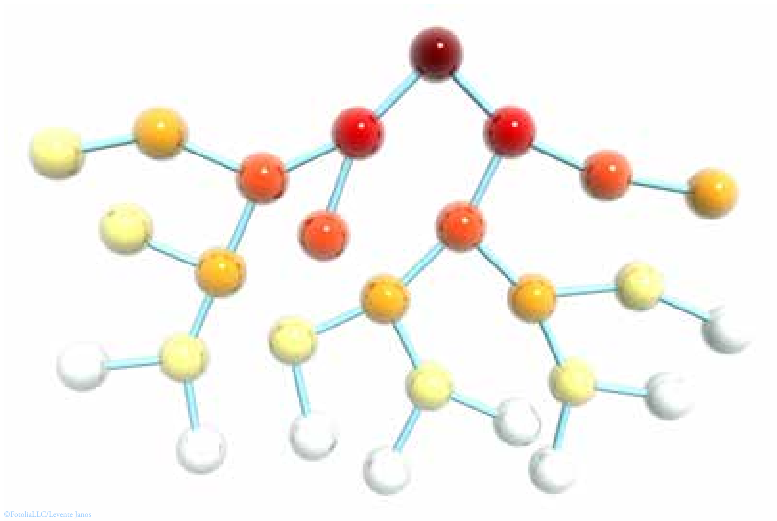
Five-minute Science Lesson: Bayesian and Frequentist Statistics – What's the Deal?

by Raymond Nelson



Statistics is the mathematical language of science. Science is the process of attempting to understand reality: how the universe works. It's a big universe, and so science tends to take things one little piece at a time, and then tries to put the pieces together. This is done with the expectation that if our knowledge about the little pieces of reality is correct then those little pieces should fit together harmoniously to explain more and more of the big universe. Still, it's a big universe and so we will most likely never know everything. As a result, science is a process of continuous learning – always adding new knowledge and information to old. Making any practical use of new knowledge requires a bit of humility – we are forced to admit that our human knowledge is at best incomplete and is possibly incorrect in some ways.

In order for us to regard our incomplete/incorrect human knowledge as trustworthy the principles of science require that the experiments we use to learn, acquire knowledge, and make conclusions must be reproducible. To accomplish that we often need to move beyond subjective observation, and to do this we have learned to ask scientific questions that can be answered objectively. Yes or no? Where? How much? Which is greater? Which is smaller? Or, what is the order? These questions often require the application of numerical quantification, mathematical analysis and logic to our data. As it happened historically, just about as soon as scientific thinkers and mathematicians began to record and calculate data with greater precision – following the renaissance, the introduction of algebra, calculus and geometry, and the first industrial revolution – they to notice unexpected variation when they compared reality to their expectations based on those precise calculations.

Frequentist statistics is the quantification of measurement error

The field of statistics originated out of a need to understand errors of measurement: measurement error. Today we have more generalized understanding of the concept of measurement error and random variation. But two-hundred and fifty years ago the great minds were still wondering this: why, with years of recorded information and precise observations and precise calculations from the best mathematical minds in history, did they observe that the planets were not exactly where their measurements and calculations suggested they would be? Is the universe fundamentally unstable? Is it all going to end in tragedy and chaos? These must have seemed like daunting and impossible questions to contend with. A more practical version of these questions is this: why is there variation in our measurement of the location of these things?

As it turns out, sometimes we have to quantify both the thing we want to quantify and the amount of variation in our data and calculations. The development of the principles of frequentist statistics allowed us to quantify observed variation in the data that is explained by factors that can be controlled or explained and the amount of variation that is uncontrolled or unexplained. In practical terms, frequentist inference allows us to better understand the location of the planets and other objects in the solar system. In broader terms it allows us to better understand the universe.

Today, frequentist inference is still very useful in all areas of science. For example: we may wish to learn about a population when it is not realistic to evaluate every member of the population. In this way, statistics is also about using knowledge from samples or small groups to make inferences about the population from which the sample was obtained.

