

New Course Request

Indiana University

Indianapolis Campus

Check Appropriate Boxes: Undergraduate credit Graduate credit Professional credit

1. School/Division Science, Biostatistics 2. Academic Subject Code BIOS
3. Course Number S530 (must be cleared with University Enrollment Services) 4. Instructor Y. Liu
5. Course Title Statistical Methods in Bioinformatics

Recommended Abbreviation (Optional) _____
(Limited to 32 Characters including spaces)

6. First time this course is to be offered (Semester/Year): Fall 2009

7. Credit Hours: Fixed at 3.0 or Variable from _____ to _____

8. Is this course to be graded S-F (only)? Yes _____ No X

9. Is variable title approval being requested? Yes _____ No X

10. Course description (not to exceed 50 words) for Bulletin publication: P: STAT 51200 and 51900, familiarity with statistical inference; or consent of instructor.
Covers statistical methods used in many areas of bioinformatics research, including sequence alignment, genome sequencing and gene finding, gene expression microarray analysis, transcriptional regulation and sequence motif finding, comparative genomics, and proteomics

11. Lecture Contact Hours: Fixed at 3 or Variable from _____ to _____

12. Non-Lecture Contact Hours: Fixed at _____ or Variable from _____ to _____

13. Estimated enrollment: 5-10 of which 100 percent are expected to be graduate students.
TBD

14. Frequency of scheduling (new program) Will this course be required for majors? elective

15. Justification for new course: _____ course for new biostatistics Ph.D. program

16. Are the necessary reading materials currently available in the appropriate library? Yes

17. Please append a complete outline of the proposed course, and indicate instructor (if known), textbooks, and other materials.

18. If this course overlaps with existing courses, please explain with which courses it overlaps and whether this overlap is necessary, desirable, or unimportant.

19. A copy of every new course proposal must be submitted to departments, schools, or divisions in which there may be overlap of the new course with existing courses or areas of strong concern, with instructions that they send comments directly to the originating Curriculum Committee. Please append a list of departments, schools, or divisions thus consulted.

Submitted by: Berlin Beh Date 3/17/09
Department Chairman/Division Director

Approved by: James M. Murphy Date 4/17/2009
Dean

Date _____
Dean of Graduate School (when required)

Date _____
Chancellor/Vice-President

Date _____
University Enrollment Services

After School/Division approval, forward the last copy (without attachments) to University Enrollment Services for initial processing, and the remaining four copies and attachments to the Campus Chancellor or Vice-President.

BIOS S530 (3 cr.)
Statistical Methods in Bioinformatics

Syllabus

A. Instructors

Yunlong Liu, PhD, Assistant Professor
Lang Li, PhD, Associate Professor
Xiaoman Li, PhD, Assistant Professor
Changyu Shen, PhD, Assistant Professor

Contact information:

Yunlong Liu, PhD
Division of Biostatistics
Indiana University School of Medicine
410 West 10th Street, Suite 5000
Indianapolis, IN 46202

Office: (317) 278-9222
Email: yunliu@iupui.edu

B. Course Description and Rationale

BIOS S530 is a graduate level course designed for students in biostatistics, statistics, bioinformatics, and other related areas. The course covers a broad range of statistical methods used in many areas of bioinformatics research, including sequence alignment, genome sequencing and gene finding, gene expression microarray analysis, transcriptional regulation and sequence motif finding, comparative genomics, , proteomics and. This course is designed to train student's skills in data analyses and communications through real life bioinformatics projects. The course's primary audiences include are graduate students in biostatistics, bioinformatics, and researchers from pharmaceutical industry.

Course materials will be covered by lectures. Knowledge gained will be reinforced by homework assignments and programming exercises. The course is part of the Biostatistics Ph.D. curriculum.

C. Prerequisites:

Students are assumed to have completed a graduate level statistics courses, such as STAT 51200 and STAT 51900, and are familiar with the basic concepts of statistical inference. Students who are uncertain about their levels of preparation are encouraged to contact the instructors.

D. Course Description for Bulletin

P: STAT 51200 and 51900, familiarity with statistical inference; or consent of instructor.
Covers statistical methods used in many areas of bioinformatics research, including sequence alignment, genome sequencing and gene finding, gene expression microarray analysis,

transcriptional regulation and sequence motif finding, comparative genomics, and proteomics.

E. Educational Objectives

The objective of this course is to provide students with an understanding of statistical methods used in bioinformatics. At the end of the course, the students are expected to understand computational molecular biology principles, statistical theories, and bioinformatics software implementation.

F. Course Content

1. Week #1: Introduction
 - General overview of the principles of molecular and cellular biology that will be used throughout the course
2. Week #2: Overview of probability theories and statistics
 - Review of discrete random variables
 - Probability density functions
 - Joint and conditional probability
3. Week #3-6: Sequence statistics and analysis
 - Statistical theories and dynamic programming
 - Sequence alignments
 - Motif discovery
 - Software implementation
4. Week #7: Midterm exam
5. Week #8-12: High throughput data analysis
 - Two-color and one color array data technology and normalization
 - Proteomics data analysis
 - Clustering and principle components
 - Empirical Bayesian models
 - Multiple comparisons and FDR
 - Molecular Signatures and Profiles
 - Software implementation
6. Week #13-14: Gene ontology, pathway analysis, and network theories
 - Hypergeometric Distribution Information
 - Gene enrichment analysis
 - Bayesian-based pathway analysis
 - Software implementation
7. Week #15: Final Exam

G. Required and Recommended Texts

The course will use the following textbook:

Statistical Methods in Bioinformatics, an Introduction by Warren J. Ewens and Gregory R. Grant, Springer.

H. Evaluation and Grading

Students' performance will be assessed by grades of two projects, homework, and two exams. Class materials and home work assignments will be posted on Oncourse and home works will be collected in class the following week. No late home works will be accepted. Make-up examinations will be given only in extraordinary situations (such as serious illness) and can be arranged after receiving prior consent from the instructor. The final course grade will be determined using the following weighting scheme:

Two projects + presentation (15% each)	30%
Midterm exam	20%
Final exam	30%
Homework	20%

Letter grades for the course are assigned using the following scale: A: 90-100; B: 80-89; C: 70-79; D: 60-69; F: less than 60. Within each letter grade, "+" and "-" will be assigned if the numeric score is in the top and bottom quintiles, respectively.

I. Bibliography

N/A

J. Cheating and Plagiarism

Academic misconduct will *not* be tolerated and all cases will be reported. Examine the IU Code of Student Rights, Responsibilities, and Conduct at <http://www.iupui.edu/code> and in particular examine the rules regarding academic misconduct at http://www.iupui.edu/code/#P2_G. Violations of these rules will result in a grade of "F" (or 0%) for the assignment in question, and may result in an "F" for the course or even expulsion from the university (see <http://life.iupui.edu/rights/undergrad/sanctions.html>).

K. Americans with Disabilities Act

If you need any special accommodations due to a disability, please contact Adaptive Educational Services at (317)-274-3241 or Joseph T. Taylor Hall (UC), Room 137.