The George Washington University
Department of Statistics

Course: Statistics 6215.10 CRN 66978 (Applied Multivariate Analysis)
Semester: Fall 2021
Class Time: Thursday 6:10 pm-8:40 pm EST
Location: MON 115

Instructor: Prof. Reza Modarres
Campus Address: Rome Hall 760B
Office Phone: 202-994-9991
E-mail: reza@gwu.edu
Office hour: Thursday 5-6 pm, or by appointment.

Grader: TBA
Email:
Office Hours:

Course Description:
This is primarily a lecture course designed to introduce you to the statistical analysis of several variables, most likely dependent, following a joint normal distribution. Stat 6215 reworks much of the material in Stat 4157—4158 using matrices and vectors (chapters 1-6 and 8). Additional topics from the literature will also be covered. The computational aspects will include the use of R and SAS/IML.

- Matrix Algebra and Random Vectors
- Multivariate Sample Geometry
- The Multivariate Normal Distribution
- Inferences about a Mean Vector
- Comparisons of Several Population Means
- Principal component Analysis

LEARNING OUTCOMES:

1. Derive properties of the multivariate normal distribution.
2. Analyze observations obtained from a multivariate normal distribution.
3. Make inferences about the mean vector.
4. Read, analyze and synthesize further methodology not covered in class.

Class instructions: 2 hours and half weekly
Homework and reading: average of 8 hours weekly

Text: Required: Applied Multivariate Analysis, 6th Ed. by Johnson and Wichern, 2007.

Grade: Your grade will be based on:

Homework	25%
Midterm	30%
Two Quizzes	15%
Final	30%

Prerequisite: Stat 4157, 4158 and Math 2184

R and SAS will be used and the computational aspects will include use of matrix algebra tools. You are expected to be familiar with R or SAS. GW labs provide access to SAS. To obtain a copy of SAS, see https://itl.gwu.edu/sas-software-distribution.

To download R, See https://cran.r-project.org/

CLASS POLICIES

Attendance policy: The material covered and the handouts distributed during both the lecture and the lab will become available on Blackboard.

Late work: will not be accepted.

Make-up exams: Makeup examinations will only be given in exceptional circumstances (e.g. well documented medical emergency).

Homework: There will be several homework sets. Each problem counts 10 points. All graded work will be returned and discussed in class. Late submissions will not be accepted. **You are expected to work individually on each problem set.**

Homework is assigned to make sure you keep up with the work. You should do the homework and ask questions in class. I am available to answer questions during class, after class, during my office hours, by email, and by appointment. If you feel you do not understand a concept or are falling

behind, let me know as soon as possible and ask for help. I will also give you a reading list for each chapter. You need to read the material before coming to class and ask questions when I go over the material.

Quizzes and Exams: Quizzes and Exams are closed notes, but open book

Blackboard

I will upload all, handouts, examples, project descriptions, old exams, etc. on the Blackboard. I will also record each lecture so that you can access it later.

SAS: *To obtain SAS for installation on their computers, students should visit* https://itl.gwu.edu/sas-software-distribution

Please be inform that SAS can only be installed on a Windows OS.

VCL: Virtual Computer Lab (VCL) is a cloud-based service that allows you to run Windows-based software through a web browser.
VCL is available only during the virtual learning period to provide special access to certain specialized academic software while students and faculty are away from campus.

See https://acadtech.gwu.edu/vcl

To access SAS via VCL visit
https://academiccommons.gwu.edu/virtualcomputerlab

CCAS Cloud: You can also access SAS (and other software) on CCAS cloud. Visit apps.ccas.gwu.edu
Login using your cloud ID. You may have to download a software called Citrix. You will see a virtual desktop and from there you can run SAS and many other apps. The Columbian College Private Cloud and CCAS Virtual

Applications. Many software programs are available through the CCAS Virtual applications. https://ots.columbian.gwu.edu/columbian-college-private-cloud

Student Blackboard Support

GW Information Technology Support Center

Phone: 202-994-4948, option 2

Email: ithelp@gwu.edu
Submit a Support Ticket

Support hours: Monday - Friday, 7:00am - 10:00pm*

*Blackboard support is available 24 hours by calling 202-994-4948, option 2

University policies

University policy on observance of religious holidays

In accordance with University policy, students should notify faculty during the first week of the semester of their intention to be absent from class on their day(s) of religious observance. For details and policy, see "Religious Holidays" at provost.gwu.edu/policies-procedures-and-guidelines

Academic Integrity Code

Academic Integrity is an integral part of the educational process, and GW takes these matters very seriously. Violations of academic integrity occur when students fail to cite research sources properly, engage in unauthorized collaboration, falsify data, and in other ways outlined in the Code of Academic Integrity. Students accused of academic integrity violations should contact the Office of Academic Integrity to learn more about their rights and options in the process. Outcomes can range from failure of assignment to expulsion from the University, including a transcript notation. The Office of Academic Integrity maintains a permanent record of the violation. More information is available from the Office of Academic Integrity at studentconduct.gwu.edu/academic-integrity.