These same principles can be used to make inferences at the level of the individual case. When this is done we have a scientific test as opposed to a scientific experiment. As a practical, consider this: if in reality there are any physiological activities that are correlated with the difference between deception and truth-telling, then perhaps we can use the principles of frequentist inference, along with some knowledge about observed data from deceptive and truthful groups, to understand the explained and unexplained variation in the individual case data. In this way we can calculate the statistical likelihood that the data for a new and unknown case was produced by a member of the population represented by the deceptive or truthful group. Of course if all physiology is mere random chaos with absolutely no correlation between any changes physiological activity and deception or truth-telling, then the premise of scientific credibility assessment and lie detection will achieve results that are no different than could be achieved by random chance alone.

Frequentist inference is based on the frequency of occurrence of things that can be observed. Frequentist inference is premised on the notion that reality and the universe exist in one way – reality is fixed. When we want to make a conclusion about something that cannot be easily observed, we use the principles of statistical inference to calculate the statistical likelihood of observing the obtained data if reality does not exists in the way we think it exists. Because it regards reality as fixed (i.e., the universe exists only one way) frequentist inference is concerned not with the probability associated with the universe, but with the probability of obtaining the observed data if our knowledge or conclusions about reality and the universe are correct.

In frequentist inference, because reality is constant, data are variable. A different set of data is expected to give slightly different information about the universe. Frequentist inference is concerned with the statistical variation in the data as an approximation of reality. But the universe is still regarded as existing only one way. For example: a person either is or is not pregnant.

Bayesian inference is the calculation of causes for the observed data or evidence

Bayesian inference – named for Thomas Bayes who together with Pierre Simon Laplace laid the foundation for area of statistical knowledge – was first conceived of as a means of as a means of calculating the most likely cause of our observations about reality and the universe. There is an old adage in science and statistics: "correlation is not causation," and Bayesian analysis is intended to help us to move beyond this apparent impasse and quantify the most likely cause for our observed data. When we have some data or evidence from our observation of reality and the universe, Bayesian analysis allows us to calculate the most likely cause for that evidence.

Bayesian analysis treats virtually everything as a probability – including reality and the universe. This does not necessarily mean that Bayesian inference actually regards reality as variable – it means only that when reality is unknown there is some probability or likelihood associated with different ways that the universe exists in reality. Returning to the previous example: a person is still either pregnant or not pregnant, not both, and not partially one or the other. For another example, consider the situation of deception or truth-telling: truth and deception are, in reality, a deep and complex philosophical and epistemological rabbit hole that are beyond the scope of this paper, but for practical purposes we can consider that a person's statements about a thing or event to be either truthful or deceptive, not both, and for practical purposes not partially one or the other.

Of course, when we can make perfect deterministic observations and conclusions then we have no need for Bayesian or frequentist statistics. Similarly, when we can make direct physical measurements we have no need for statistical estimation. The difference between Bayesian inference and frequentist inference is that when reality is unknown the Bayesian approach is to simply treat the unknown as a probability. The purpose of Bayesian inference is to help us make the best decision when our knowledge is uncertain. What is the probability that a person is pregnant or not pregnant? Or what is the probability that a person is deceptive or truthful? Whereas frequentist inference treats reality as fixed and data as variable, the Bayesian inference treats unknown reality as a probability.

Bayesian analysis says simply that when reality is unknown we think of the different possibilities. In practice it is seldom the case that we have absolutely no information about reality. More often we have some uncertain or incomplete knowledge or belief about the situation or phenomena of interest to us. Bayesian inference requires us to declare our knowledge or beliefs in the form of an *a priori* or prior probability, and then uses the observed data and evidence to mathematically update or modify our prior probability into a posterior probability. Posterior probabilities have obvious practical value, but it is important to remember that we can have a posterior probability only when we are willing to declare or prior knowledge in the form of a prior probability (actually a prior probability distribution).

