

## CHƯƠNG TRÌNH HỌC PHẦN

### 1. Thông tin chung về học phần

Tên học phần: Thống kê máy tính

Mã học phần:

Số tín chỉ: 3

Học phần tiên quyết: Xác suất thống kê

Đào tạo trình độ: Đại học

Giảng dạy cho ngành: Công nghệ Thông tin

Bộ môn quản lý: Bộ môn Hệ thống thông tin

Phân bổ tiết giảng của học phần:

- Nghe giảng lý thuyết: 30 tiết
- Thực hành: 15 tiết (yêu cầu phòng máy có cài đặt phần mềm R, MS Office)
- Tự nghiên cứu: 90 tiết

### 2. Mô tả tóm tắt học phần

Học phần trang bị cho người học kiến thức về phương pháp luận, các công cụ toán học và khoa học máy tính phục vụ cho việc xử lý số liệu bằng máy tính điện tử. Người học được trang bị phương pháp, kỹ năng thu thập, phân tích số liệu đưa ra các dự đoán hỗ trợ quyết định dựa vào các phần mềm xử lý thống kê chuyên dụng tiên tiến, ví dụ phần mềm R.

### 3. Chủ đề và chuẩn đầu ra của học phần

#### 3.1. Danh mục vấn đề của học phần

1. Thu thập dữ liệu
2. Phần mềm nguồn mở R và ngôn ngữ R cho phân tích số liệu và thống kê.
3. Sử dụng phần mềm R giải quyết các vấn đề trong thống kê.
4. Sử dụng R trong bài toán phân lớp, phân cụm, xử lý dữ liệu thiếu.

#### 3.2. Chuẩn đầu ra của quá trình dạy - học từng chủ đề của học phần

**Chủ đề 1:** Thu thập dữ liệu

Nội dung	Mức độ
<b>Kiến thức</b> 1. Giới thiệu thống kê máy tính 2. Lấy mẫu ngẫu nhiên đơn giản 3. Lấy mẫu ngẫu nhiên phân tầng 4. Xác định dữ liệu nhiễu (outlier) trong tập mẫu 5. Sử dụng các công cụ trực quan: Histogram, quantile plot, box plot	2
<b>Thái độ</b> 1. Thống kê liên quan đến việc trích xuất kiến thức từ một tập dữ liệu thô. 2. Lấy mẫu ngẫu nhiên: Chọn ngẫu nhiên một cá thể từ quần thể hoặc phân tầng và áp dụng lấy mẫu đơn giản cho từng tầng. 3. Phân loại dữ liệu và áp dụng các công cụ thích hợp để loại bỏ nhiễu.	

<b>Kỹ năng</b> 1. Phân biệt giữa thống kê truyền thống và thống kê máy tính. 2. Áp dụng thành thạo các phương pháp lấy mẫu. 3. Sử dụng thành thạo các công cụ hỗ trợ lọc bỏ nhiễu.	3
---	---

**Chủ đề 2:** Phần mềm nguồn mở R và ngôn ngữ R cho phân tích số liệu và thống kê

Nội dung	Mức độ
<b>Kiến thức</b> 1. Cú pháp ngôn ngữ R 2. Nhập dữ liệu trong R 3. Sử dụng R cho các tính toán đơn giản 4. Sử dụng R cho các phép tính ma trận 5. Tính toán xác suất và mô phỏng. 6. Kiểm định giả thuyết và trị số R 7. Phân tích số liệu bằng biểu đồ	3
<b>Thái độ</b> 1. Phần mềm R là phần mềm mã nguồn mở hiện đang được sử dụng rộng rãi trong phân tích và xử lý số liệu. 2. Phần mềm R được sử dụng thông qua giao diện dòng lệnh (gồm các lệnh và các hàm)	
<b>Kỹ năng</b> 1. Sử dụng R thành thạo cho các tính toán căn bản 2. Sử dụng tốt R để tính xác suất, kiểm định giả thuyết, phân tích số liệu	3

**Chủ đề 3:** Sử dụng phần mềm R giải quyết các vấn đề trong thống kê.

Nội dung	Mức độ
<b>Kiến thức</b> 1. Phân tích thống kê mô tả 2. Phân tích hồi qui tuyến tính 3. Phân tích phương sai 4. Phân tích hồi qui logistic 5. Phân tích biến cố 6. Phân tích tổng hợp 7. Ước tính cỡ mẫu	3
<b>Thái độ</b> Các mục kiến thức từ 1 đến 7 là các vấn đề cơ bản trong thống kê. Sử dụng R để giải quyết các vấn đề này, từ đó có thể tự tìm hiểu R để giải quyết các bài toán phân tích khác	
<b>Kỹ năng</b> 1. Áp dụng thành thạo phân tích số liệu bằng biểu đồ 2. Sử dụng R để giải quyết được các vấn đề trong thống kê	3

**Chủ đề 4:** Sử dụng R trong bài toán phân lớp, phân cụm, xử lý dữ liệu thiếu.

Nội dung	Mức độ
<b>Kiến thức</b>	2

<ol style="list-style-type: none"> <li>1. Các phương pháp xử lý dữ liệu thiếu.</li> <li>2. Phân lớp dữ liệu với R</li> <li>3. Phân cụm dữ liệu với R</li> </ol>	
<b>Thái độ</b> <ol style="list-style-type: none"> <li>1. Thực tế dữ liệu thường không đầy đủ</li> <li>2. Ngoài các tác vụ thống kê, R có thể xử giải quyết các bài toán khác như phân lớp, phân cụm.</li> </ol>	
<b>Kỹ năng</b> <ol style="list-style-type: none"> <li>1. Thực hiện một tác vụ phân lớp, phân cụm, xử lý dữ liệu thiếu với sự hỗ trợ của R</li> <li>2. Lập trình và hàm với R</li> </ol>	3

#### 4. Phân bổ thời gian chi tiết

Vấn đề	Phân bổ số tiết cho hình thức dạy - học					Tổng
	Lên lớp			Thực hành, thực tập	Tự nghiên cứu	
	Lý thuyết	Bài tập	Thảo luận			
Vấn đề 1	5					5
Vấn đề 2	10			4		14
Vấn đề 3	10			8		18
Vấn đề 4	5			3		8

#### 5. Tài liệu

TT	Tên tác giả	Tên tài liệu	Năm xuất bản	Nhà xuất bản	Địa chỉ khai thác tài liệu
1	Geof H. Givens and Jennifer A. Hoeting	Computational Statistics	2009	Wiley	Online
2	Nguyễn Văn Tuấn	Phân tích số liệu và biểu đồ bằng R	2006	NXB Đại học Quốc gia TP HCM	Thư viện
3	Peter Dalgaard	Introductory Statistics with R, second edition	2008	Springer	Thư viện
4	Robert I.Kabacoff	R In Action	2011	Manning	Thư viện

#### 6. Các hoạt động đánh giá

TT	Các chỉ tiêu đánh giá	Phương pháp đánh giá	Trọng số (%)
1	Tham gia học trên lớp: chuẩn bị bài tốt, tích cực thảo luận...	Quan sát, điểm danh	10
2	Tự nghiên cứu: hoàn thành nhiệm vụ giảng viên giao trong tuần, bài tập nhóm/tháng/học kỳ...	Chấm báo cáo, bài tập...	20
4	Kiểm tra đánh giá cuối kỳ	Viết	20
5	Thi kết thúc học phần	Viết, vấn đáp,	50

		tiểu luận....	
--	--	---------------	--

**TRƯỞNG KHOA**

**TRƯỞNG BỘ MÔN**