

Mục tiêu: Trang bị các công cụ thống kê nhiều chiều nhằm xử lý các số liệu thực tế phức tạp trong kinh tế - xã hội. Các công cụ này, là mở rộng tự nhiên của thống kê cổ điển, có thể sử dụng cho nhiều mục đích như chiết suất thông tin, khai phá dữ liệu, mô hình hóa, dự báo... trong nhiều bài toán khoa học, kỹ thuật, kinh tế...

Objective: To provide the tools of multivariate statistics for treatment of practical big data. These tools, being natural extensions of respecting ones in classical statistics, can be used for many objectives as extracting information, data meaning, modeling, forecasting... in many problems of sciences, techniques, economy...

Nội dung: Học phần bao gồm các khái niệm của thống kê nhiều chiều: xử lý số liệu, miền tin cậy khoảng tin cậy đồng thời, phân tích phương sai nhiều chiều, hồi quy tuyến tính nhiều chiều, phân tích thành phần chính, phân tích nhân tố trực giao, phân lớp và phân cụm.

Content: The course includes concepts of multivariate statistics: treatment of data; confidence region; joint confidence interval; multivariate variance analysis, multivariate linear regression; principal component analysis; factor analysis; classification and clustering.

1. THÔNG TIN CHUNG

Tên học phần:	Phân tích số liệu (Data analysis)
Mã số học phần:	MI4024
Khối lượng:	2(2-1-0-4) <ul style="list-style-type: none">- Lý thuyết: 30 tiết- Bài tập/BTL: 15 tiết (nếu có bài tập lớn cần ghi rõ)- Thí nghiệm: 0 tiết
Học phần tiên quyết:	-
Học phần học trước:	- MI2020/MI3030 Xác suất thống kê, MI1141 Đại số
Học phần song hành:	-

2. MÔ TẢ HỌC PHẦN

Môn học cung cấp cách tiếp cận về phân tích thống kê dựa trên các số liệu mẫu nhiều chiều, với các ý tưởng và thuật toán tính toán tương ứng. Thống kê nhiều chiều đóng vai trò trung tâm của xử lý số liệu của nhiều vấn đề thực tiễn kinh tế và xã hội. Ngoài kỹ năng tính toán và mô hình hóa thực tế, sinh viên còn được rèn luyện phong cách làm việc nhóm, khả năng thuyết trình trong trình bày một vấn đề khoa học sau này.

3. MỤC TIÊU VÀ CHUẨN ĐẦU RA CỦA HỌC PHẦN

Sinh viên hoàn thành học phần này có khả năng:

Mục tiêu/CĐR	Mô tả mục tiêu/Chuẩn đầu ra của học phần	CĐR được phân bổ cho HP/ Mức độ (I/T/U)
[1]	[2]	[3]
M1	Khả năng hiểu biết và áp dụng các kiến thức cơ sở tìm kiếm, phân tích thông tin	1.1; 1.2; 1.3; 2.2; 4.3
M1.1	Khả năng xác định, phân tích và giải quyết các vấn đề liên quan đến thông tin và xử lý thống tin	1.1 (I) 1.2 (T)

Mục tiêu/CDR	Mô tả mục tiêu/Chuẩn đầu ra của học phần	CDR được phân bổ cho HP/ Mức độ (I/T/U)
M1.2	Khả năng phân tích dữ liệu, thiết kế, xây dựng hệ thống thông tin	1.3 (TU)
M1.3	Có tư duy tổng thể, có ý tưởng giải pháp và mô hình hóa, triển khai ứng dụng CNTT	2.2 (T) 4.3 (U)
M2	Năng lực phân tích, tìm kiếm ý tưởng mới, thiết kế, triển khai và điều hành hệ thống thông tin lớn	1.3; 4.1; 4.3
M2.1	Tham gia thiết kế, xây dựng, triển khai và điều hành hệ thống thông tin	1.3 (T)
M2.2	Tìm kiếm ý tưởng, thiết kế và xây dựng, cài đặt hệ thống xử lý thông tin cụ thể và cải tiến	4.1; 4.3 (TU)
M3	Kỹ năng phát hiện vấn đề, thích ứng nhanh chóng với thực tế, khả năng tự nghiên cứu	2.4; 4.1; 4.3
M3.1	Năng lực phát hiện vấn đề, phân tích và khai thác thông tin	4.1; 4.3 (TU)
M3.2	Kỹ năng tự học, tự nghiên cứu để thích ứng nhanh với sự phát triển của khoa học và công nghệ	2.4 (U)

