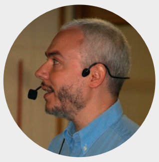
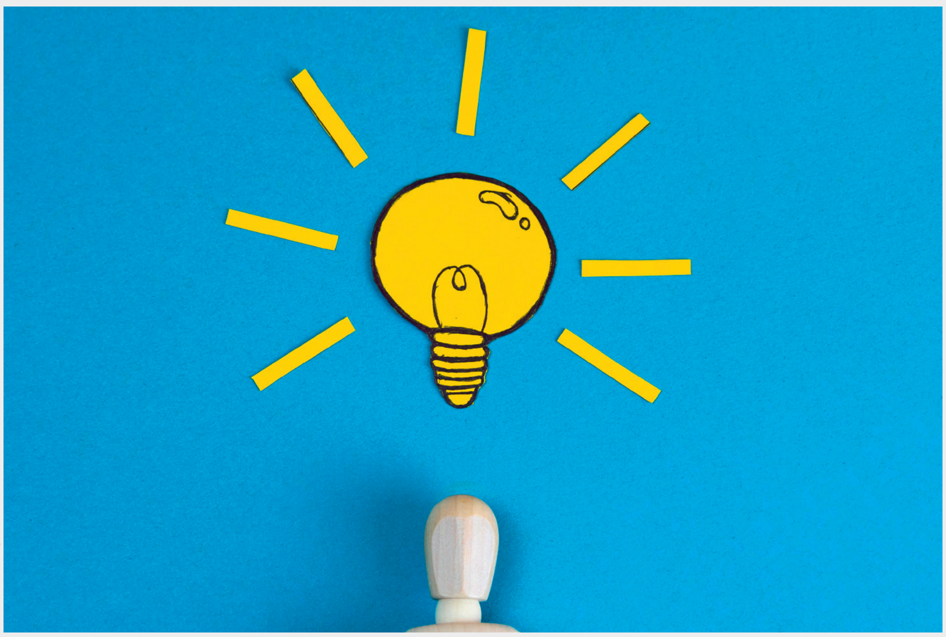


Angel Ricardo Plastino is professor at Universidad Nacional del Noroeste de la Prov. de Buenos Aires (UNNOBA), Argentina, and a member of Argentina's National Research Council (CONICET).

Angelo Plastino is emeritus professor at Universidad Nacional de La Plata, Argentina.

What's the big idea? Cramér–Rao inequality and Rao distance

Angel Ricardo Plastino and **Angelo Plastino** give a brief introduction to two key developments in statistics that originate with C. R. Rao's 1945 paper, "Information and accuracy attainable in the estimation of statistical parameters"



uspmen/Bigstock.com

In 1945, C. R. Rao broke ground with a manuscript that introduced inequality relations which set limits on the accuracy attainable in the estimation of the parameters characterising a probability density.[1] Rao also derived a distance between probability densities, introducing concepts from Riemannian geometry into statistics. These developments had implications beyond the field of statistics and are currently regarded as essential in the analysis of fundamental aspects of physics.

Rao's manuscript dealt with the problem of estimating statistical parameters from observations.[2] When applying statistical ideas to natural or social phenomena, it often happens that one knows, from previous empirical or theoretical considerations, the form of a probability density, but this density depends on parameters whose values are unknown and have to be estimated from observations. Rao discovered fundamental limits on the accuracy with which these parameters can be estimated.

Suppose, for example, that a biologist is investigating a population of animals. These animals have different weights, distributed according to a Gaussian distribution (also called a normal distribution). This distribution has a bell-like shape characterised by two parameters: its centre and its width. Let us assume that the width is known, and we have to estimate the centre, which corresponds to the mean weight of the animals. This quantity has to be estimated from a sample of the population. Rao discovered that there is a basic limitation on the accuracy with which our unknown parameter can be estimated. This limitation imposes a lower bound on the variance of the estimated parameter, given by the inverse of a quantity known as the Fisher information. This bound holds provided that the estimated parameter is computed from the sample data in an unbiased way.

The inequality relating the variance of the estimated parameter and the Fisher information is the celebrated Cramér–Rao inequality. This inequality is central to statistics because it sheds light on a fundamental problem that arises in any application of statistical methods, and because it constitutes a universal result largely independent of any specific details of the problem at hand.

In more general situations, a probability density depends on a set of n unknown parameters, $(\theta_1, ..., \theta_n)$. Rao derived limitations on the accuracy attainable in the estimation of each parameter. These studies motivated Rao to consider the problem of defining a proper distance between two probability densities. The probability densities studied by Rao, characterised by n parameters, can be regarded as points in an abstract n-dimensional space. To define the concept of distance in this space, Rao employed mathematical tools introduced by the mathematician Bernhard Riemann. Interestingly, these mathematical tools are closely related to the ones used by Einstein in the development of general relativity.

It often happens with ground-breaking developments that they admit important applications way beyond the ones originally intended. Rao's 1945 contributions are no exception. Fisher information, in conjunction with the Cramér–Rao inequality, and Rao's distance, are now important tools in the study of the foundations of both quantum mechanics and statistical physics. For instance, the Cramér–Rao inequality has been applied to the study of uncertainty principles in quantum mechanics, and to the analysis of the arrow of time in connection with irreversible processes. ■

References

1. Rao, C. R. (1945) Information and accuracy attainable in the estimation of statistical parameters. *Bulletin of the Calcutta Mathematical Society*, **37**(3), 81–91.

2. Amari, S. and Nagaoka, H. (2000) *Methods of Information Geometry*. Oxford: Oxford University Press.