A note on confidence intervals

Elliott Sober

If you construct a 95% confidence interval, drawing your data from a population in which the average height is x, then the probability is 0.95 that the confidence interval you construct will include (will "cover") the true value x. Above are two examples of this fact. In figure 1, the true population average is 6; in figure 2, the population average is 4. In both figures, the interval [5,7] is shown in red. In Figure 1, [5,7] includes the true value 6; in Figure 2, the interval[5,7] does not include the true value 4.

Suppose you drew data from a population and didn't know what the true value is of the average height. You construct [5,7] – it's a 95% confidence interval. What you can infer is that

(1) Pr(your observations sanction a [5,7] confidence interval | population mean is 6) > Pr(your observations sanction a [5,7] confidence interval | population mean is 4).

But (1) does not entail that

(2) Pr(population mean is 6 | your observations sanction a [5,7] confidence interval) > Pr(population mean is 4 | your observations sanction a [5,7] confidence interval).

This is why it is a mistake to interpret a 95% confidence interval as having a probability of 0.95 of including the true value.

This is like the gremlin example. Let E = you hear noise coming from the attic. Let H = there are gremlins bowling in the attic. The noise allows you to infer that

(3) $\Pr(E \mid H) > \Pr(E \mid \text{not}H)$.

But (3) does not entail that

(4) $\Pr(H \mid E) > \Pr(\text{not}H \mid E)$.

The reason (1) does not entail (2) is the same as the reason that (3) does not entail (4). You can't derive facts about the posterior probabilities of hypotheses just from facts about their likelihoods; you need, in addition, information about their prior probabilities. Confidence intervals are part of frequentism, which does not talk about prior probabilities. This is why (2) is no part of the confidence interval framework.

Figure 1: Each line represents a 95% confidence interval for a different data set drawn from the same population in which the mean height is 6. The red line is the interval [5,7].
.

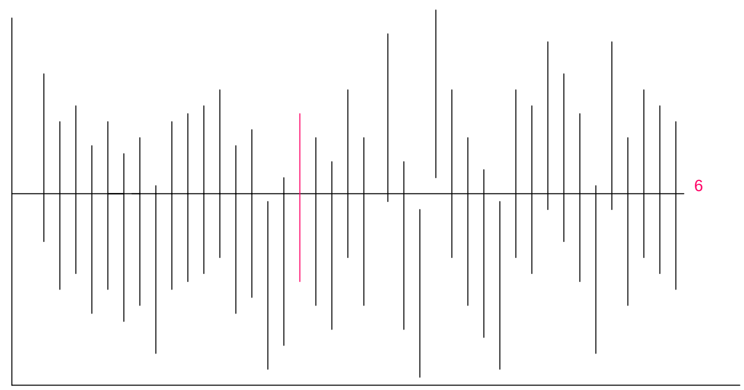

Figure 2. Each line represents a 95% confidence interval for a different data set drawn from the same population in which the mean height is 4. The red line is the interval [5,7].