4. TÀI LIỆU HỌC TẬP

Giáo trình:

[1] Nguyễn Hữu Dư, Nguyễn Văn Hữu, *Thống kê nhiều chiều*, Nhà XB Đại học Quốc gia, Hà Nội, 2004

Sách tham khảo:

[1] Johnson R.A. and D.W. Wichern, *Applied Multivariate Statistical Analysis*, Prentice Hall, 1998

[2] Cressie N, *Statistics for Spatial Data*, Wiley, N.Y., 1993

[3] Hardle W. and L. Simar, *Applied Multivariate Statistics Analysis*, MD*Tech, Berlin, 2003

5. CÁCH ĐÁNH GIÁ HỌC PHẦN

Điểm thành phần	PP đánh giá cụ thể	Mô tả	CDR được đánh giá	Tỷ trọng
[1]	[2]	[3]	[4]	[5]
A1. Điểm quá trình	Đánh giá quá trình A1.1 Thảo luận và bài tập A1.2 Bài tập nhóm	Thuyết trình Báo cáo	M1.1; M1.2; M2.2; M1.3; M2.2; M3.1	30%
A2. Điểm cuối kỳ	A2.1 Thi cuối kỳ	Thi vấn đáp/ Viết/BTN	M1-M3	70%

* Điểm quá trình sẽ được điều chỉnh bằng cách cộng thêm điểm chuyên cần. Điểm chuyên cần có giá trị từ -2 đến +2, theo Quy chế Đào tạo đại học hệ chính quy của Trường ĐH Bách khoa Hà Nội.

6. KẾ HOẠCH GIẢNG DẠY

Tuần	Nội dung	CDR học phần	Hoạt động dạy và học	Bài đánh giá
[1]	[2]	[3]	[4]	[5]
1	Chương 0: Cơ sở <ul style="list-style-type: none"> - Ma trận và vecto; các khái niệm Ma trận xác định dương, ma trận căn hai Phân tích phổ Khoảng cách tổng quát - Ma trận và vecto ngẫu nhiên 	M1.1 M1.2 M3.1	Giảng bài Đọc tài liệu	A2.1
2	<ul style="list-style-type: none"> - Các bất đẳng thức ma trận Bất đẳng thức Cauchy – Schwarz Bổ đề cực đại hóa Chương 1: Mẫu thống kê nhiều chiều <ul style="list-style-type: none"> - Ma trận mẫu, mẫu ngẫu nhiên 	M1.1	Giảng bài Bài tập Đọc tài liệu	A1.1-2 A2.1
3	<ul style="list-style-type: none"> - Phương sai mẫu tổng quát - Phân phối chuẩn nhiều chiều Mật độ giá trị mật độ Phân phối có điều kiện 	M1.1	Giảng bài Bài tập Đọc tài liệu	A1.1-2 A2.1
4	<ul style="list-style-type: none"> - Hàm hợp lý và ước lượng hợp lý cực đại Phân phối Wishart Trường hợp mẫu lớn - Kiểm định tính phân phối chuẩn 	M1.1	Giảng bài Bài tập Đọc tài liệu	A1.1-2 A2.1
5	Chương 2: Phân tích các vecto trung bình <ul style="list-style-type: none"> - Kiểm định giả thuyết về trị trung bình Thống kê Hotelling - Kiểm định tỷ số hợp lý Thống kê Lambda Wilk 	M1.1 M3.1	Giảng bài Bài tập	A1.1-2 A2.1
6	<ul style="list-style-type: none"> - Miền tin cậy Khoảng tin cậy đồng thời Khoảng tin cậy Bonferroni - Trường hợp phục hồi quan sát thiếu - So sánh vecto trung bình So sánh cặp; so sánh dùng quan sát lặp So sánh nhiều vecto trung bình MANOVA 	M1.1 M1.2 M3.1	Giảng bài Bài tập Thực hành	A1.1-2 A2.1
7	Chương 3: Mô hình hồi quy tuyến tính nhiều chiều <ul style="list-style-type: none"> - Mô hình hồi quy tuyến tính bội - Ước lượng bình phương cực tiểu 	M1.1 M1.2 M3.1	Giảng bài Bài tập Đọc tài liệu	A1.1-2 A2.1
8	<ul style="list-style-type: none"> - Phân tích mô hình hồi quy bội - Kiểm định mô hình hồi quy bội - Mô hình hồi quy tuyến tính nhiều chiều Kiểm định tỷ số hợp lý cho các tham số hồi quy Dự báo bằng hồi quy tuyến tính nhiều chiều	M1.1 M1.2 M3.1	Giảng bài Bài tập Đọc tài liệu	
9	Kiểm tra giữa kỳ - Giao bài tập lớn	M1.1	Giảng bài	A1.1-2

