

STA 70500 Multivariate Statistical Methods

Instructor Information:

Name: Prof. Yu (Ryan) Yue; **Department:** Paul H. Chook Department of Information Systems and Statistics;
Email: yu.yue@baruch.cuny.edu ; **Office:** VC 11-266; **Office hours:** Mondays and Wednesdays, 3 ~ 4 pm.

Course Description:

This course provides a comprehensive survey of widely used multivariate statistical methods intended for the user of advanced statistical methodology. The multivariate normal distribution and related distributions, including Hotelling T^2 and Wishart distributions, are introduced and their uses illustrated in statistical estimation and hypothesis testing in multivariate normal models. Additional topics introduced and applied include multivariate analysis of variance and covariance, multivariate regression, canonical correlation, principal component analysis, factor analysis, discriminate analysis, and cluster analysis. The lectures will be accompanied with SAS software package presentations. Students are required to complete a group project, in which they prepare their own database, apply these methods, and prepare a written report on their findings.

Course Materials:

Textbook: *Methods of Multivariate Analysis*, Alvin Rencher, 2nd edition.

Software: SAS (Statistical Analysis System) 9.3 or later.

Lecture Notes: will be prepared by professor and posted on Blackboard.

Blackboard Website: Go to <http://www.baruch.cuny.edu/bctc/blackboard/>

Course Requirements:

- **Homework:** There will be 6 assignments posted on the course blackboard. These assignments are to assist in learning and understanding the course materials. The solution to each assignment will be posted on Blackboard shortly after submission. **No late submission will be accepted.**
- **Exams:** There will be one midterm exam and one final exam. The midterm will be in class with closed book and notes, but you can bring a formula sheet made by your own, a calculator and (clean) scrap paper. The formula sheet is restricted to exactly one letter-sized sheet of paper on which you can write on both sides. The final exam will be take-home. Students will use SAS to analyze the real-data problems given in the test using appropriate methods learned in class. A written report is required to submit answers to test problems.
- **Group project:** Students will form groups on their own, but each group size will be restricted. Professor will provide a list of multivariate techniques that students may use to analyze the data sets found by themselves. The data sets cannot be those that have been analyzed in class. The project report must have following parts: objective(s) of analysis, data description, introduction of the methods, implementation of the methods, interpretations of the results and conclusions, and appendix where you put SAS codes and outputs. More information will be provided in a separate document.
- **Grading disputes:** Any grading dispute must be submitted in writing to the instructor within one week after the grade is posted. No changes will be made after the deadline. It is student's responsibility to check his or her grades on the course webpage for any discrepancy.

Evaluation Criteria:

Weighted Total = Homework (10%) + Project (20%) + Midterm (35%) + Final (35%). Your letter grade for this course will be determined by the weighted total ONLY (**no extra credit or rounding up**).

Attendance Policy:

Students are encouraged to attend every lecture. Excessive absences will certainly jeopardize course grade. While attending class, students are expected to show consideration for all persons present in the room. Any behavior (e.g., talking to your classmates, leaving classroom during lecture) that causes distraction is not allowed and will get penalized at the instructor’s discretion, with at least one percentage deduction in the final grade per instance. Serious cases will be turned over to the relevant office at Baruch College.

Academic Integrity:

Academic dishonesty is unacceptable and will not be tolerated. Cheating, forgery, plagiarism and collusion in dishonest acts undermine the college's educational mission and the students' personal and intellectual growth. Baruch students are expected to bear individual responsibility for their work and to uphold the ideal of academic integrity. Any student who attempts to compromise or devalue the academic process will be sanctioned. Please see the Baruch College Website for Further Information:

http://www.baruch.cuny.edu/academic/academic_honesty.html

Counseling and Student Health:

Students may occasionally have personal issues that arise in the course of pursuing higher education that may interfere with academic performance. If you are facing problems affecting your coursework you are encouraged to seek confidential assistance at the Baruch College Counseling Center 646–312-2158 or contact the Office of Graduate Programs 646-312-1300.

TENTATIVE SCHEDULE

Lecture	Topic	Sections
1	Introduction of Multivariate Data Analysis; Overview of Matrix Algebra; Random Vectors and Covariance Matrices	1.1-1.4, 2.1-2.11, 3.1-3.7
0.5	The Multivariate Normal Distribution	4.1-4.3
1	Multivariate Tests on One or Two Mean Vectors	5.1-5.4, 5.7
1.5	Multivariate Analysis of Variance	6.1-6.3, 6.5
1.5	Multivariate Regression	10.1, 10.2, 10.4, 10.5
1	Canonical Correlation	11.1-11.4
Midterm Exam		
1	Discriminant Analysis: Description of Group Separation	8.1-8.2, 8.4-8.6
1	Classification Analysis: Allocation of Observations to Groups	9.1-9.5, 9.7
1	Principal Component Analysis	12.1, 12.2, 12.4-12.6, 12.8
1.5	Factor Analysis	13.1-13.5
2	Cluster Analysis	14.1-14.5
Final Exam		