ĐẠI HỌC QUỐC GIA TP. HỒ CHÍ MINH

TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN

ĐỀ CƯƠNG MÔN HỌC

XỬ LÝ DỮ LIỆU THỐNG KÊ

1. THÔNG TIN CHUNG (General information)

Tên môn học (tiếng Việt):	Xử lý Dữ liệu Thống kê
Tên môn học (tiếng Anh):	Analysis of Statistical Data
Mã môn học:	IE201
Thuộc khối kiến thức:	Đại cương ☐; Cơ sở nhóm ngành ☐;
	Cơ sở ngành ☐; Chuyên ngành ☑; Tốt nghiệp ☐
Khoa, Bộ môn phụ trách:	Khoa học và Kỹ thuật Thông tin
Giảng viên biên soạn:	PGS. TS. Đỗ Phúc
	Email: phucdo@uit.edu.vn
Số tín chỉ:	3
Lý thuyết:	3
Thực hành:	0
Tự học:	6
Môn học tiên quyết:	Không
Môn học trước:	Xác suất Thống kê
	Giải tích 1

2. MÔ TẢ MÔN HỌC (Course description)

Môn học nhằm trình bày những phương pháp sơ cấp cơ bản để xử lý, mô tả, phân tích, phán đoán, các dữ liệu thống kê, cùng với cơ sở lí luận xác suất của các phương pháp đó.

3. MỤC TIÊU MÔN HỌC (Course goals)

Sau khi hoàn thành môn học này, sinh viên có thể:

Bảng 1.

Mục tiêu môn học
Những phương pháp sơ cấp cơ bản để xử lý, mô tả, phân tích, phán đoán, các dữ liệu thống kê, cùng với cơ sở lí luận xác suất của các phương pháp đó.
Vận dụng những kiến thức đã học vào việc phân tích các dữ liệu thống kê.
Sau khi học môn này, sinh viên sẽ đam mê phân tích dữ thống kê, có khả năng nghiên cứu phân tích dữ liệu thống kê.

4. CHUẨN ĐẦU RA MÔN HỌC (Course learning outcomes)

Bảng 2.

CĐRMH	Mô tả CĐRMH (Mục tiêu cụ thể)
G1 (LO 1, LO 2)	Những phương pháp sơ cấp cơ bản để xử lý, mô tả, phân tích, phán đoán, các dữ liệu thống kê, cùng với cơ sở lí luận xác suất của các phương pháp đó.
G2 (LO 5)	Vận dụng những kiến thức đã học vào việc phân tích các dữ liệu thống kê.
G3 (LO 5)	Sau khi học môn này, sinh viên sẽ đam mê phân tích dữ thống kê, có khả năng nghiên cứu phân tích dữ liệu thống kê.

5. NỘI DUNG MÔN HỌC, KẾ HOẠCH GIẢNG DẠY (Course content, lesson plan)

Bảng 3.

Buổi học (45 tiết)	Nội dung	CĐRMH	Thành phần đánh giá
Buổi 1 – 5 (15 tiết)	**Chương 1: Mô tả đầy đủ các phân phối** 1. Quan trắc và rút mẫu có hoàn lại 2. Phân phối tổng quát và phân phối mẫu 3. Định lý *Glivenkod-Cantellli* 4. Biểu đồ tổ chức và liên hệ với hàm mật độ lý thuyết	G1	A2, A4
Buổi 6 – 8 (9 tiết)	**Chương 2: Mô tả đặc trưng của phân phối** 1. Định vị phân phối 2. Độ phân tán phân phối, độ lệch và độ nhọn của phân phối, các vấn đề về phân phối nhiều chiều	G1	A2, A4
Buổi 9 – 12	**Chương 3: Phân tích tương quan và hồi quy** 1. Hồi quy của kỳ vọng	G2	A4

(12 tiết)	2. Hồi quy tuyến tính bội 3. Phân tích tương quan		
Buổi 13 – 15 (9 tiết)	**Chương 4: Tiêu chuẩn chi-bình phương** 1. Phương pháp kiểm định tính chuẩn của một phân phối xác suất 2. Phương pháp kiểm định tính độc lập, tính thuần nhất	G2	A4

6. ĐÁNH GIÁ MÔN HỌC (Course assessment)

Bảng 4.

Thành phần đánh giá	CĐRMH	Tỷ lệ (%)
A1. Quá trình (Kiểm tra trên lớp, bài tập, đồ án, …)		0%
A2. Giữa kỳ	G1, G2	30%
A3. Thực hành		0%
A4. Cuối kỳ	G1, G2	70%

7. QUY ĐỊNH CỦA MÔN HỌC (Course requirements and expectations)

- Giảng viên cung cấp nội dung bài giảng trên lớp, sinh viên chủ động học tập và trao đổi với nhau và với giảng viên.
- Giảng viên cung cấp chủ đề cho sinh viên tìm hiểu ở nhà và theo nhóm, sinh viên trình bày tại lớp và thảo luận.
- Sinh viên cần chủ động tự tìm hiểu thêm từ các nguồn tài liệu khác, nhất là trên Web.

8. TÀI LIỆU HỌC TẬP, THAM KHẢO

Giáo trình

1. Prof. Cynthia Rudin (Fall 2011). *Statistical Thinking and Data Analysis*. MIT Course.
2. Nguyễn Bác Văn (1997). *Xác suất và xử lý số liệu thống kê*. NXB Giáo dục.

Tài liệu tham khảo

1. Nguyễn Bác Văn (1997). *Mở đầu thống kê xác suất*. NXB Giáo dục.
2. Tamhane, Ajit C., and Dorothy D. Dunlop (1999). *Statistics and Data Analysis: From Elementary to Intermediate*. Prentice Hall.

9. PHẦN MỀM HAY CÔNG CỤ HỖ TRỢ THỰC HÀNH

1. Microsoft (2010). *Visual Studio 2010*.
2. Microsoft (2008). *Microsoft SQL Server Management Studio*.

Tp.HCM, ngày 03 tháng 03 năm 2016

Trưởng khoa/bộ môn	**Giảng viên biên soạn**
(Ký và ghi rõ họ tên)	(Ký và ghi rõ họ tên)
TS. Nguyễn Gia Tuấn Anh	PGS. TS. Đỗ Phúc