

[4] In DI, Levi puts it this way: "Assuming that coherence and consistency are sacrosanct, either conditionalization or the rule of direct inference must be sacrificed" (21); in CC he writes, "Kyburg's account of direct inference cannot

With the help of some distinctions introduced by Levi, we can make clear what is at issue here. Again consider a finite set of statements, the epistemological probability function EP_K, and the set of fitting Q_K-functions. For any statement S in the set, either EP_K or the set of fitting Q_K functions will constrain the odds that the agent may accept or offer on direct bets on S. When we turn to more complicated bets, the situation changes.

Called-off bets: Consider a bet on S which is called off if T fails to be true. According to Levi, the set of functions $Q_{K'} = Q_K(S \wedge T)/Q_K(T)$ determines the appropriate odds for such bets. This is a matter of credal coherence, and does not involve confirmational conditionalization.

Conditional bets: These are bets on S that are conditional on the acceptance of T into the corpus K. If confirmational conditionalization is accepted, the same set of probabilities $Q_{K'}$ already mentioned will be used to evaluate these bets. On my view, the odds are determined by the conditional probability $CEP_K(S, T)$, which need *not* correspond to the set of coherent $Q_{K'}$-functions determined by the multiplication theorem. According to Levi, the odds for both sorts of bets should be the same; and according to him, I must prescribe differently, on pain of abandoning coherence.

Levi is perfectly correct in noting that there are circumstances under which I would prescribe differently for called-off bets and for conditional bets. His feeling that this is unacceptable stems, I think, from his universalism: his feeling that an agent's credal state should determine (if only approximately) the odds at which he would bet on any proposition in the language and that his confirmational commitment should similarly provide guidance for any set of hypothetical bets. But I see no reason in general for demanding that the odds that I would offer on S, if and when T should become an ingredient of my rational corpus, should be the same as the odds I

consistently satisfy the principles of credal consistency, coherence, and convexity and also the principle of confirmational conditionalization—provided that confirmational commitment is defined for every consistent and deductively closed set of sentences K" (730–731). He writes later, in CC, that "if Kyburg is to have his position, he is . . . under strong pressure to abandon confirmational conditionalization." I never endorsed confirmational conditionalization. In discussing independence I claim that it is often the case that a general hypothesis h is relevant to a bit of evidence e, but not vice versa—a clear violation of Bayesian conditionalization. I in fact offer counterexamples myself to the application of the multiplication theorem. The significance of the red herrings of Malmo is unclear. But I think the question at issue is not whether EP satisfies confirmational conditionalization, but whether my rules of direct inference can be grafted onto a treatment of credal states satisfying confirmational conditionalization. Levi shows decisively that they cannot be.

offer now, when my corpus lacks T, for a bet on S which is to be called off unless T becomes added to my corpus in advance of S.[5] I may be compelled (or bribed) to make a book now (lacking knowledge of T) on the algebra whose atoms are $S \wedge T$, $S \wedge \sim T$, $\sim S \wedge T$, and $\sim S \wedge \sim T$. If the book includes called-off bets, the odds will conform to one of the functions $Q_{K'}$. On the other hand, I may be required to make a *hypothetical* book on the algebra whose atoms are S and $\sim S$, which is the book I am *now* committing myself to use later on, on the condition that T becomes part of my rational corpus. I see no reason why these two sets of odds should be related in the way required by confirmational conditionalization.

What set of odds I use depends, of course, on the particular set of options I am confronted with. If all I am considering is the called-off bet, so that I need not post odds on the whole algebra generated by S and T, then it seems quite reasonable for me to use the odds appropriate to the corpus consisting of the deductive closure of K and T, and thus the odds determined by the conditional epistemic probability $CEP_K(S, T)$. In this case I may treat the called-off bet as a conditional bet. It is only where I must make book on the whole algebra that the distinction is important.

I shall return to Levi's arguments in section IV.

II

Before discussing the significance of the failure of conditionalization, let us look at the role that conditional measures play in the system of LFSI.[6] We shall see that, in many cases, these conditional measures do the work traditionally assigned to conditional probabilities, at least by non-Bayesians.

