

STATISTICS 516
STATISTICAL DESIGN AND ANALYSIS OF GENE EXPRESSION EXPERIMENTS
SPRING 2011

Location: 1322 Hoover Hall
Time: 9:30–10:45 a.m. Tuesday and Thursday
Instructor: Dan Nettleton dnett@iastate.edu <http://www.public.iastate.edu/~dnett/>
Office: 2115 Snedecor Hall (phone: 515-294-7754)
Hours: 10:00–10:50 a.m. Monday and Wednesday or other times by appointment

Purpose of the Course:

STAT 516 will cover application and practical use of statistical methods for designing and analyzing experiments that involve the measurement of gene expression. Current statistical research topics related to statistical design and analysis of such experiments will be emphasized throughout the course. Students completing STAT 516 should be able to provide expert advice on gene expression experimental design, perform appropriate analyses in collaboration with biological researchers, and be ready to consider research problems in the statistical design and analysis of gene expression experiments. Furthermore, students should gain a sound understanding of the statistical principles important for good experimental design and analysis involving high-dimensional data.

Prerequisites:

STAT 500 is the prerequisite for STAT 516.

Grading:

Homework	25%
Midterm Exam	25% Thursday, March 3, 9:30–10:45 a.m.
Project	15% In-class presentations on April 19, 21, 26, 28
Final Exam	35% Monday, May 2, 9:45–11:45 a.m.

On exam days, be sure to arrive early with a calculator and pen or pencil. No make-up exams will be given unless circumstances are exceptional. Your final course percentage will be determined by the following formula:

$$\text{Course Score} = 0.25 H + 0.25 M + 0.15 P + 0.35 F$$

where H, M, P, and F represent the score out of 100 possible on homework, the midterm exam, the project, and the final exam, respectively. Letter grades will be assigned according to the course scores. Students with similar course scores (as judged by the instructor) will receive similar letter grades. The order of the course scores will match the order of the letter grades. All students have the opportunity to earn good grades in Statistics 516 by learning the course material well. Please let your instructor know of any difficulties as soon as possible. If a student has a disability that qualifies under the Americans with Disabilities Act and Section 504 of the Rehab Act requires accommodations, he/she should contact the Disability Resources (DR) office for information on appropriate policies and procedures. DR is located on the main floor of the Dean of Students Building, Room 1076; their phone is 515-294-7220.

Homework Guidelines:

- The purpose of all homework assignments in Statistics 516 is to help you learn the course material.
- All homework assignments are posted at <http://www.public.iastate.edu/~dnett/microarray/hw.shtml>.
- Due dates are posted on the web site and written on the homework documents.
- No late homework is accepted, but your lowest homework score will not be used to compute your grade.
- Answers to homework assignments should be neatly written or typed and well organized with no extraneous information.
- You are welcome to work with other students on homework problems. You must, however, write your own answers to the questions. Copying another's work is prohibited.

Intended Course Coverage:

1. Introduction to the basic biology of gene expression
2. The basics of cDNA microarray technology
3. The basics of oligonucleotide microarray technology
4. Microarray image processing
5. Normalization of two-color microarray data
6. Normalization and construction of expression measures using Affymetrix microarray data
7. Introduction to microarray experimental design
8. Split-plot experimental designs in microarray experimentation
9. Incomplete block experimental designs
10. Latin square experimental designs
11. Introduction to mixed linear models
12. Choosing between competing experimental designs
13. Multiple testing methods for the analysis of microarray data
14. Mixture modeling of the p -value distribution
15. Power and sample size calculations for microarray experiments
16. Introduction to randomization tests
17. Significance Analysis of Microarrays (SAM)
18. Estimation of gene-specific variances
19. Parametric empirical Bayes methods for microarrays
20. Cluster analysis of microarray data
21. Identifying differentially expressed gene categories
22. Next Generation Sequencing Experiments

Some of the topics above may be replaced with other current research topics on the statistical design and analysis of gene expression experiments.

We will use the free statistical software R throughout the course to help with statistical computing. SAS may be used occasionally.