Reichenbach on simplicity and Reichenbach's straight rule

Elliott Sober [February 9, 2011]

Why should simplicity be used to choose between two curves that fit the available data equally well? Reichenbach (1937) defends this policy by arguing that a certain procedure will converge on the true curve (if there is a true curve) as the data set is made larger and larger. The procedure he describes is: connect the data points with straight lines and then smooth a little. There is a problem here. Even if the claim about this method's converging on the truth is correct, that doesn't solve the problem as stated, which involves a finite data set and the question of which curve will be more predictively accurate when new data are obtained.

Notice, first of all, that "connect the dots and then smooth" is not what friends of simplicity do in curve-fitting problems. They often prefer simple smooth curves that fail to pass exactly through all the data points over more complex curves that fit the data better. So the method that Reichenbach describes is not the method used when simplicity and fit-to-data are both taken into account to decide which curves are better and which are worse.

But there is an additional problem with Reichenbach's argument: *If one method will converge on the truth in the infinite limit, there are others that will do so as well, and these methods will often disagree among themselves about what to say about finite data.*

There is a parallel point about the rule of induction that Reichenbach called the straight rule:

(SR) If you observe m apples and n of them are green, you should infer that n/m of all apples are green.

This rule will converge on the truth as m is made larger and larger. What does this fact about the asymptotic limit show about what you should think when you look at 100 apples and 55 are green? Nothing, for consider the following unstraight rule:

(UR) If you observe m apples and n of them are green, you should infer that (n/m)+c of all apples are green (where c is a corrective factor that shrinks to zero as m gets bigger).

Notice that UR will converge on the truth if SR will. But UR and SR give very different advice about what to say about finite data sets – e.g., the one described above (Salmon 1966, pp. 85-89).

Comments: UR isn't a single rule. It is a family of rules, the members of which differ in terms of the specifics they provide in characterizing the corrective factor c. Also, SR and UR should be considered as applying to the process of gathering data where there is sampling with replacement (so that there is no finite bound on the number of observations you can make).

An exercise: Describe a policy in curve-fitting, different from Reichenbach's policy of connect-the-dots-and-then-smooth, which will converge on the truth if Reichenbach's policy will, but which disagrees with Reichenbach's policy in terms of what one should say about which curves are best when there is a finite data set.

Question: Ethicists contrast *act* and *rule* utilitarianism. How does the above point about induction and simplicity connect with that contrast?

References

Hans Reichenbach, *Experience and Prediction*, University of Chicago Press, 1937.

Wesley Salmon, *Foundations of Scientific Inference*, University of Pittsburgh Press, 1966.