Tuần	Nội dung	CDR học phần	Hoạt động dạy và học	Bài đánh giá
[1]	[2]	[3]	[4]	[5]
	Chương 4: Phân tích thành phần chính - Thành phần chính lý thuyết Thành phần chính với các biến quy chuẩn Thành phần chính của ma trận hiệp phương sai có cấu trúc đặc biệt	M2.2	Bài tập	A2.1
10	- Thành phần chính mẫu Số lượng thành phần chính Giải thích thành phần chính Quy chuẩn các thành phần chính mẫu	M1.2 M2.2 M3.2	Giảng bài Bài tập Thực hành	A1.1-2 A2.1
11	Chương 5: Phân tích nhân tố - Mô hình nhân tố trực giao Nhân tố chung và nhân tố đặc biệt Cấu trúc hiệp phương sai của mô hình nhân tố trực giao	M1.1 M2.2 M3.1	Giảng bài Bài tập Đọc tài liệu	A1.1-2 A2.1
12	- Phương pháp ước lượng tham số Phương pháp thành phần chính Phương pháp hợp lý cực đại - Các tiếp cận cải biên ước lượng tham số - Phép quay nhân tố	M1.1 M2.2 M3.1	Giảng bài Bài tập	A1.1-2 A2.1
13	Chương 6: Phân lớp và phân cụm - Phân lớp cho 2 tập nền (có và không có phân phối chuẩn) - Hàm phân lớp Fisher - Quy tắc nhân lớp cho nhiều tập nền	M2.2 M3.1	Giảng bài Bài tập	A1.1-2 A2.1
14	- Hàm phân lớp Fisher cho nhiều tập nền - Độ đo khoảng cách và phân cụm - Các phương pháp phân cụm phân tầng	M2.2 M3.2	Giảng bài Bài tập	A1.1-2 A2.1
15	- Các phương pháp phân cụm không phân tầng - Một số thuật toán thực hành	M1.2 M2.2 M3.2	Giảng bài Bài tập Thực hành	A1.1-2 A2.1
16	<i>Tổng ôn và giải đáp</i>			

7. QUY ĐỊNH CỦA HỌC PHẦN

(Các quy định của học phần nếu có)

Tuân thủ quy định học tập trong Quy chế đào tạo của ĐHBK Hà Nội.

8. NGÀY PHÊ DUYỆT:

Chủ tịch Hội đồng

Nhóm xây dựng đề cương

9. QUÁ TRÌNH CẬP NHẬT

