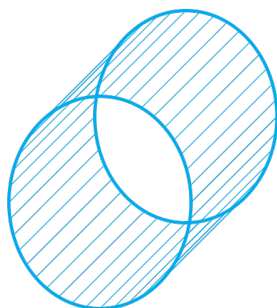


TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN TP.HCM  
CAO HỌC KHÓA 30

—\*—



Khoa Toán - Tin học  
Fac. of Math. & Computer Science

**Bài tập lần 2**  
**MÔ HÌNH HÓA THỐNG KÊ**

Giảng viên hướng dẫn: **TS. Nguyễn Thị Mộng Ngọc**

Nhóm thực hiện: **Nhóm 4**

TP. Hồ Chí Minh – Tháng 01, 2021

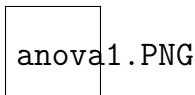
**BẢNG PHÂN CÔNG CÔNG VIỆC**

Thành viên	Công việc	Mã số học viên
1. Đặng Khánh Thi	— — —	20C29038
2. Đinh Thị Nữ	— — —	20C29013
3. Lý Phi Long	— — —	20C29028
4. Phan Thị Thùy An (Nhóm trưởng)	— —	20C29002

# BÀI 1

- $X_1$ : áp lực công việc
- $X_2$ : kỹ năng quản lý
- $X_3$ : mức độ hài lòng với chức vụ của mình
- $Y$ : mức độ lo lắng (biến phụ thuộc)

Bảng ANOVA:



1. Tính tổng bình phương hồi quy trên  $X_1, X_2$  và  $X_3$ ?

$$SSR = SSR_{X_1} + SSR_{X_2|X_1} + SSR_{X_3|X_1, X_2} = 981.326 + 190.232 + 129.431 = 1299.989$$

2. Xác định tỷ lệ phần trăm sự biến thiên của mức độ lo lắng được giải thích bởi các biến độc lập.

3. Có thể kết luận rằng trong tất cả ba biến giải thích đều có ảnh hưởng đáng kể đến mức độ lo lắng hay không? Chỉ rõ kiểm định nào được dùng.

4. Nếu chúng ta chỉ xét biến giải thích  $X_1$ , hãy lập bảng ANOVA ?

5. Kiểm định giả thiết sau với mức ý nghĩa 5%

(a)  $H_0 : \beta_1 = 0$  cho mô hình  $Y = \beta_0 + \beta_1 X_1 + \epsilon$

(b)  $H_0 : \beta_2 = 0$  cho mô hình  $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon$

(c)  $H_0 : \beta_3 = 0$  cho mô hình  $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \epsilon$

6. Xác định hệ số xác định cho mỗi mô hình trong câu 5.

7. Trong các mô hình trên, mô hình nào thích hợp nhất để giải thích sự biến động mức độ lo lắng của các giám đốc ?

## BÀI 2

- $Y$  : mức độ bền dẻo của nhựa
- $X_1$ : độ dày của vật liệu
- $X_2$ : mật độ của vật liệu

1. Tìm 2 phương trình đường thẳng hồi quy và 1 phương trình siêu phẳng (nếu có) ?
2. Xác định tỷ lệ phần trăm sự biến thiên của biến phụ thuộc cho từng mô hình có thể có trên.
3. Nếu chúng ta chỉ quan tâm đến cả 2 biến giải thích, hãy lập bảng ANOVA?
4. Kiểm định giả thiết sau với mức ý nghĩa 5%

$$H_0 : \beta_1 = \beta_2 = 0$$

5. Xác định khoảng tin cậy với mức ý nghĩa 5% cho  $\beta_1$  trong trường hợp mô hình chỉ có biến độc lập là độ dày của vật liệu ( $X_1$ ).
6. Với khoảng tin cậy vừa tìm được ở câu 5, chúng ta có thể khẳng định rằng hồi quy tuyến tính là có ý nghĩa giữa mức độ bền dẻo của nhựa và độ dày của vật liệu và mật độ của vật liệu không? Chứng minh điều khẳng định của bạn.

## BÀI 3

1. Viết các mô hình tuyến tính với 2 biến độc lập (có thể).
2. Ước lượng các hệ số hồi quy trong từng mô hình tuyến tính ở câu 1.
3. Với độ tin cậy 95%, tìm khoảng tin cậy cho các tham số trong mô hình với 2 biến độc lập  $x_1$  và  $x_2$ .
4. Xác định hệ số xác định cho mỗi mô hình trong câu 1.
5. Trong các mô hình trên, mô hình nào thích hợp nhất để giải thích sự biến thiên của  $Y$  ?
6. Viết mô hình tuyến tính dưới dạng ma trận với số biến độc lập nhiều nhất có thể, và xác định kích thước của ma trận.
7. Ước lượng các hệ số hồi quy trong mô hình tuyến tính ở câu 6.
8. Trong mô hình tuyến tính ở câu 6, tính ước lượng của  $V(\epsilon)$  và  $V(\hat{\beta})$ .
9. Với độ tin cậy 95%, tìm khoảng tin cậy cho  $V(\epsilon)$ .
10. Khi thêm 2 biến độc lập  $x_3$  và  $x_2$  vào mô hình chỉ với 1 biến độc lập  $x_1$  thì làm cho chất lượng ước lượng cao hơn không?