Objective Evidence and Absence: Comment on Sober

Michael Strevens

November 2008

Abstract

Elliott Sober argues that the statistical slogan "Absence of evidence is not evidence of absence" cannot be taken literally: it must be interpreted charitably as claiming that the absence of evidence is (typically) not very much evidence of absence. I offer an alternative interpretation, on which the slogan claims that absence of evidence is (typically) not *objective* evidence of absence. I sketch a definition of objective evidence, founded in the notion of an epistemically objective likelihood, and I show that in Sober's paradigm case, the slogan can, on this understanding, be sustained.

1. Confirmation and Couture

Statisticians wear T-shirts advancing the following thesis:

Absence of evidence is not evidence of absence.

What proposition is expressed by this fashion statement? Sober (2009) advances one suggestion; in this commentary, I will fabricate an alternative. Science, I will propose, makes a distinction between merely having reasons to believe a hypothesis, on the one hand, and having objective evidence for that hypothesis, on the other; the latter but not the former is founded on certain objective epistemic probabilities. The T-shirt should be read, I submit, as proclaiming that absence of evidence is not *objective evidence* of absence. This paper is

1

concerned not so much to argue against Sober's interpretation as to investigate this other interesting possibility.

I do not have the space, unfortunately, to discuss the other theme of Sober's rich paper, the fine-tuning argument. Even then, my remarks will be confined to Sober's central example, the bearing of the discovery or otherwise of fossilized intermediate forms on a hypothesis of common ancestry; I hope that the reader will see the way to a fruitful generalization. Finally, I will follow Sober in supposing that the law of likelihood lies at the heart of scientific epistemology.

* * *

The hypothesis on the table is that two species have a common ancestor. You set up an experiment to test this hypothesis: you send off an expedition to look for fossil intermediates, that is, fossils of a form intermediate in character between the two species. The expedition has now returned having completed its digging plan; shortly, its leader will report to you whether or not fossil intermediates were found. Let me suppose that the leader delivers one of two possible reports: either fossil remains of an intermediate were discovered, or no such remains were discovered. The event of the report's turning out the former way is e; the event of its turning out the latter way—of there being an "absence of evidence"—is therefore $\neg e$. (My e, then, is Sober's $O(e)$.)

I have set things up in this way to avoid two complications. First, there is no prospect of finding fossils of more than one intermediate form. Second, the only possible sense in which there can be an absence of evidence is the sense in which an otherwise successful dig fails to turn up a fossil intermediate. I therefore ignore two other ways in which you might have an absence of evidence: you might have not heard from the expedition (perhaps it has not yet returned), or you might have heard that the expedition failed to realize its digging plans (perhaps the truck broke down on the way to the fossil fields). It is surely uncontroversial that these latter two kinds of non-evidence do not constitute evidence of absence; I follow Sober, then, in supposing that they are not the kinds of absence of evidence concerning which the T-shirt makes its claim—if they were, it would be an insubstantial garment, of no interest either to philosophers or to creationists.

2

As Sober explains, according to the law of likelihood *e* confirms the common ancestry hypothesis relative to its negation—the hypothesis that ancestry is in some way or other separate—because the probability of finding a fossil intermediate is greater on the common ancestry hypothesis than on the separate ancestry hypothesis. (The foundation of this difference in likelihoods will be the topic of section 4; for now, I refer you to Sober.)

2.   The Sober Interpretation

According to Sober, the T-shirt says something like this:

> *Absence of evidence is (in many interesting and salient cases) only very weak evidence of absence.*

This is a striking interpretation: it attributes to the T-shirt an underlying meaning that contradicts its apparent meaning. Who knew that the Straussians had infiltrated statistics? In any case, the T-shirt's wearers come out as looking rather disingenuous—duplicitous, even. If Sober's understanding of their message is correct, why do they not say explicitly that failure to find (some) fossil intermediates is only very weak evidence against common ancestry? It may not sell many T-shirts (as Sober's concluding paragraph remarks), but neither does it promulgate a doctrinaire line that is by statisticians' own lights in fact a falsehood, threatening to hand creationists a rhetorical victory with which to sweeten, and perhaps even to mask entirely, their epistemological defeat.