There are two quite natural ways in which we become concerned about conditional probability. In one, we are concerned with a measure in a subset: the conditional probability of getting an ace on a role of a die, given that we have got an odd number. There we want the measure of aces among odd numbers ($\frac{1}{3}$), which is just the measure of aces and odd numbers (the measure of aces, $\frac{1}{6}$), divided by the measure of odd numbers ($\frac{1}{2}$). In the second, we are

---

[5] It might be that the conflict of intuitions here is related to the conflict of intuitions among those who adopt an intensional object-language approach to probability which is reflected in the dispute about whether the probability of a conditional should or should not be the same as a conditional probability. In one case we are concerned with the probability of the conditional: What are the odds that if T is true, S will be true? in the other with the conditional probability: If I know T, what should the odds on S be?

[6] In ch. 11 of LFSI I distinguish between epistemological independence and stochastic independence. I go on to contrast conditional probability and conditional measure. I should, perhaps, have continued the same terminology, and discussed epistemological conditionalization and stochastic conditionalization.

concerned with a measure in a product set: the conditional probability of getting a red chip on the second draw from an urn containing three red and three black chips, given that we got a red chip on the first draw, is the measure of the set of pairs of red chips divided by the measure of the set of pairs in which the first chip is red.

It is the latter sort of conditional measure which has been of greatest interest to those who have been concerned with inductive problems. Both sorts of conditional measure play an important role in applications of statistical knowledge. This role is reflected in the fact that (epistemological) conditional probabilities, under a wide range of circumstances, are measured by the usual conditional measures. This is proved in a number of metatheorems in Chapter Eleven of LFSI. For example, (MT 11.1) suppose that S, T, and the conjunction of S and T have probabilities determined by the same reference class (this is true of the example of the die; it is true wherever epistemic conditionalization leads us from a measure in a set to a measure in a subset); then there are values in the corresponding probability intervals which do satisfy the multiplication theorem. More generally, it follows from MT 11.3 that if the appropriate condition concerning randomness is met, the epistemological conditional probability will reflect the usual conditional measure and, thus, will satisfy the principle of confirmational conditionalization. To put the matter another way, whenever the addition of T to our corpus K does not lead to a different statistical model for the assessment of S, confirmational conditionalization will be satisfied. These are loose ways of putting the matter, but they should indicate that when the multiplication theorem is violated, it is violated for a reason.

An example will show that, in a large and important class of inductive cases involving statistical inference, confirmational conditionalization is satisfied. Suppose that we know (it is part of the contents of K), for every i, that the proportion of red balls in urn i is p_i. We select an urn and then a sample of balls from that urn. The principle of confirmational conditionalization would have us take the probability of the statement that we have urn i to be the prior probability that we have urn i, multiplied by the probability that we obtained a sample of the sort we did obtain given that we have urn i, divided by the general probability of getting a sample of that composition. Put more simply, the probability that we have urn i is to be the conditional probability that we have urn i, given that we obtained a sample of the sort we did obtain. This holds unproblematically for conditional epistemological probability provided

that a single condition is satisfied: that relative to the corpus containing a description of the sample, the urn be a random member of the set of urns yielding such samples in experiments of the sort described. Clearly this will generally be the case.

Conditional epistemological probability will be determined by a conditional measure when the knowledge on which we are conditionalizing does not interfere with appropriateness of the natural reference class. The natural reference class is the generic subset determined by the data. We can say this because in instances of the second sort—draws from urns—we can put everything in terms of cartesian products. For example, the initial draw of a red ball can be construed as a draw of a pair (or a sequence) of which the first item is red. When conditional epistemological probability is determined by the natural conditional measure, then confirmational conditionalization will be satisfied. Levi's example—and mine in LFSI—are set up in such a way that the data on which we are conditioning do *not* lead us to the generic subset. In general this occurs when our statistical knowledge of the generic subset is relatively vague, but does not differ (in my technical sense) from our relatively precise knowledge of some other potential reference class.

Conditionalization involving continuous distributions is a little trickier, as we shall see in the discussion of Seidenfeld's paper. To the extent that we have knowledge of continuous distributions, we may be said to have knowledge of the corresponding conditional measures—there is no problem about the existence of the conditional measures. But when we come to apply these conditional measures, the question of the way we approach them as limits of the discrete measures we may plausibly claim to know about may be crucial. In these cases it is not the *failure* of conditionalization that concerns us, but the fact that we may not find it easy to decide what conditional measures are appropriate.

