

STT 855 – EPI 855: Introduction to Statistical Genetics
Fall 2019 (3 credits)

Time: T & Tr 12:40PM – 2:00 PM

Place: 311 Ernst Bessey Hall

Instructors contact information & office hours:

Yuehua Cui	Gustavo de los Campos
C432 Wells Hall	1311 IQ – building (775 Woodlot Dr.)
Tel: (517) 432-7098	Tel: (517)884-7607
Email: cuiy@msu.edu	Email: gustavoc@msu.edu
Web: http://www.stt.msu.edu/~cui	Web: https://quantgen.github.io/
Office hours: Th 2-3pm (2 nd half of the course)	Office hours: 2-3PM Statistics department, 4 th floor C Wells Hall (1 st half of the course)

Textbook and other supporting materials: The first half will follow Introduction to Quantitative Genetics (4th Edition) 4th Edition by Douglas S. Falconer & Trudy F.C. Mackay (ISBN 0-582-24302-5). For the 2nd half, Dr. Cui will make the lecture notes available in D2L and no textbook is required.

Lecture notes, homework and other materials will be made available through [D2L](#).

Prerequisite: STT863 + STT442 or STT861 or equivalent.

Basic skills necessary for this course: Calculus; Basic probability distribution theory; t-test; Chi-square test; linear and logistic regression; ANOVA; Maximum likelihood estimation; Computer software (e.g., R)

Course objectives: The course is designed for Master and PhD students in statistics and/or life science majors. Students will learn basic quantitative genetics concepts and basic and advanced statistical methods for genetic data analysis.

Course description

The course is organized in two modules: (I) quantitative genetics principles and (II) statistical methods for analyses of genetics data.

Content by Module

I-Introduction to quantitative genetics (by *G. de los Campos*) will cover the following topics (topics 9 and 10 will be covered if time permits).

1. Introduction:
 - Mendel, laws of inheritance and a brief history of Quantitative Genetics
 - The “gene” (Sept 3rd, prof. David Arnosti).
2. Allele and genotype frequencies (CH1 Falconer & Mackay, F&M):
 - Allele frequencies and genotype frequencies
 - Hardy-Weinberg equilibrium
 - Deviations from Hardy-Weinberg equilibrium
3. Sources genetic variation (mutation, selection and drift, CH2, F&M)
4. Values and means in a single locus model (CH6 & CH7, F&M):
 - Additive, dominance and allele substitution effects
 - Variance decomposition in the single-locus model
5. Extension to multi-locus models (selected publications)
6. Resemblance between relatives (CH9):
 - Kinship: definition and computation from pedigree and from SNP data
 - Heritability estimation using parent-offspring- and sib-regressions.
7. Correlations between traits (if time permits)
8. Association of alleles at pairs of loci:
 - Meiosis and recombination
 - Linkage and Linkage Disequilibrium
 - Metrics (D , D') definition, estimation and testing
 - Physical versus map distance

II-Statistical methods for analyses of genetic data (by *Y. Cui*) will cover the following topics

9. Mapping quantitative traits using marginal association analyses (aka GWAS)
 - Estimation and inference in a single-locus model: OLS, chi-square, logistic regression.
 - Multiple-testing (type-I error rate and FDR control)
 - Diagnostics (qqplots, p-value inflation)
 - Population stratification and controlling for population structure
10. Haplotype analysis and gene-gene interaction analysis
11. Mapping quantitative traits using family data
 - Family-based linkage mapping
 - TDT test
12. Sequencing data and rare variants analysis
13. Genetic risk prediction
14. Genomic Analyses of multi-locus models
 - Heritability estimation using mixed effects models
 - Best Linear Unbiased Prediction (BLUP)
 - Genomic prediction

Textbook:

Required: Introduction to Quantitative Genetics by Falconer & Mackay (4th ed)

Other references:

- [*The Fundamentals of Modern Statistical Genetics*](#) by Nan M. Laird and Christoph Lange, Springer-Verlag, New York.
- *Applied Statistical Genetics with R For Population-based Association Studies* by Andrea S. Foulkes. Use R! series, Springer
- [*Statistical Genetics of Quantitative Traits: Linkage, Map and QTL*](#) by Rongling Wu, Chang-Xiang Ma and George Casella, Springer-Verlag, New York.
- [*Mathematical and Statistical Methods for Genetic Analysis*](#) by Kenneth Lange, Springer-Verlag, New York.
- *Genetics and Analysis of Quantitative Traits* by Michael Lynch, Bruce Walsh, Sinauer Assoc.

Grading

The final grade will be the average of the grade in the first and 2nd modules (see above).

Module 1:

- 30% homework
- 70% midterm

Module 2:

- 30% homework
- 70% Individual data analysis project & report
 - Research project: The instructor will conceive of multiple problems related to the topic of the 2nd module of the course. Students can work singly or as a group. ***Each student/group will choose one of the problems as their final project.*** Students with different backgrounds are encouraged to work together.
 - Alternatively, ***students are also encouraged to base their class projects on one of their own problems (with real data).*** But they should consult with the instructor first.
 - Paper review: Students can choose to review one or several papers with similar topics and write a report.

Format of written report (Follow the format of a refereed journal, e.g., *Genetics* or *Biometrics*): Introduction, Methods, Results, Discussion, References. Page limit: ≤ 20, 12pt font double-spaced pages (including tables, figures and references).