There is some reason, then, to look for an alternative interpretation of the T-shirt. As Sober shows, however, if you subscribe to the law of likelihood (as Bayesians and many others do), there is a wonderfully simple argument that makes it impossible for you to wear the T-shirt literally—an argument that establishes beyond question that absence of evidence is to some degree at least evidence of absence. The argument turns on a simple mathematical truth:

$$P(e|h) > P(e|\neg h) \quad \text{iff} \quad P(\neg e|h) < P(\neg e|\neg h)$$

You might call this the "risk principle": any experiment that has a chance of confirming a hypothesis must also have a chance of disconfirming it. (It surely

deserves a T-shirt of its own.)

Are scientists and statisticians then compelled either to reject the law of likelihood or to adopt a transformative hermeneutics of fashion, on which the real meaning of their T-shirts denies the apparent meaning? Perhaps not: the T-shirt might say something weaker than, but not contrary to, what it seems to say.

3.   Objective Evidence

I will distinguish two kinds of evidence, objective and subjective evidence, and I will propose the following interpretation of the T-shirt:

Absence of evidence is (typically) not objective evidence of absence.

In other words, to the extent that absence of evidence is, as Sober has shown, evidence of absence, it is *subjective* evidence of absence. The implicit assumption is, of course, that objective evidence is the really good stuff. Subjective evidence is an inferior epistemic fuel; indeed, it is questionable enough that science should ignore it—officially, at least—altogether. In a scientific context, then, having only subjective evidence is like having no evidence at all. Thus the elision of the term *objective* is not unwarranted; worn scientifically, the T-shirt pretty much means what it says.

What is the difference between objective and subjective evidence? Let me give you an answer within the framework of the law of likelihood, the principle according to which the relative impact of evidence e on two hypotheses h_1 and h_2 is entirely determined by the likelihoods $P(e|h_1)$ and $P(e|h_2)$.

The objectivity of a piece of evidence is, I stipulate, proportional to the objectivity of the relevant likelihoods (or better, proportional to the objectivity of their ratio; see section 4). By objectivity I mean *epistemic* objectivity; an objective likelihood may be, but does not have to be, a physical probability (although confusingly, one name for physical probability is "objective probability"). It may help to say right away that epistemic objectivity comes in degrees. So the question of the objectivity of evidence concerns a spectrum rather than

a strict dichotomy; the problem with absence of evidence will turn out to be that, as evidence of absence, it is *not objective enough*.

What, then, makes for objectivity or subjectivity in a likelihood? Since I am attributing an interest in this brand of objectivity to a statistical T-shirt, there is some value to starting with the known ideology of statisticians. Consider in particular the reasons that classical statisticians give for rejecting Bayesianism. There is an official argument against Bayesianism that I want to ignore, namely, that no sense can be given to the probabilities that Bayesians assign to hypotheses. This is a paradigm of philosophical overreach, of making a grand metaphysical claim to establish a conclusion that is believed for more modest, in this case methodological, reasons. The problem with probabilities for hypotheses is not that they do not exist. Quite the contrary: it is much easier to understand the basis of subjective probabilities or degrees of belief than to understand the basis of the probabilities assigned by, say, statistical mechanics. What bothers classical statisticians and other anti-Bayesians—including Sober 2008, an excellent source for this point of view—is that these probabilities are not epistemically objective. In the first instance, this means simply that there is no agreement about their values. (This is hardly a sophisticated characterization of objectivity, but it is adequate for my purposes in this paper.) As a consequence, by Bayesian lights, there is no single conclusion that all scientists should agree is warranted by the outcome of a given experiment.

The objection can be targeted more finely, as follows. According to the Bayesian methodology, for any particular experimental outcome various calculations yield what you might call a *Bayesian multiplier* for each hypothesis on which the experiment bears. You take this Bayesian multiplier and apply it to your prior probability for the hypothesis: if the multiplier is 1.2, for example, you multiply your probability for the hypothesis by 1.2. The experiment, in this case, confirms the hypothesis, increasing its probability by 20%.

