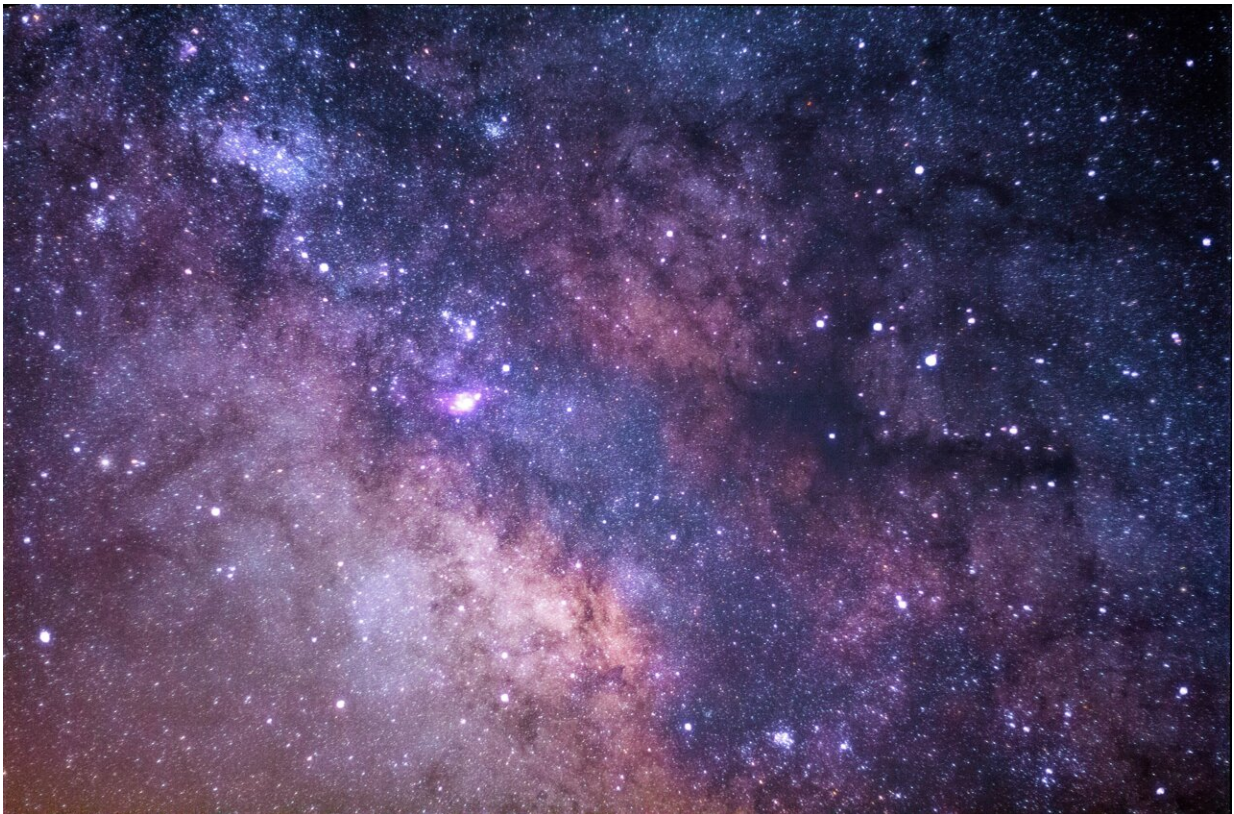# Big data astronomy: Using statistics in a new way to decipher the universe

June 5 2023, by Russ Nelson

Credit: Unsplash/CC0 Public Domain

The digital age has been a tremendous boon to the fields of both statistics and astronomy. However, according to Dr. Max Bonamente, a professor of physics and astronomy at The University of Alabama in

Huntsville (UAH), most astronomers are not sufficiently trained to realize the substantial benefits to be gained by putting these disciplines together. He and his colleagues are working to change all that through pioneering research in the burgeoning field of astrostatistics.

Dr. Bonamente published a paper in the *Monthly Notices of the Royal Astronomical Society* showcasing an innovative new twist in [probability distributions](#) that promises to revolutionize the ways cosmological data can be interpreted.

"Traditionally, astronomers have been poor statisticians—we like to 'make up the statistics as we go,'" the researcher explains. "My latest paper is a new method to account for systematic errors. It describes a new probability distribution method I developed that hadn't been thought of before. It's nerdy stuff, but has real-life implications in terms of making conclusions from observations. Many astronomers don't have the necessary math background to do statistics carefully. It's hard, because statistics is hard math at its core. Few people want to take the extra time to do it. Of course, not everyone feels that way."

This is evidenced by the success of a workshop called iid2022: Statistical Methods for Event Data and subtitled, Illuminating the Dynamic Universe, recently hosted by UAH, a part of The University of Alabama System. Dr. Bonamente and his colleague, Dr. Lingling Zhao, an assistant professor of space science, organized the workshop.

The gathering was designed to train young scientists in proper statistical methods for the analysis and interpretation of data and included hands-on collaborative analysis of sample problems employing advanced software. The gathering also provided a forum for astronomers and researchers in related fields to exchange recent advances in the analysis of event data.

"Event data" are the collection of individual events—in astronomy, typically light photons, but also neutrinos or other particles. These events can be studied through statistical applications as a function of location (images), time (such as light curves) or energy or wavelength (spectra). Events can also be defined as ensembles of quantities, such as gravitational wave events or galaxy clusters detected through measurements of the Cosmic Microwave Background, which is the cooled remnant of the first light that could ever travel freely throughout the universe.

A native of Italy, Dr. Bonamente moved to the U.S. in 1997, and is a UAH alumnus, earning both an M.S. and Ph.D. in physics at UAH, where he has developed the use of a statistical method called Markov chain Monte Carlo (MCMC) for analysis of cosmological events. MCMCs comprise a class of special algorithms used in probability distributions, a mathematical function that gives the probabilities of the occurrence of different possible outcomes for an experiment.

"These methods have made it possible to analyze data faster and with greater accuracy," the researcher notes. "Nowadays, machine learning is everywhere in astronomy. We used MCMCs to measure the Hubble constant, for example, which was a big deal at the time." The Hubble constant is one of the most important numbers in cosmology, because it tells us how fast the universe is expanding.

Astrostatistics represents the future of big data management and analysis in astronomy, as the latest technologies are producing staggering amounts of data of truly mindboggling complexity. The challenge to analyze this data is only growing exponentially as new data-gathering mechanisms evolve in radio, microwave, infrared, X-ray, gamma ray, interferometer and optical instruments that will require new statistical algorithms and techniques to make sense of it all.

"Most astronomers or physicists don't know much of probability theory, let alone statistics," Dr. Bonamente points out. "A scientist's job ought to be that of being careful, and not to give in to the desire to find a great new result when it's not there. So, marrying math and astronomy is the natural direction for me."

More information: Massimiliano Bonamente, Systematic errors in the maximum-likelihood regression of Poisson count data: introducing the overdispersed χ2 distribution, *Monthly Notices of the Royal Astronomical Society* (2023). DOI: 10.1093/mnras/stad463

Provided by University of Alabama in Huntsville