Identifying mistakes in Interpreting Confidence Intervals

Some variety in the language is acceptable, and some is not.

Part 1. Correct interpretations

Following are correct interpretations of confidence intervals from two different studies. (The information in these statements came from correctly designed and implemented studies.)

- I am 90% confident that the population proportion of lie detector tests that miss identifying one lie is between 0.23 and 0.46.
  or
  I am 90% sure that the population proportion of lie detector tests that miss identifying one lie is between 0.23 and 0.46.
- I am 90% confident that the average commute time for workers in Atlanta is between 17.2 and 19.2 minutes.

Part 2. Meaning of 90% confidence:

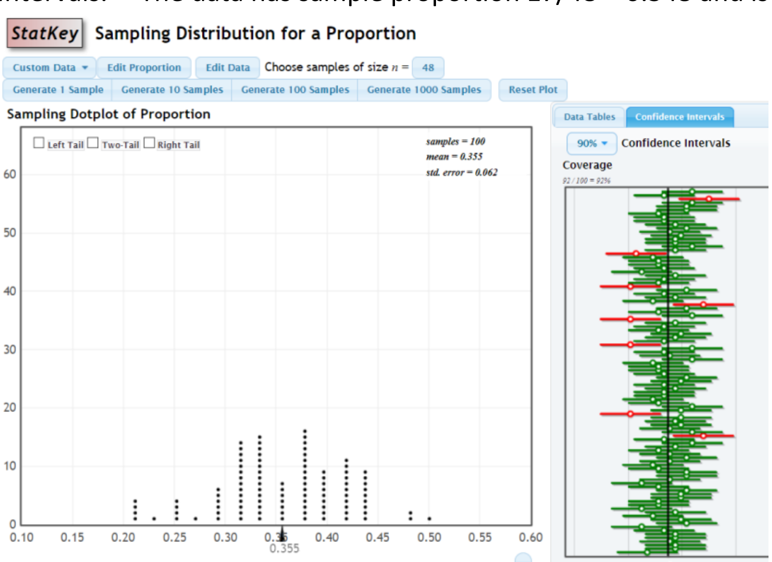
If

- we do collect another set of sample data from the same situation, and
- compute the sample proportion, and then
- compute a 90% confidence interval based on that data, and
- continue to do that for additional samples,

then

- on the average, we expect about 90% of those intervals to contain the true population proportion.

To remember this, it is very useful to visualize it. To reconstruct the lie detector example, use StatKey > Sampling Distribution > Proportion and use the setting in the upper left picture, called "Confidence Intervals." The data has sample proportion 17/48 = 0.348 and is from a sample of size 48.



Part 3: Typical mistakes that people make:

1. There is a 90% probability that the population proportion of lie detector tests that miss identifying one lie is between 0.23 and 0.46.
   Why is this a mistake?
   Answer: "probability" refers to something being random, and the population proportion isn't Prandom, but is a fixed, but unknown, quantity. What's random here is the endpoints. But, when we have formed the confidence interval, we use specific endpoints, so it's hard to remember that those are random. We have to remember the picture above to remind ourselves that the 100 intervals have different sets of random endpoints.
2. I am 90% confident that a worker in Atlanta has a typical commute time of between 17.2 and 19.2 minutes.
   Why is this a mistake?
   Answer: This statement is about individual sample values, not about the population mean.
3. I am 90% confident that all worker's commute times are between 17.2 and 19.2 minutes.
   Why is this a mistake?
   Answer: This statement is about individual sample values, not about the population mean.
4. I am 90% confident that most worker's commute times are between 17.2 and 19.2 minutes.
   Why is this a mistake?
   Answer: This statement is about individual sample values, not about the population mean.
5. I am 90% confident that worker's commute times are between 17.2 and 19.2 minutes.
   Why is this a mistake?
   Answer: This statement is about individual sample values, not about the population mean.
6. I am 90% confident that the sample mean of commute times is between 17.2 and 19.2 minutes.
   Why is this a mistake?
   Answer: This statement is about the sample mean, not about the population mean.
   (Also note that from our method of computing confidence intervals, we can be 100% confident that the sample mean is in our interval!)
7. I am confident that 90% of all worker's commute times are between 17.2 and 19.2 minutes.
   Why is this a mistake?
   Answer: This statement is about individual sample values, not about the population mean.

This last mistake is an attempt, I think, to actually describe what the 90% is 90% OF. And that is a good goal – but this is 90% OF the wrong thing! To correct it, we have to go back to Part 2, where we describe (and use simulation to investigate) WHAT we have 90% OF in our process.

Part 4. Conclusion

To remember the meaning of a confidence interval statement, think about the PROCESS, including what probability the 90% is describing.

Keep the visualization in mind to remember that it is the endpoints that are random in the interpretation of the confidence interval.

Part 5. More advanced idea. Bayesian statistics does give probability intervals

In Bayesian statistics, our parameters are considered random. Thus our interval estimates of the parameters are probability intervals. They are called "credible" intervals, instead of "confidence" intervals, to keep this distinction clear.