Virtual academic support

A full range of academic support is offered virtually in AY 21-22. See coronavirus.gwu.edu/top-faqs for updates.

Tutoring and course review sessions are offered through Academic Commons in an online format. See academiccommons.gwu.edu/tutoring
Writing and research consultations are available online. See academiccommons.gwu.edu/writing-research-help

Coaching, offered through the Office of Student Success, is available in a virtual format. See studentsuccess.gwu.edu/academic-program-support

Fall Semester 2021

Classes Begin	Monday, August 30, 2021
Labor Day (no classes)	Monday, September 6, 2021
Fall Break/Reading Day (no classes)	Friday, October 22, 2021
Thanksgiving Break (no classes)	Wednesday, Nov. 24 – Sat., Nov. 27
Designated Friday	Tuesday, December 7, 2021
Last Day of Classes Saturday,	December 11, 2021
Make-Up Day	Monday, December 13, 2021
Final Examinations	Tuesday, Dec. 14, 2021 – Wed., Dec.

Stat 6215: Tentative Schedule

Class	Date	Topic
1	Sep 2	Chapter 1: Aspects of Multivariate Analysis
2	Sep 9	Chapter 2: Matrix Algebra and Random vectors
3	Sep 16	Chapter 2: Multivariate Distributions and SVD
4	Sep 23	Chapter 2: Matrix inequalities and maximization **Quiz I**
5	Sep 30	Chapter 3: Sample Geometry and Random Samples Sample mean, covariance and correlation
6	Oct 7	Chapter 4: The Multivariate Normal Distribution
7	Oct 14	Chapter 4:Sampling Distribution of the mean and covariance
8	Oct 21	Chapter 4: Sampling from $N(0,\sum)$
9	Oct 28	**Midterm Exam 7:10-9:40 P.M.**
10	Nov 4	Chapter 5: Inference about a Mean Vector
11	Nov 11	Chapter 5
12	Nov 18	Chapter 5, **Quiz II**
	Dec 2	Chapter 8: Principal Component Analysis
13	Dec 9	Chapter 11; Review for Final
15	Dec 16	**Final Exam 6:10-8:40 P.M.**

We may also study classification and clustering techniques as time permits.

STAT 6215 OUTLINE

1. Aspects of Multivariate Analysis
   - What is multivariate analysis?
   - Differences with Univariate Analysis.
   - Founding Fathers and their contributions
   - Data visualization of multivariate Observations: Chernoff faces, Parallel plots, Andrews plot
   - SAS IML and R matrix operations
2. Matrix Algebra and Random Vectors
   - Basic properties of vectors and vector spaces, and projections
   - Linear and quadratic form.
   - Positive definite quadratic forms matrices. Spectral decomposition.
   - Rotations and orthogonal matrices. Direction cosines.
   - Sample variance as a quadratic form, geometric interpretation
   - Partitioned vectors and matrices
   - Bivariate distributions, Box-Muller Transformation Singular value Decomposition, Cholesky Decomposition
   - Matrix inequalities and maximization

3. Sample Geometry and Random Sampling
   - The geometry of the sample mean
   - Deviation vectors
   - Mean and variance of the sample mean vector
   - Generalized sample variance and total sample variance
   - Linear combination of variables
4. The Multivariate Normal Distribution
   - The multivariate normal density and its properties
   - P-dimensional Geometry, regression lines and ellipses of constant density
   - Maximum likelihood estimation
   - Properties of the sample mean and covariance from a multivariate normal distribution, large sample properties.
   - Independence of the sample mean vector and covariance matrix
   - General covariance of the mean and variance, lack of correlation and independence
   - Conditional distributions and regression
   - Wishart distribution, Sampling distribution of the correlation coefficient
   - Assessing the normality assumption
   - QQ plot, Kolmogorov-Smirinov, Shapiro-Wilk, Cramer Von Mises, and Anderson-Darling tests, Filliben test
   - Assessing bivariate and multivariate normality
   - Measures of multivariate skewness and kurtosis
   - Box-Cox transformation and its properties
5. Inferences about a Mean Vector

- Hotelling T^2 distribution and the likelihood ratio test
- Confidence regions and simultaneous comparison of component means
- Large sample inference about a population mean vector
- Multivariate Quality control charts
- Missing observation and the EM algorithm
- Power of the Hotelling T^2 distribution for large n
- Multivariate outlier detection techniques

6. Principal Component Analysis

- Population principal components
- Principal components from standardized variables
- Covariance with special structure
- Interpretation of principal components
- Number of principal components
- Large sample properties of the sample eigenvalues and eigenvectors
- Testing for equal correlation structure
- Interpoint distances, properties and applications
- Graphical comparison of groups of observations