The important point to remember is that although conditional epistemological probability is not a numerical function of epistemological probability, nor even a numerical function of the Q_K-functions corresponding to epistemological probability, we have in our rational corpora all the conditional *measures* that anybody could want. The question at issue is only when those conditional measures determine the values of conditional epistemological probabilities.

III

With this as background, let us turn to Seidenfeld's discussion of sufficiency and statistical inference. Seidenfeld's first point is that the "principle of sufficiency" does not hold for epistemological probability, "with dire consequences" for my treatment of inverse

inference. Since sufficiency is not a notion popularly bandied about among philosophers, but since it is so intimately connected with conditionalization and relevance, which are, it will be valuable to explain this in detail.

Suppose we have a population, actual or potential, and a random quantity X defined for members of that population. For simplicity, suppose that X is real-valued. A distribution of X in that population specifies for every real number r the measure of the subset of that population satisfying the condition $X(x) \leq r$. The object of statistical inference is to infer from a sample of that population something about this distribution: for example, that it is D, that it is highly probable that it is a distribution in a certain set of distributions $\{D\}$, that it is a distribution with a mean falling within certain limits, etc. The data for this inference consists of a sample from the population. The set of possible samples from the population we call the *sample space*. The inference is based on a function defined on the sample space which may be real-valued (for example, it may be the mean value of X for the objects in the sample—the sample mean) or vector-valued (for example it may be the observed value of the mean, together with the observed value of the variance in the sample) or the order statistic (the set of X-values observed in the sample ordered according to magnitude) or maybe even the identity function (the set of values observed in the order in which they are observed).

Sufficiency is defined only relative to a family of distributions. Here is a typical definition:

A statistic $T = t(X)$ is said to be sufficient for a family of distributions if and only if the conditional distribution of X given the value of T is the same for all members of the family.[7]

I shall call the notion of sufficiency just defined *statistical* sufficiency. What a sufficient statistic does is to partition the sample space into subsets in such a way that any two members of a given subset yield the same conditional distribution of the (real- or vector-valued) parameter θ indexing the family. It thus provides a criterion of *generic* relevance, given a family of distributions. For making an inference about the mean of a normal population of known variance, for example, the sample mean is sufficient. This is to say that the sample variance, for example, is irrelevant. Under ordinary circumstances the sample mean preserves all the relevant information in the sample. For a normal distribution, the sample median normally does not.

[7] Bernard W. Lindgren, *Statistical Theory* (New York: Macmillan, 1968), p. 228.

This may suggest, mistakenly, that statistical sufficiency is an epistemological notion: it appears in the company of such terms as 'evidential import', 'relevance', etc. The suggestion is misleading. What are referred to in the definition are general functions T and X, general conditional distributions, and so on. It is the general function $T = t(X)$ which is said to be sufficient. There is no reference to specific circumstances, particular background knowledge, or the like. In particular, it says nothing about inference. For that we need a *principle* of sufficiency, to which we will come in due course.

Two things should be observed about sufficient statistics. First of all, a sufficient statistic always exists, namely, the identity function that partitions the sample space into the unit sets of its members. The interest for statisticians of sufficiency is twofold: if one is interested in reducing, by a general rule, a complex body of data to simple form for the purpose of making an inference about a certain family of distributions, one is interested in making that reduction by means of a statistic sufficient relative to that family of distributions; and one is interested in making the maximum possible reduction—finding a *minimal* sufficient statistic—relative to the family of distributions in question. Note that if T is a sufficient statistic and U is a statistic from which $T = f(U)$ can be recovered, then U is also sufficient.

Second, note that the definition of statistical sufficiency refers to the conditional distribution of X given T. It does not speak of probability, and in general one focuses on densities rather than probabilities. It is a purely statistical notion, and quite appropriately makes no mention of background knowledge.