Another difference between Bayesian inference and frequentist inference is that Bayesian inference treats the available or observed data as fixed. The available data is all the information we have to use to modify our knowledge and conclusions about what we think we know about reality and the universe when our knowledge is uncertain. When we obtain more data we can again modify or update our knowledge and conclusions. But unless we achieve some perfect deterministic observation, or unless we achieve a direct physical measurement, our knowledge and conclusions about reality exist only as a probability that the observed evidence was caused by reality as our conclusions would have us think of it.

Because it goes more directly to conclusions about causes, Bayesian inference has tremendous practical application. It has been used to improve the accuracy of artillery trajectories, locate submarines, find lost or missing persons, and even to estimate the number of German tanks and crack the encryption codes of the enigma system during WWII. It is used routinely in medicine, epidemiology, finance, econometrics, data and business analytics, forensics, internet search engines, word prediction apps on mobile devices, sports statistics and sports betting, and, perhaps most importantly - filtering out email-spam for male-enhancement products or long-lost wealthy relatives in far away countries. Virtually every form of classification and prediction scheme in use today has made use of Bayesian inference in some way.

Summary

Science is a systematic way of thinking and acting to acquire and make use of new knowledge about reality and the universe. One difference between scientific knowledge and pseudoscience is that science is expected to keep learning and to keep making use of new knowledge, and any reasonably intelligent and diligent person can attain knowledge. In contrast, pseudoscience tends to rely on dogma that is sometimes static, with nothing new to learn about reality because it is not actually connected to reality, and sometimes accepted only from one source – the owner/originator/guru who can be the only true fountain of esoteric wisdom. While science is about knowledge and reality and tends to diversify, pseudoscience is primarily about economics and power and tends to centralize.

Another difference between science and pseudoscience is that science expects us to quantify our knowledge even when it is difficult to do so, whereas pseudoscientific dogma must be accepted on faith without quantification. Where pseudoscience and dogma are completely satisfied with the memorization of a set of words in a particular order as an acceptable final answer for which we need not and must not inquire further, science holds the goal of understanding the practical use and limitations of a concept or piece of knowledge. Quantification of scientific knowledge and scientific conclusions, when it is difficult or impossible to achieve a direct physical measurement relies on statistics. Both frequentist inference and Bayesian inference are fundamental to the practice of science.

Statistics, as the language of science, allows us to move beyond impressionistic and subjective statements towards quantifying some of the complex phenomena that are very real though also very difficult or impossible to observe directly with our senses. As it turns out some of the most interesting things that we may wish to measure are actually very difficult or impossible to measure – things like a person's personality, level of intelligence, or deception and truth-telling. Mundane things like a person's height and weight are easy to measure. Measurement requires two things: 1) a physical phenomena and 2) a defined unit of measurement. When we lack these two things and still desire to quantify some phenomena, we use statistics.

Statistics is related to probability and chance. The practical application of statistics in science goes back about 2 or 3 centuries. But the practical application of probabilities outside of science may go back much further – in gaming and wagering. Early statisticians like Bayes and Laplace would have been uninterested in gaming, though today statistics and probabilities play an important role in sports, sports betting and wagering of all kinds in addition to their important roles in business, science, and even politics.

Believable or not, some people have, at times, suggested that statistics is not fun, that statistics is difficult, and even that statistics is boring. In fact, statistics is all of these – statistics is not for the weak minded or faint-of-heart. Statistics is not for people who cannot find patience or attention for details. In fact, statistics is not for wimps. However, learning about statistics is much less boring and difficult – or at the very least it is more tolerable – when there is some very interesting or very important problem to solve. The logic of frequentist inference is not immediately intuitive for some, but most people can develop their intuition for statistics by learning a few basic concepts. Fortunately for most professionals in field practice versus research or academic work, it is rarely necessary today for humans to do the actually statistical math – for that we have computers. It is very important, however, that professionals who desire to serve their agencies and communities as experts and not mere practitioners should feel some obligation to learn the basic principles of science and statistics – beginning with the fundamentals of frequentist and Bayesian inference.

©FotoliaLLC/Karen Roach