The subjectivity of Bayesianism consists in the subjectivity of this multiplier: for the very same hypothesis and experiment, it may take different values for different investigators. Most dramatically, it might be less than one for one investigator and greater than one for another, meaning that the evidence confirms the hypothesis for one investigator and disconfirms it for the other.

How can this be? The multiplier for a hypothesis h on evidence e is

$$\frac{P(e|h)}{P(e)}.$$

Thus, e confirms h if the likelihood $P(e|h)$ is greater than $P(e)$ and disconfirms it if the likelihood is less than $P(e)$. The likelihood is usually understood to be objective—certainly it is in those cases where the hypothesis h explicitly assigns a physical probability to the evidence.[1] The subjectivity comes from the variability of $P(e)$: different investigators may assign different values to this probability, and so two scientists who agree on the likelihood may nevertheless disagree on the qualitative significance of the evidence, since one but not the other assigns a value for $P(e)$ that is greater than the likelihood.

Why is the value of $P(e)$ subjective? The theorem of total probability provides the neatest explanation:

$$P(e) = P(e|h_1)P(h_1) + P(e|h_2)P(h_2) + \cdots$$

where $h_1, h_2, \ldots$ are a set of competing hypotheses. (Technically, they must form a mutually exclusive, exhaustive set—a partition of the possibilities.) The likelihoods $P(e|h_i)$, I have assumed, are objective: anyone who understands the content of one of these hypotheses agrees on the probability that it assigns to the evidence. But the other probabilities that appear in the formulation, the $P(h_i)$s, vary according to the scientist. One scientist may think that a certain hypothesis that is very favorable to e is likely to be true, while another may find that same hypothesis quite implausible; ceteris paribus, the former investigator will have a higher subjective probability for e itself.

This, then, is why classical statisticians and their philosophical fellow-travelers find Bayesianism abhorrent: it is a methodology according to which the evidence does not speak univocally, that is, according to which the same piece of evidence may have different significance for different scientists. I will not try to say any more here about why evidence ought to speak univocally.

---

1. This in virtue of the principle of rationality dictating that in normal circumstances, conditional subjective probabilities should be set equal to corresponding physical probabilities (Lewis 1980).

Bayesians have replies to the classical statisticians' concerns with objectivity, most notably in the form of their convergence results, and yet the convergence results make some rather strong assumptions, which may not apply in many real scientific situations. It is a complex issue. But it is the statisticians who wear the T-shirts, and they are steeped in the classical ideology. So I think it is reasonable to put aside the question whether statisticians are right to place such high value on objectivity in the interpretation of evidence, and simply take this attitude as part of the sociocultural context in which the T-shirts make their claim.

One way to get around the subjectivity problem while retaining some of the flexibility of Bayesianism is to follow Sober in becoming a likelihood theorist. Since on this view, the only probabilities that are in play in determining the significance of evidence, both qualitative and quantitative, are likelihoods, and the likelihoods are typically treated as the objective pole star in the otherwise ever-shifting firmament of Bayesian probabilities, you might find this move irresistible. But—and here we are closing in on the T-shirt—likelihoods are not necessarily objective.

Consider, for example, the likelihoods that answer the following question: "To what degree does failing to find an intermediate fossil count against the existence of the corresponding intermediate form?". (Note that I am temporarily putting aside the deeper question of common ancestry.) There are two such likelihoods: the probability of failing to find the fossil conditional on there being such an intermediate form, and the probability of failing to find the fossil conditional on there *not* being such an intermediate form. In symbols, if, as above, e is the event of finding a fossil and h posits the intermediate form, the important likelihoods are $P(\neg e | h)$ and $P(\neg e | \neg h)$.[2] The degree to which not finding the fossil counts, for the likelihood theorist, against the existence of the intermediate form is determined by the ratio of these two likelihoods.