Seidenfeld's statement of sufficiency and the sufficiency principle is different from the formulation I have just given. Seidenfeld's definition of sufficiency involves reference to background knowledge K and is stated in terms of conditional probabilities, rather than in terms of conditional distributions. In this he is following Dennis Lindley, and other subjectivists and personalists, and of course this is crucial for the plausibility of his sufficiency principle:

> If and only if t is sufficient for d with respect to θ, then inference from t (alone) preserves all the relevant evidence contained in d concerning θ.[8]

[8] DIII, p. 710. There are notational ambiguities which should not be allowed to get in the way of our understanding. In the statement of sufficiency just quoted T is a *function* defined on the sample space; the sample space itself consists of vectors or n-sequences of real numbers. It is the function T that is said to be sufficient. Seidenfeld writes of making an "inference from t," and writes placeholders "t" and "d" in propositional locations; we must therefore read "t" (some-

The trouble is, it is now no longer a simple matter to know that a reduction of the data d is sufficient. Seidenfeld says that "for the binomial distribution with statistically independent trials a sufficient set of statistics for sample data (with respect to inference about the binomial parameter) is the pair: number of trials and frequency of outcome" (716). It is perfectly true that ⟨number of trials, frequency of success⟩ is a statistically sufficient statistic for discriminating among binomial distributions. But it may or may not be the case that this statistic contains all the "evidential import" or all the "relevant information" embodied in the particular sample actually observed. (We might know, for example, that the parameter p has a value equal to the reciprocal of the ordinal number of the first success.[9])

One way of responding to this situation would be to say that we have got hold of the wrong statistical model: we should be using one that reflects our knowledge of the particular case at hand, rather than one appropriate to the *generic* problem of binomial inference.

The point is that, as Seidenfeld construes it, sufficiency is not a statistical notion at all, but an epistemological or credal notation. This is clear from his characterization (716):

[t is sufficient just in case] $Q_K(d/t \ \& \ H_\theta) = Q_K(d/t)$

For this identity to make sense at all, d and t must be construed as representing *statements*, not numbers or functions, and whether it holds or not depends on the details of the agent's beliefs, and not on statistical generalities. The confident assertion that the sample mean is sufficient for an inference about the mean of a normal population of known variance is warranted only for the *statistical* notion of sufficiency.

For example, suppose that it is known in K that we are sampling from a normal population of unknown mean and known variance, and that d is "The sequence of observed sample values is ⟨r_1, . . . , r_n⟩" and that t is "the sample mean is $\bar{r}$." I may perfectly well regard the median, or the fourth observation, as relevant to H_θ, even in the presence of t. That is, $Q_K(d/t \wedge H_\theta) = Q_K(d/t)$ might

---

times) as "The T-value of the particular sample we have observed is so and so," and "d" as "The particular sample we have observed consisted of the sequence of real numbers ⟨r_1, . . . r_n⟩."

[9] One might note that for serious personalists and subjectivists there *are* no "statistically independent" trials; the outcomes on the first n trials have an important bearing on the probability of success on the $n + 1$st trial. In this case exchangeability will do in lieu of independence; nevertheless it is clear that no general *statistical* considerations can determine whether or not a given "reduction of the data" entails the loss of "relevant" information.

be false, even when t represents the canonical reduction of d. Of course this can be explained away: we can say that this shows that the agent does not regard the sample as representing exchangeable events or quantities, or that he does not regard it as a sample from a single population, and therefore that the canonical model is inappropriate. But however the failure is *explained*, the fact remains that ordinary statistical sufficiency, which concerns distributions or measures or chances, does not ensure credal or epistemological sufficiency, which has to do with beliefs.

Moreover, there is a perfectly natural definition of *epistemological* sufficiency according to which the sufficiency principle holds (trivially) and which is related to the statistical notion of sufficiency by a natural condition concerning randomness.

Recall that what a sufficient statistic does, essentially, is to induce a partition of the sample space into subsets which are generic-relevance equivalence classes—subsets such that any two points in the sample space belonging to the same subset have the same generic evidential relevance to the inferences concerning the family of distributions to which the distribution X belongs. We may easily relativize this to a body of knowledge K:

> A statistic $T = t(X)$ is epistemologically sufficient for a family of distributions, relative to K, just in case the partition of the sample space induced by T is such that if X_1 and X_2 belong to the same equivalence class [i.e., if $t(X_1) = t(X_2)$] then the probability of a statement about the distribution of X (that it has a certain mean, that it is one of a set of distributions, etc.) will be the same relative to K and X_1 as it is relative to K and X_2.

The condition of randomness is simply that the sample be a random member of the sample space with respect to the property in question relative to the truncated corpus containing only the reduction of the data to t, if and only if it is so relative to the full corpus containing the data d.

Alternatively: Let $T = t(X)$ be a statistically sufficient statistic for a family of distributions. If X_1 is a random member of the sample space, with respect to differentiating among members of the family, then $T_1 = t(X_1)$ is epistemologically sufficient for that purpose.

