2016-2017 SAMSI Program on
Statistical, Mathematical and Computational Methods for Astronomy (ASTRO)

G. Jogesh Babu

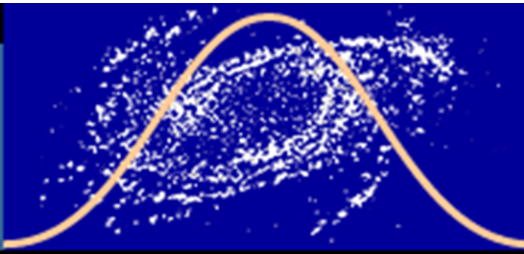
Statistical and Applied Mathematical Sciences Institute

samsi
NSF•Duke•NCSU•UNC

Eberly College of Science

PennState

Center for Astrostatistics



Why astrostatistics?

Astronomers encounter a surprising variety of statistical problems in their research:

- The sky has vast numbers of stars & galaxies and gas on all scales.
- Most stars have orbiting planets, most galaxies have a massive black hole
- Astronomers acquire huge datasets of images, spectra & time series of planets, stars, galaxies, quasars, supernovae, etc.
- Various properties of cosmic populations observed and empirically studied with all kinds of telescopes (n>>p)
- Properties are measured repeatedly but with irregular spacing.
- Spatial distributions in sky (2D), space (3D), and parameter space (pD) is complex (MVN assumption usually inapplicable)

Eric Feigelson and I started collaborating in late 1980s and the term `Astrostatistics' was coined in mid 1990s, when we published a book by the same name.

Astrostatistics at SAMSI

Astrostatistics Program at SAMSI January 2006

- Opening workshop January 18-20, 2006
  - Bayesian astrostatistics
  - Nonparametric inference
  - Astronomy for Statisticians
- Working groups
  - Exoplanets, Surveys & Population studies
  - Gravitational Lensing,
  - Source detection & feature detection
  - Particle Physics
- Concluded with SCMA IV at Penn State in June 2006

Astrostatistics at SAMSI

Astrostatistics sub Program Fall 2012

Statistical and Computational Methodology for Massive Datasets

- Workshop September 19-21, 2012
  - The search for transients
  - Missions (Fermi, SDSS. DES, Plank, LSST, LIGO)
  - Sparsity (*high-dimensional data, but low-dim signal*)
  - Data Mining
- Working groups
  - Discovery & Classification in Synoptic Surveys;
  - Inference & Simulation in Complex Models,
  - Stochastic Processes& Astrophysical Inference
  - Graphical Models & Graphics Processors

Astrostatistics at SAMSI

Exoplanets Summer 2013

Modern Statistical and Computational Methods

for Analysis of Kepler Data

June 10-28, 2013

Astrostatistics at SAMSI

Statistical, Mathematical and Computational Methods for Astronomy 2016-2017

- Opening workshop August 22-26, 2016
  - Time Domain Astronomy
  - Exoplanet Data analysis, Hierarchical Modeling
  - Uncertinity, selection effects etc., in GW
  - Pulsar Timing Arrays and Detection of GWs
  - Non-stationary, non-Gaussian, Irregularly sampled processes.
  - Statistial issues in Cosmology

SAMSI ASTRO Program is timely

- Gravitational Waves Detected 100 years after Einstein's Prediction

- First GW (*GW150914*) was detected by Laser Interferometer Gravitational-Wave Observatory (LIGO) Confirming a major prediction of Albert Einstein's 1915 general theory of relativity. (GW from colliding Black Holes).

- The planning meeting on September 21, 2015 included LIGO scientists who had only just learned of the candidate detection, and had to keep it secret until confirmed and announced on February 11, 2016.

  A second GW event (*GW151226*) was announced on June 15, 2016. It was recorded on December 26, 2015.

Working Groups

I. Uncertainty Quantification and Astrophysical Emulation

II. Synoptic Time Domain Surveys

III. Multivariate and Irregularly Sampled Time Series

IV. Astrophysical Populations

V. Statistics, computation, and modeling in cosmology

Uncertainty Quantification and Astrophysical Emulation

- Uncertainty Quantification (UQ) and Reduced Order Modeling (ROM) are at the core of many problems in gravitation and cosmology, from direct simulations of the Einstein equations to the inverse problem (inference problem).

- The goal of this working group is to leverage expertise from areas such as generalized polynomial chaos and simulator emulation based on stochastic processes, and domain-areas such as gravitation, astrophysics, and cosmology.

  Group Leaders: Derek Bingham (Simon Fraser), Earl Lawrence (LANL)

Synoptic Time Domain Surveys

- TDA has been getting richer in terms of datasets that span several years, many bands, and include dense and sparse light-curves for hundreds of millions of sources.
- The variety, volume etc. squarely fall in Big Data regime.
- The light-curves often have large gaps, are heteroskedastic, and the intrinsic variability – often poorly understood – adds an element of uncertainty when multi-band data are not obtained simultaneously.
  - Kepler-type planet search/characterization (also in other WGs)
  - binary black-hole searches from CRT Transient Survey

Group Leaders: Ashish Mahabal (Astro, Caltech), G. Jogesh Babu (Stat, PSU)

Ashish Mahabal will discuss some of the work under various subgroups of this WG.

Multivariate and Irregularly Sampled Time Series

- Ground-based photometric surveys, with billions of irregular light curves, detect stellar variability from a multitude of sources, from planetary transits, to supernovae. How do we effectively separate these classes in irregularly sampled data?
- How can we maximize the sensitivity of space-based photometric surveys for detecting small planets in the presence of stellar activity?
- How can we measure planet masses robustly despite observational data with outliers (due to stellar activity affecting either Doppler data or transit times)?

Group Leaders: Ben Farr (Astro, U. Chicago), Soumen Lahiri (Stat, NCSU)

Statistics, Computation, and Modeling in Cosmology

- Will bring together leading researchers in cosmology, computational spatial and Bayesian statistics, experimental design, and computer modeling to develop methodology necessary for answering fundamental questions about the origin and large scale structure of the universe.
- How can we make inferences for deterministic nonlinear dynamical systems (inference of initial conditions and model parameters)?

Group Leaders: Jeff Jewell (Astro, JPL), Joe Guinness (Stat, NCSU)

Jeff Jewell will discuss the applications areas under this WG.

Astrophysical Populations

- Improve the statistical methodology for interpreting detections of exoplanets, gravitational waves (GW), as well as using those to infer the underlying population of planetary systems and GW sources.
- The exoplanets community is particularly interested in developing techniques to robustly detect and characterize planets in the presence of stellar activity from Doppler Surveys for which we do not have a first principles model.
- The GW community is interested in detecting gravitational wave sources for which the details of the primary GW signal and/or backgrounds are unknown.
- Both applications require developing algorithms to efficiently explore high-dimensional parameter spaces and to establish confidence in detections, despite complex and unknown sources of background signals that "noise".

Group Leaders: Jessi Cisewski (Stat, Yale); Eric Ford (Astro, Penn State)

Jessi Cisewski will now discuss some of the progress made under this WG.