The value of $P(\neg e | \neg h)$ is, we can all agree, more or less one. No intermediate form, no fossil. But what about $P(\neg e | h)$? Surely it is greater than zero. After

---

2. All such probabilities should be regarded as implicitly conditioned on the dig's having been completed as planned, for the reasons given in section 1.

that, however, it is hard to say. There are many factors that enter into the determination of such a probability. First, there are a number of different ways that specimens of the intermediate form might have been fossilized, and a number of different ways in which each of these causal routes to fossilization might have fallen through. Second, there are a number of ways that a fossil might come to be, or might fail to come to be, in a place where your expedition will find it. The likelihood $P(\neg e|h)$ is a function of the probability distributions over every one of these factors. Its value will depend, for example, on probability distributions over various hypotheses about the way the putative intermediate form lived: its environment, its habits, its predators, and so on. Such probabilities are exactly the kind of Bayesian priors that give Bayesian methodology its subjectivity. For the same reason that a probability such as $P(e)$ is subjective, then, the likelihood $P(\neg e|h)$ is subjective, and so any methodology whose epistemic prescriptions hinge on the likelihood is itself possessed of the sort of subjectivity objectionable to the classical statistical mindset.

The same point can be made about many other likelihoods, drawing on the commonplace associated with the names of Duhem and Quine. A hypothesis that we are interested in testing will seldom, if ever, assign a probability to a piece of evidence single-handedly. Rather, it does so in consultation with one or more auxiliary hypotheses, some perhaps theoretical and some concerning relevant initial conditions. Unless the truth of the auxiliaries is secured in advance, there will be rival auxiliaries to take into account. But in taking them into account, their relative plausibility will also have to be taken into account. In Bayesian terms (again invoking the theorem of total probability):[3]

$$P(e|h) = P(e|ha_1)P(a_1) + P(e|ha_2)P(a_2) + \cdots$$

for the various possible auxiliaries a_1, a_2, and so on (as before, a partition of the possibilities). The prior probabilities of the auxiliaries $P(a_i)$ introduce the same subjectivity to the likelihood that the prior probabilities of the main hypotheses

3. Here I assume for simplicity's sake that the hypothesis is independent of the auxiliaries, so that $P(a_i|h) = P(a_i)$. This is in fact a rather tendentious assumption (Strevens 2001, note 7); however, its dubious status does no harm to my argument here.

8

introduced to the probability of the evidence in the Bayesian case, a subjectivity that undermines the experimental outcome ¬*e*'s status as objective evidence.

As you can see, I am attributing to the statisticians a certain objectivist scientific methodology: all probabilities that play a role in statistical inference, whatever their nature, must have an objective foundation. You can espouse this methodology without denying altogether the epistemic significance of subjective evidence, that is, of evidence that bears on your hypotheses by way of partly subjective probabilities. What you might say is this: some epistemic weight ought to be given to subjective evidence. As its name implies, it too is a kind of evidence. But the weight is not of the right kind for subjective evidence to play a role in science, given the importance of certain kinds of consensus in the scientific process. *Scientific* evidence is objective evidence. A scientist who does not have any objective evidence in a very real sense does not have any evidence.

I propose, then, that the T-shirt has the following to say about the fossils: if you find an intermediate fossil, that is objective evidence for the existence of the corresponding intermediate form; if you fail to find one, that is some subjective reason to believe that there is no intermediate form, but it is not objective evidence—it is not scientific evidence—against the form's existence. Absence of evidence for intermediate forms is not scientific evidence for their absence. A fortiori, it is not scientific evidence against common ancestry.

Consider next two objections to this statistical semiotics.

4.   The End of Objectivity?

I argued that failure to discover a fossil was not objective evidence against the existence of the corresponding intermediate form because the likelihood $P(\neg e|h)$ had no objective value. But the same argument applies to $P(e|h)$: its value, too, depends on the probabilities of various possible routes to fossilization and then discovery, probabilities to which different scientists assign different values. The objectivity of this latter likelihood is instrumental to determining the fossil's status as objective evidence for the existence of the form; if my

argument goes through, then, fossil hunting is apparently unable to provide any scientific evidence for or against intermediate forms![4]

Let me consider both a special and a general strategy for meeting this objection.

