New Course Request **Indiana University** Indianapolis Campus

Check Appropriate Boxes: Undergraduate credit ☐ Graduate credit X ☐ Professional credit ☐

1. School/Division Science, Biostatistics 2. Academic Subject Code BIOS
3. Course Number S646 (must be cleared with University Enrollment Services) 4. Instructor J. Harezlak
5. Course Title Advanced Generalized Linear Models

   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: BIOS S546, familiarity with statistical inference. The theory of classical and modern approaches to the analysis of clustered data, repeated measures, and longitudinal data: ramdom effects and growth curve models, generalized estimating equations, statistical analysis of multivariate categorical outcomes, estimation with missing data. Discussion of computational issues: EM algorithm, quasi-likelihood methods, Bayesian methods for both traditional and new methodologies.
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 TBD of which 100 percent are expected to be graduate students.
14. Frequency of scheduling: new program Will this course be required for majors? Yes
15. Justification for new course: Required 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: Barbara Bailey 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.

UPS 724

University Enrollment Services Final—White; Chancellor/Vice-President—Blue; School/Division—Yellow; Department/Division—Pink; University Enrollment Services Advance—White

BIOS S646 (3 cr.)
Advanced Generalized Linear Models

Syllabus

A. Instructors:

Jaroslaw Harezlak, PhD, Assistant Professor
Changyu Shen, PhD, Assistant Professor
Sujuan Gao, PhD, Associate Professor

Contact information:
Jaroslaw Harezlak, PhD
Division of Biostatistics
Indiana University School of Medicine
410 West 10th Street, Suite 3000
Indianapolis, IN 46202

Office: (317) 274-2682
Email: harezlak@iupui.edu

B. Course Description:

This course presents the theory of classical and modern approaches to the analysis of clustered data, repeated measures, and longitudinal data. Topics include random effects and growth curve models, generalized estimating equations, statistical analysis of multivariate categorical outcomes, and estimation with missing data. The class also discusses computational issues including EM algorithm, quasi-likelihood methods and Bayesian methods for both traditional and new methodologies. The course is part of the Biostatistics Ph.D. curriculum.

C. Prerequisites:

Bios 546 and familiarity with concepts and theory of statistical inference. Students who are uncertain about their level of preparation are encouraged to contact the instructors.

D. Course Description for Bulletin:

P: BIOS S546, familiarity with statistical inference. The theory of classical and modern approaches to the analysis of clustered data, repeated measures, and longitudinal data: random effects and growth curve models, generalized estimating equations, statistical analysis of multivariate categorical outcomes, estimation with missing data. Discussion of computational issues: EM algorithm, quasi-likelihood methods, Bayesian methods for both traditional and new methodologies.

E. Educational objectives:

At the end of the course students:

- will know the theory of modern statistical methods for the analysis of repeated measures, correlated outcomes and longitudinal data

- will be able to apply the methods using modern computational algorithms

F. Course outline (Estimated Number of Lectures) :

1. Introduction (1)
   Reading: DHLZ §1-3

2. Linear Mixed Models for Correlated Data (6)
   LMM; MLEs; Fixed effects; Random effects; BLUPs; EM; Restricted MLEs (REMLs); Variance components; Variance/covariance modeling.
   Reading: DHLZ §4-6, VM1 §3-10

3. Generalized Linear Models and Quasi-likelihood (6)
   Exponential family; Mean-variance relationship; Link functions; MLEs; Weighted least squares; Hypothesis testing; Over-disperson; Estimating equations; Quasi-likelihood Functions.
   Reading: MN §2, 4, 9

4. Generalized Estimating Equations (4)
   GEE1, GEE2, Hypothesis Testing in GEE.
   Reading: DHLZ §7-8, VM2 §6, 8, 10

5. Generalized Linear Mixed Models (6)
   Relationship between GEEs and GLMMs; Likelihood; Approximate inference (PQL); MLEs; EM algorithm; Conditional inference; Bayesian inference.
   Reading: DHLZ §9, 11, VM2 §13-16

6. Modeling Longitudinal Data with Dropouts (6)
   Missing mechanisms (MCAR, MAR and NMAR); Likelihood-based approaches (selection models and pattern-mixture models); Weighted GEEs.
   Reading: DHLZ §13, VM1 §15-21, VM2 §26-32

G. Meeting Times:
Lectures twice per week

H. Required and Recommended Text:
For the most part the course will use the following textbook:
(DHLZ) Diggle, Heagerty, Liang and Zeger (2002). Analysis of Longitudinal Data, Oxford University Press.

We will supplement the textbook with instructors' own lecture notes,

Additional texts:
(MN) McCullagh, P. and Nelder, P.(1989). *Generalized Linear Models.* Chapman & Hall/CRC.

(VM1) Verbeke,G. and Molenberghs, G. (2000). *Linear Mixed Models for Longitudinal Data.* Springer.

(VM2) Verbeke, G. and Mollenberghs, G. (2005) Models for Discrete Longitudinal Data. Springer

I. **Evaluation and Grading:**
Students will be evaluated based on their performance on the homework assignments (30%), take home mid-term exam (30%), and the term paper and presentation (30%). 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.

J. **Cheating and Plagiarism: 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. Joseph T. Taylor Hall (UC), Room 137.