INFERENCE

More Cautions About Interpreting Confidence Intervals

So What *Should* We Say?
Since 90% of random samples yield an interval that captures the true mean, we *should* say, "I am 90% confident that the interval from 29.5 to 32.5 mph contains the mean speed of all the vehicles on Triphammer Road." It's also okay to say something less formal: "I am 90% confident that the average speed of all vehicles on Triphammer Road is between 29.5 and 32.5 mph." Remember: *Our uncertainty is about the interval, not the true mean.* The interval varies randomly. The true mean speed is neither variable nor random—just unknown.

Confidence intervals for means offer new tempting wrong interpretations. Here are some things you *shouldn't* say:

- ***Don't say,*** *"90% of all the vehicles* on Triphammer Road drive at a speed between 29.5 and 32.5 mph." The confidence interval is about the *mean* speed, not about the speeds of *individual* vehicles.
- ***Don't say,*** "We are 90% confident that *a randomly selected vehicle* will have a speed between 29.5 and 32.5 mph." This false interpretation is also about individual vehicles rather than about the *mean* of the speeds. We are 90% confident that the *mean* speed of all vehicles on Triphammer Road is between 29.5 and 32.5 mph.
- ***Don't say,*** "The mean speed of the vehicles is 31.0 mph 90% *of the time.*" That's about means, but still wrong. It implies that the true mean varies, when in fact it is the confidence interval that would have been different had we gotten a different sample.
- Finally, ***don't say,*** "90% *of all samples* will have mean speeds between 29.5 and 32.5 mph." That statement suggests that *this* interval somehow sets a standard for every other interval. In fact, this interval is no more (or less) likely to be correct than any other. You could say that 90% of all possible samples will produce intervals that actually do contain the true mean speed. (The problem is that, because we'll never know where the true mean speed really is, we can't know if our sample was one of those 90%.)
- ***Do say,*** "90% of intervals found in this way cover the true value." Or make it more personal and say, "I am 90% confident that the true mean speed is between 29.5 and 32.5 mph."

INTRO STATS - Deveaux, Velleman & Bock.

The Frequentist Definition of Probability

In 1872, John Venn, the British philosopher, had proposed a formulation of mathematical probability that would make sense in real life. He turned a major theorem of probability on its head. This is the law of large numbers, which says that if some event has a given probability (like throwing a single die and having it land with the six side up) and if we run identical trials over and over again, the proportion of times that event occurs will get closer and closer to the probability.

Venn said the probability associated with a given event is the long-run proportion of times the event occurs. In Venn's proposal, the mathematical theory of probability did not imply the law of large numbers; the law of large numbers implied probability. This is the frequentist definition of probability.

Probability Versus Confidence Level

JERZY NEYMAN 1894 – 1981

Neyman's procedure does not break down, regardless of how complicated the problem, which is one reason it is so widely used in statistical analyses. Neyman's real problem with confidence intervals was not the problem that Fisher anticipated. It was the problem that Bowley raised at the beginning of the discussion. What does probability mean in this context? In his answer, Neyman fell back on the frequentist definition of real-life probability. As he said here, and made clearer in a later paper on confidence intervals, the confidence interval has to be viewed not in terms of each conclusion but as a process. In the long run, the statistician who always computes 95 percent confidence intervals will find that the true value of the parameter lies within the computed interval 95 percent of the time. Note that, to Neyman, the probability associated with the confidence interval was not the probability that we are correct. It was the frequency of correct statements that a statistician who uses his method will make in the long run. It says nothing about how "accurate" the current estimate is.

THE LADY TASTING TEA – DAVID SALSBERG.