The Special Strategy Because of my choice of h (asserting the existence of an intermediate form rather than common ancestry), the likelihood $P(e|\neg h)$ is zero: as remarked earlier, if there was no intermediate form, you will not find a fossil of such a form. While the likelihood $P(e|h)$ is subjective, everyone will agree that it is greater than zero. The degree to which h is confirmed by e is proportional to the ratio of these two likelihoods, with the zero likelihood as the denominator. Thus, regardless of the value of the non-zero likelihood, everyone will agree that h is massively, in principle even conclusively, confirmed by the discovery of the fossil. There is, then, complete and well-founded agreement about the evidential significance of e for h; for this reason, e is objective evidence for h.

Why does the same argument not work for $\neg e$? Presumably everyone will agree that the likelihood of $\neg h$ on $\neg e$ is higher than the corresponding likelihood of h—you are more likely not to turn up a fossil of an intermediate form if the form never existed. So it is agreed that $\neg e$—the absence of evidence—is evidence against h to some degree. But there is no objectivity to the degree of confirmation, which depends on the ratio between the likelihoods, a ratio that is subjective because both likelihoods have non-zero values and at least one has a subjective value. Without some agreement on the degree of confirmation, however, agreement on the direction of confirmation is worthless. After a search has failed to turn up fossil intermediates, what should we conclude? That the intermediates probably never existed? Or that we have not looked nearly hard enough? Without the magnitude, there is no way to choose between these two options and never will be, thus no real progress can be made on the question at hand, whether there was an intermediate form. Or at least, no progress can be made that has an *objective* foundation.

---

4. Equally, you might note that as a matter of mathematical truth $P(\neg e|h) = 1 - P(e|h)$; one of these two likelihoods cannot, then, be more objective than the other.

Qualitative objectivity alone is not enough, then; science also needs quantitative objectivity. In the special case where the likelihood of ¬*h* on *e* is zero, you can have that objectivity for *e* while lacking it for ¬*e*.

The General Strategy   Suppose now that *h* is the hypothesis of common ancestry, and so that the likelihood $P(e|¬h)$ is greater than zero (since the separate ancestry of two species does not rule out the possibility that one species was in the past more similar than it is now to the other's present-day form). The special strategy will not be applicable, since the two likelihoods that determine the degree to which *e* confirms *h* are both non-zero and subjective.

That the likelihoods are subjective, however, does not entail that their ratio is subjective. Let me factor the likelihood of *h* into several distinct parts:

$$P(e|h) = P(dis.fos.ex|h)$$

where *ex* is the event of this particular intermediate form's existing at some stage in the past, *fos* is the event that a specimen of such a form is fossilized, and *dis* is the event that such a fossil is discovered by the expedition in question.

The likelihood ratio that determines the degree to which *e* confirms *h* relative to ¬*h* can now be written as follows:

$$\frac{P(dis.fos.ex|h)}{P(dis.fos.ex|¬h)}$$

Suppose, more or less following Sober (§3), that *ex* screens off *h* from *dis* and *fos*, that is, that the probability of a fossil intermediate's turning up, given that the form once existed, is the same regardless of whether or not the common ancestry hypothesis is true.[5] Then the contributions of the probabilities of the *dis* and *fos* events to the ratio cancel out, giving you

$$\frac{P(ex|h)}{P(ex|¬h)}$$

---

5. This is not quite Sober's assumption; his concerns the probability of finding some fossil intermediate or other, whereas mine concerns the probability of finding a fossil representative of some particular intermediate form. The difference is significant for the justification of the assumption; see the end of this section.

as the ratio that determines the degree to which *e* confirms the common ancestry hypothesis. It was uncertainties in the probabilities of fossilization and discovery that created most of the variation in the likelihoods of *h* and ¬*h* on *e*. The canceling out results, then, in a likelihood ratio that is considerably less subjective than either of the likelihoods.