In cases of statistical inference, the "property in question" is ordinarily something like the representative property that Seidenfeld mentions. But in inferences concerning the unknown mean of a normal distribution, it is not merely statements of the form $|\bar{x} - \mu| \leq \epsilon$ that we can make probability statements about, but

statements of the form $\bar{x} - \mu \in B$, where B is any Borel set at all.[10] We have to select among such statements only when we are concerned to detach one from its evidence, accept it into a lower-level rational corpus, and regard it as "practically certain" and as a datum for further inferences. Under these circumstances we want to embody as much information in the detached statement as possible, and hence we seek something like a shortest or most informative interval.

Seidenfeld argues that epistemological probability fails to satisfy "the" principle of sufficiency, because it fails to satisfy the principle of conditionalization and because the failure of sufficiency entails unacceptable difficulties for statistical inference. I have argued that we should distinguish three kinds of sufficiency: *statistical*, for which we can demonstrate that the sample mean is sufficient for the population mean; *credal*, for which we cannot tell when a statistic is sufficient, except by looking at a belief function for the whole body of knowledge K; and *epistemological*, which concerns beliefs, but is intimately and naturally connected to statistical sufficiency through a condition concerning randomness. Epistemological probability *does* satisfy the epistemological principle of sufficiency.

But what of Seidenfeld's demonstration and his examples? His demonstration is of the fact that credal conditionalization entails a principle of credal sufficiency, and cuts no ice since I reject conditionalization as a general principle and since I find credal sufficiency a principle too hard to apply in practice. The examples, on the other hand, show that the principle of epistemological sufficiency does not apply where I thought it did—and where it clearly should. Since the examples are quite different, I shall comment on them one by one.

The first concerns repeated measurements of the weight of a body of liquid. We suppose that we know that the weighings are normally distributed about the true value w, with known variance σ^2. The results of the weighings are $w_1, \ldots, w_n$. The mean of these numbers is $\bar{w}$; the median w_m. The true value is w. We know both the distribution of $\bar{w} - w$ and the distribution of $w_m - w$. As Seidenfeld points out, the frequency with which $|\bar{w} - w|$ is less than ϵ is not the same as the frequency with which $|w_m - w|$ falls in a corresponding Borel set—i.e., the interval $(w_m - \bar{w} - \epsilon, w_m - \bar{w} + \epsilon)$. The rules concerning randomness thus prevent the ordinary inference from going through. This demonstrates clearly a deficiency in

---

[10] Seidenfeld nowhere denies this, but his discussion focuses so exclusively on "representativeness" that one might form the impression that this is the only statistical property for which *EP* provides probabilities.

those rules. A modification (and simplification) of those rules has been developed[11] which does allow the classical inference in this case and which furthermore allows an inference based on the median w_m when we do *not* know the value of the mean $\bar{w}$.

It is not quite clear, incidentally, whether Seidenfeld would want to accept an inference concerning w when all that is known of the weighings is w_m. The statistic w_m is *not* statistically sufficient for inferences about w; if we require ordinary statistical sufficiency, an inference that is intuitively perfectly reasonable is barred.

The second example concerns the Behrens-Fisher problem. As Seidenfeld indicates, the solution to this problem is controversial within statistics, and thus I am not bothered by the fact that Fisher's solution turns out to be epistemologically invalid. I can, however, explain what goes wrong. Suppose that ζ is a nuisance parameter, and that data d enable us to make inferences about ζ. Under certain circumstances, we may be able to assign a precise probability to all statements of the form $\zeta < z$. This *seems* like knowing the distribution of ζ in a certain population, and Fisher (sometimes) took it this way, though it is difficult to see what the population would be. If there were such a population and if the distribution of ζ in that population were known, there would naturally be no difficulty in "integrating out" ζ for the purpose of making an inference about ϕ which depends on ζ. But epistemological probability statements, even if there are a lot of them, do not constitute a frequency distribution over which we can integrate.

The third example concerns the combination of data of two kinds. We can make an inverse inference about H_ϕ given d; we can make an inverse inference about H_ϕ given d'. Seidenfeld directs our attention to the case in which "d and d' represent different kinds of information which *cannot* be combined into a single report for one inverse inference about ϕ" (720). Why can't we combine the two reports into one? Just staple them together, and renumber the pages, for example? Suppose that d represents the mean of a set of weighings, and d' the median of a different set of weighings. That they "cannot be combined" just means that there is no standard statistical technique for dealing with d and d' simultaneously. This is a crucial problem—particularly in the theory of measurement— for which I have no solution. Subjectivist Bayesians have a solution, but I find that "solution" unacceptable for reasons I have detailed elsewhere. I think an objective solution is worth seeking.

