**Researches about theory (R)**

1\_R. Give your best description of the many reaching out of statistics, in its various form, as a branch of math (Probability theory, etc.), as a set of methodologies used in many other disciplines, as an essential tool to deal with any sort of data, make reports and provide governance tools. Discuss whether it can be considered a "science" and what is the "scientific method" (what is a "theory" and what is a "hypothesis"). What is the role of Statistics in Math and Science?

Just to understand what we are talking about, let’s take two definitions:

“Statistics is the discipline that concerns the collection, organization, analysis, interpretation, and presentation of data in applying statistics to a scientific, industrial, or social problem, it is conventional to begin with a statistical population or a statistical model to be studied. Populations can be diverse groups of people or objects […]. Statistics deals with every aspect of data, including the planning of data collection in terms of the design of surveys and experiments.” (Wikipedia, <https://en.wikipedia.org/wiki/Statistics>)

“Statistics is a mathematical science pertaining to the collection, analysis, interpretation or explanation, and presentation of data.” (Science Daily, <https://www.sciencedaily.com/terms/statistics.htm>)

So, it seems obvious we are talking about data, information. As the Science Daily writes, Statistics is a mathematical science, but we can use it in many different fields: computer science, biology, physics… Anything involves data, so anything can be descripted and analyzed using Statistics.

Statistics is not so rarely confused with Probability, that is another branch of math. As prof. Steven Skiena from Stony Brook University says (<https://www3.cs.stonybrook.edu/~skiena/jaialai/excerpts/node12.html>), Probability and statistics are related areas of mathematics which concern themselves with analyzing the relative frequency of events. Still, there are fundamental differences in the way they see the world:

* Probability deals with predicting the likelihood of future events, while statistics involves the analysis of the frequency of past events.
* Probability is primarily a theoretical branch of mathematics, which studies the consequences of mathematical definitions. Statistics is primarily an applied branch of mathematics, which tries to make sense of observations in the real world.

Since Statistics is about “Data”, and it’s a branch of math, let’s define what the word data means. But before talking about data, let’s define ‘dataset’. The dataset is the object of the statistical analysis, and it is composed by a collection of points / information. We call this collection ‘Statistical population’ and each point can be referred as ‘Statistical unit’. Each statistical unit could have different information related to itself, for example if we consider a student we can list the weight, height, hair’s color, age and so on… Each single piece of information is what we refer as data. Be aware that the dataset is quite always represented as a table, having as rows (or columns) the statistical units and as columns (or rows) the qualitative or quantitative observations of each unit. Based on the structure of the dataset we talk about different aspects of statistics. If we consider only one attribute, we are talking about Univariate Statistics; if we consider multiple attributes (referred to statistical units) we talk about Multivariate statistics.

We have two types of statistics methodologies and tools: Inferential Statistics and Descriptive Statistics. The Descriptive one, as the name suggests, refer to all the methodologies used to describe a statistical population and its data. With the other, the inferential population, we refer to all the methodologies used to carry out some conclusions on a bigger population, not involved in the analysis. The important difference between the two is the fact that with the second we’re trying to describe a bigger population with an analysis on a part of the whole population.

As we can see in the Application part of the related Wikipedia page (<https://en.wikipedia.org/wiki/Statistics#Applications>) we can list some applications of Statistics, like: Statistical Computing, Business Statistics, Statistics applied to mathematics or the arts …etc..

A ’science’ is “the intellectual and practical activity encompassing the systematic study of the structure and behaviour of the physical and natural world through observation and experiment.” From the Oxford Languages.

So of course, we can consider Statistics as a science, that is composed by mathematic tools to analysis, comprehend and manage any kind of information.

The scientific method instead is “a method of procedure that has characterized natural science since the 17th century, consisting in systematic observation, measurement, and experiment, and the formulation, testing, and modification of hypotheses.”