How subjective? The numerator in the ratio is equal to one. When the common ancestry hypothesis is tested relative to some particular theory of separate ancestry, rather than the simple denial of common ancestry, it is very plausible, I think, that the denominator will also have a relatively objective value. Thus the degree of confirmation will be objective, and the evidence will be objective evidence for common ancestry.

The same canceling out cannot be obtained when it is ¬*e*'s significance that must be ascertained; here the relevant ratio is

$$\frac{\sum \left(P(\neg dis.fos.ex \,|\, h) + P(\neg fos.ex \,|\, h) \right)}{\sum \left(P(\neg dis.fos.ex \,|\, \neg h) + P(\neg fos.ex \,|\, \neg h) + P(\neg ex \,|\, \neg h) \right)}$$

where the summation is over all possible intermediate forms. There is no serious simplification in sight. So in a cognitively diverse field, there will be little intersubjective agreement, little objectivity, to be found in ¬*e*'s evidential bearing on *h* or ¬*h*. You may have absence of evidence for, but you have no objective evidence against, the hypothesis of common ancestry.

A loose end: I supposed above that *ex* screens off *h* from *fos* and *dis*. Is this a reasonable assumption? The most obvious way that *h* might be relevant to *fos* and *dis*, conditional on the past existence of the fossilized intermediate form, is if greater numbers of an instantiated form are likely to have existed given common ancestry than given separate ancestry. But although greater numbers of intermediate forms (plural) are likely to have existed given common ancestry (see Sober's note 8), there is I think no reason to believe that greater numbers of any particular form are more likely on one hypothesis than on the other, given that the form did at one time exist. So the screening-off assumption holds roughly true.

5.   Heads I Win, Tails We Flip Again?

A complementary difficulty: on the objectivist story presented above, it seems that you can *never* have objective evidence against the existence of intermediate forms—or more accurately, that no matter how hard you look for an intermediate fossil, your failing to find one *never* constitutes objective evidence, thus never constitutes scientific evidence, against the existence of the corresponding form. The lesson generalizes: on the objectivist way of thinking, you find yourself with the Popperian conclusion that no amount of fruitless searching will count as scientific evidence against an existential claim (unless of course the search space is completely explored).

    Now would be a good time to remind you that I am here to interpret, not to defend, the T-shirt. The above consequence of objectivism is indeed, I think, a rather extreme doctrine. But now is also a good time to observe: *it is exactly what the T-shirt says*. That absence of evidence is not evidence of absence is not an abstruse consequence that I have extracted with great difficulty from a controversial interpretation of the classical statistical methodology. It is a slogan that statisticians shout to the heavens and sell online. Apparently, they really mean it. That my interpretation of their thinking has the consequence that, by their lights, they *ought* to mean it, is evidence for, not against, that interpretation.

    Here is the T-shirt again, in a slightly different form:

Absence of evidence is not scientific evidence of absence

where scientific evidence means objective evidence. What the T-shirt is saying, then, is that absence of evidence never provides the kind of objective handle on the relevant existential hypothesis that you would need to make an objective decision as to whether to accept or reject the hypothesis. The objectivist rationale for this view is not crazy. Indeed, I think that it deserves serious attention. It may, however, go a little too far.

References

Lewis, D. (1980). A subjectivist's guide to objective chance. In R. C. Jeffrey (ed.), *Studies in Inductive Logic and Probability*, volume 2. University of California Press, Berkeley, CA.

Sober, E. (2008). *Evidence and Evolution: The Logic behind the Science*. Cambridge University Press, Cambridge.

———. (2009). Absence of evidence and evidence of absence: Evidential transitivity in connection with fossils, fishing, fine-tuning, and firing squads. *Philosophical Studies* 144.

Strevens, M. (2001). The Bayesian treatment of auxiliary hypotheses. *British Journal for the Philosophy of Science* 52:515–538.