In the fourth example, we are to consider "singular predictive

---

[11] "Randomness," forthcoming. This development owes much to Seidenfeld's acute and patient criticisms of earlier versions of the rules for randomness.

inference." This is a very important example, because here I think it is easiest to see how our intuitions are being misled, not only in this example, but in the second example. To fix our ideas, suppose that d and d' each represent the outcome of a set of tosses of a coin : d represents the outcome of a sample of 100 tosses we have collected, of which 46 have yielded heads; d' consists of a single toss. We are to give a probability that d' will consist of a head, relative to a corpus containing d. Suppose there is no information about coins in K. Then the probability that d' will consist of a head, relative to Kd, is the trivial probability $(0, 1)$ and would remain so, were d to consist of 10,000 tosses. If K does contain information about tosses of coins, say to the effect that half of them land heads, the probability that d' will consist of a head is $\frac{1}{2}$, and remains so, relative to Kd.

Is this inappropriate or counterintuitive? I don't think so. Seidenfeld invites us to compute the probability that d' will consist of a head by integrating over θ, the true chance that the coin will yield heads. But this is something we don't know, and I fail to see how we can attain something equivalent to knowledge by integrating over ignorance.[12] I say "equivalent to knowledge," since on my view to say that the probability that d' yields heads is (p, q) is to say that there is some general statistical *knowledge* in Kd on which to base this probability. As Kd has been described, there is no general statistical knowledge concerning coins in it. The only plausible probability for the statement that d' yields heads is the *fully* indeterminate $(0, 1)$.

The intuition that Seidenfeld's example asks us to account for is that statistical data should influence the probabilities we assign to further instances. He correctly points out that epistemological probability does not do this directly. My response is that it should not do this directly, and that the influence of data on the probabilities of future instances should be exercised only through the mediation of accepted statistical hypotheses. Thus in order to account for the intuition we must consider two levels of rational corpus and a rule of acceptance. If we do this, we can see how data influence our probabilities.

Suppose that K is the Ur-corpus, and contains no information other than d bearing on the behavior of coins. Suppose our rule of acceptance entails that we may accept into the corpus of level p the most informative statistical hypothesis about the behavior of coins which has a probability of at least p. Suppose that d concerns

---

[12] Note that this is just like the second example, where we "integrate" over the unknown nuisance parameter ζ.

not a hundred tosses, but ten thousand; the evidence will therefore support the acceptance of a strong statistical hypothesis at a high level. Let K' be the rational corpus of this level. K' will contain d, but will also contain a statistical hypothesis about the frequency of heads in tosses of this coin. The probability that d' will consist of a head, relative to the higher corpus K, is still what it was before: $(0, 1)$; but, relative to the lower-level corpus K', the probability that d' will consist of a head will be determined by the strongest statistical hypothesis acceptable at level p about the behavior of the coin in question. But there is still no summing over weighted products—nor should there be.

In general, I think, it is the case that we must *accept* some general (statistical) hypothesis H on the basis of data, before we can use the data to determine probabilities. (Perhaps, with Aristotle, I think all scientific knowledge is essentially general.) This entails, of course, that when the highest corpus that contains the data is r, it is only in some *lower*-level corpus that we can accept a generalization based on the data and, thus, only in some lower-level corpus that these data can determine the probabilities we assign to individual instances. I suspect that much confusion about the use of probability and the role of data has stemmed from a failure to distinguish probabilities from levels of acceptance.

