LTCC Proposed Course

- Title: Algebraic Statistics

- Basic Details: An introduction to the application of computational commutative algebra and geometry to statistics, with special emphasis on ideals of points (designs) and toric ideals applied to categorical models.

  - Core Audience: statistics and pure with an interest in applied algebra.
  - Course Format: 10 hours intensive: 5 lectures plus 5 computer classes using symbolic algebra software: Maple, CoCoa, Singular.

- Course Description: This is an introduction to the fairly recent field in which computational algebraic geometry, and Groebner bases in particular, are applied in statistics. They arise principally in two ways: via considering experimental designs as solutions of polynomial equations and via probability models expressed via polynomial identities, particularly in contingency table and graphical models.

  - Keywords: Groebner bases, toric ideals, varieties, ring theory, graphical models, experimental design

  - Syllabus:

    Lecture 1 plus class. Rings, ideals, varieties, bases, quotients, Groebner bases. This is a a very basic introduction to the material in Cox, Little and O'Shea.

    Lecture 2 plus class. Experimental designs and ideal of points. Use of G-bases to find candidate models as order ideals. Explanation of aliasing. Indicator function representation. Corner cut models.

    Lecture 3 plus class. From loglinear models to toric ideals. Power product representation. Saturation. Independence and conditional independence, graphical models. Moments and maximum likelihood.

    Lecture 4 plus class. Sufficient statistics and margins. Conditional exact tests. Markov bases and simulation.

    Lecture 5 plus class. Hilbert functions and resolution. Minimal free resolution. Application to inclusion-exclusion identities and bounds and system reliability.

  - Recommended reading:

    Ideals, Varieties, and Algorithms An Introduction to Computational Algebraic Geometry and Commutative Algebra Series: Cox, David; Little, John; O'Shea, Donal 3rd ed., 2007, Springer.

    (The above is the basic text and familiarity with it BEFORE the start of the course would be very useful).

    Algebraic Statistics: Pistone, Giovanni; Riccomago, Eva; Wynn, Henry P. 2001, CRC.

    Algebraic Statistics for Computational Biology Edited by L. Pachter and B. Sturmfels University of California, Berkeley: 2005. Cambridge.

Search "algebraic statistics" on Google to get a lot of other information.

- Additional Optional reading: search "algebraic statistics" on Google to get a lot of other information.

- Prerequisites: Undergraduate abstract algebra and around MSc level statistics is useful but not absolutely essential.

- Format:

  - No of problem sheets: 5

  - Electronic lecture notes: 20-30 pages

  - Necessary support facilities: Computer rooms for five sessions, contiguous with the lectures.

  - necessary software requirements for computing facilities. Maple, CoCoa, Singular, 4ti2 (the latter three are free).

  - Proposed timing: Late Autumn

  - Lecture/computer session. split: 50/50

- Lecturer Details:

  - Lecturer: Professor Henry Wynn aided by Dr. Hugo Maruri-Aguilar

  - Lecturer home institution: LSE

  - Lecturer e-mail: h.wynn@lse.ac.uk

  - Lecturer telephone number: 0207 955 6063