CI Explanation from an Anonymous Source – Do Not Quote

There is some confusion in the literature regarding the interpretation of a classical confidence interval. For example, Arnold (1990, p. 568) states that *nothing* can be said about a classical confidence interval that has been computed in one particular sample. This criticism assumes that the interpretation of a classical confidence interval must use the *relative frequency* definition of probability. The relative frequency definition of probability is defined as the proportion of trials in which some event occurs as the number of trials approaches infinity. This purely objective view of probability is typically used in statistics texts that cover the theoretical properties of classical statistical methods (see e.g., Hogg & Tanis, 2001). Although a computed confidence interval has no meaning in terms of the relatively frequency definition of probability, this notion of probability should not be used to interpret a computed confidence interval (Pearson, 1947). Instead, we can use a *subjective degree of belief* definition of probability to interpret a computed confidence interval.

Subjective probabilities can be criticized because two different experts can assign different probabilities to a particular outcome such as a sales forecast or a possible merger. However, experts can agree about the probability of certain events in highly simplified settings. For example, suppose a jar contains green and red marbles of equal size and weight and it is known that 95% of the marbles are green. Imagine also that a marble turns white after it is removed and its original color will never be known. If the marbles are thoroughly mixed and one marble, with eyes closed, is removed from the jar, most people would assign a subjective probability of .95 to the claim that the selected marble was green.

Bonett and Wright (2007) explain how this marble example is analogous to the interpretation of a confidence interval that has been computed in one sample. If a 95% confidence interval is computed from every possible sample of a given size, then we know from statistical theory that about 95% of these intervals will include the unknown population parameter value. This is analogous to knowing that 95% of the marbles are green. If a confidence interval is computed from one randomly selected sample, which is analogous to drawing one marble from a thoroughly mixed jar, then our subjective probability that the computed confidence interval includes the unknown population parameter value is .95.

*From Bonnet and Wright (2007, p. 28):*

Suppose a jar contains 1000 marbles in which 950 are blue and 50 are red. All marbles are the same size, weight and texture and have been thoroughly mixed. Close your eyes and remove one marble. Knowing that 95 per cent of the marbles are blue and that every marble has the same chance of being selected, with eyes still closed, you would form a degree of belief regarding the color of the marble in your hand. Your degree of belief that the marble in your hand is blue should be identical to your degree of belief that the 95 per cent confidence interval in the commitment example (0.5, 3.8) has captured the difference in the two subpopulation means. We now give some hypothetical examples to illustrate how hypothesis testing results can be replaced with confidence interval results for several elementary types of statistical analyses. In some examples, the hypothesis testing results are ‘‘non-significant’’ and in other examples the hypothesis testing results are ‘‘significant’’. In each example, the use of a confidence interval, when contrasted with the more traditional hypothesis testing approach, provides important additional information.