How should we tally up the scorecard, then? With regard to sufficiency, I claim that in one clear sense—a sense more closely related to the standard statistical notion than that employed by Seidenfeld—a natural principle of sufficiency *is* satisfied by epistemological probability. But the mean/median example seems to be a clear win for Seidenfeld in the sense that it points to a defect in the rules regarding randomness: sufficiency does not seem to operate the way it should in a standard situation.[13] Inverse inference with nuisance parameters is controversial within statistics, and many statisticians would agree that "integrating out a nuisance factor" is valid only under special circumstances—namely, when we know the distribution of that nuisance factor in a family of populations. But under these circumstances, integrating out will generally be valid for epistemological inverse inference as well. Inverse inference with data of different kinds represents a serious problem in statistics, and a particularly serious problem in the theory of measurement (in which we wish to be able to connect direct measurements of area with indirect measurements obtained through measurements

---

[13] But the data are not all in; Stephen Spielman claims to have discovered a way out in my original framework.

of length and geometric considerations). I do not find the Bayesian treatment of this problem acceptable, and, so far as I know, there is as yet no standard non-Bayesian treatment. Epistemological probability does not warrant the "singular predictive inference"— the example is correct. But I do not think the singular predictive inference can be warranted in a nonsubjective way, nor do I think it *should* be warranted.

IV

Levi regards confirmational conditionalization as sacrosanct (CC, 732). I regard direct inference as sacrosanct. One way to adjudicate such conflicting concerns is to see what consequences they lead to in particular cases. In this regard, Levi's examples are of little help. They show that *were* I to try to endorse confirmational conditionalization and construe epistemological probability as a constraint on Levi's "credal states," I would be led to bizarre consequences. Since I don't try to endorse confirmational conditionalization, I am led to nothing more bizarre than the claim that, under the circumstances described, the probability that Petersen is a Protestant is 0.9, which seems to me eminently sensible. I do not find the examples compelling.

Levi also offers arguments in support of confirmational conditionalization. The arguments pertain to choices among gambles, and embody the useful distinction between "called-off" and "conditional" bets. But these arguments are peppered with provisos:

". . . insofar as new circumstances and considerations do not intervene to warrant reconsideration of the matter" (733)
". . . all other factors remain constant" (733)
". . . with everything else remaining constant" (734)
". . . mandated only when all other relevant factors other than changes in the corpus remain fixed" (734)

It is clear that although Levi regards confirmational conditionalization as sacrosanct, he does not regard confirmational commitment as sacrosanct. This is the point of his distinction between temporal conditionalization (which he rejects) and confirmational conditionalization. The former is mandated only when "all relevant factors other than a change in corpus remain fixed. Of course they need not remain fixed. Not all legitimate changes in temporal credal state are temporal credal conditionalizations or their inverses" (734). Now a commitment that can be altered so easily does not seem to me much of a commitment, but perhaps that is beside the point. Being a necessitarian, I am willing to claim that there is only one rational confirmational commitment to have. I therefore hold this

confirmational commitment sacrosanct, and give up confirmational conditionalization when it conflicts with it.

The truth of the matter is not that Levi regards confirmational conditionalization as sacrosanct (since we can change its *import* by changing our confirmational commitment) and that I blaspheme against it, but that I want to save a simple notion of direct inference for the sake of a generally applicable standard of rationality, whereas Levi wants to save the universal applicability of the betting model, even if it means changing confirmational commitments in ways that (at least so far) seem to admit of no rational reconstruction.

The heart of our disagreement does not lie in any of these technical considerations, however. It is deeper, more interesting, and more philosophical. It emerges explicitly when Levi writes: "My own view is that the coffin of empiricism is already sealed tight" (737). This is obviously not my view. But in disputes between pragmatists and empiricists, the pragmatist does have an advantage; anything the empiricist can succeed in establishing can be regarded as an ingredient of pragmatism—all pragmatism requires is a body of shared agreement—and anything that the empiricist has failed to establish to the satisfaction of the pragmatist can be regarded as artificial and wrong-headed. The pragmatist can feel vindicated when he can point to a problem that the empiricist hasn't solved; the empiricist can achieve vindication only by solving all the problems.

Nevertheless, I remain an empiricist. Russell said somewhere that postulating the existence of entities of a certain sort had the same advantages over constructing them logically as theft has over honest toil. The advantages of postulating shared agreement about matters of empirical fact over constructing shared agreement on the basis of empirical data are the same.

<div style="text-align: right">HENRY E. KYBURG, JR.</div>

University of Rochester

NOTES AND NEWS

The National Endowment for the Humanities announces its program of Summer Seminars for College Teachers for 1980; twelve college teachers will be selected to attend each of 120 eight-week seminars; participants will receive a stipend of $2,500 to cover travel expenses to and from the seminar location, books and other research expenses, and living expenses. The purpose of the program is to provide opportunities for faculty at