

TRƯỜNG ĐẠI HỌC BÁCH KHOA TP. HCM  
Khoa KTXD - Bộ môn Kỹ thuật & Quản lý Tài nguyên nước

# PHÂN TÍCH DỮ LIỆU



Giảng viên: PGS. TS. NGUYỄN THÔNG

E-mail: [nguyenthong@hcmut.edu.vn](mailto:nguyenthong@hcmut.edu.vn) or [nthong56@yahoo.fr](mailto:nthong56@yahoo.fr)

Web: <http://www4.hcmut.edu.vn/~nguyenthong/index>

TÉL. (08) 38 691 592- 098 99 66 719

PHÂN TÍCH DỮ LIỆU  
Chương 8: Hồi quy tuyến tính

Chương 1. Thống kê mô tả (ôn).  
 Chương 1bis. Xác suất & phân phối thống kê (ôn).  
 Chương 2. Khoảng tin cậy.  
 Chương 2bis: Kỹ thuật chọn mẫu.  
 Chương 3. Kiểm định thống kê.  
 Chương 4. Phân loại dữ liệu (Classification).  
 Chương 5. Phân nhóm dữ liệu (Cluster).  
 Chương 6. Phân tích thành phần chính (PCA).  
 Chương 7. Phân tích chuỗi thời gian.  
 Chương 8. **Hồi quy tuyến tính**.  
 Chương 9. Xử lý số liệu thực nghiệm.  
 Chương 10. Giới thiệu phần mềm SPSS or R.

PGS. TS. Nguyễn Thông

PHÂN TÍCH DỮ LIỆU  
Chương 8: Hồi quy tuyến tính

## GIỚI THIỆU HỒI QUY

Đây là phương pháp nghiên cứu (định lượng) nhằm giải thích một biến **ĐINH LƯỢNG** nhờ vào 1 hoặc nhiều biến **ĐINH LƯỢNG** dưới dạng một phương trình toán học.

9/4/2018 PGS. Dr. Nguyễn Thông 3

PHÂN TÍCH DỮ LIỆU  
Chương 8: Hồi quy tuyến tính

→ Để có thể biểu diễn hồi quy **CÓ NGHĨA** → Các biến phải có tính **NHÂN QUẢ** (có ý nghĩa liên quan kinh tế, xã hội, vật lý,...).

→ Lưu ý, phân tích phương sai dùng xem xét quan hệ biến giữa biến giải thích dạng **ĐINH TÍNH** & biến cần được giải thích dạng **ĐINH LƯỢNG**.

9/4/2018 PGS. Dr. Nguyễn Thông 4

PHÂN TÍCH DỮ LIỆU  
Chương 8: Hồi quy tuyến tính

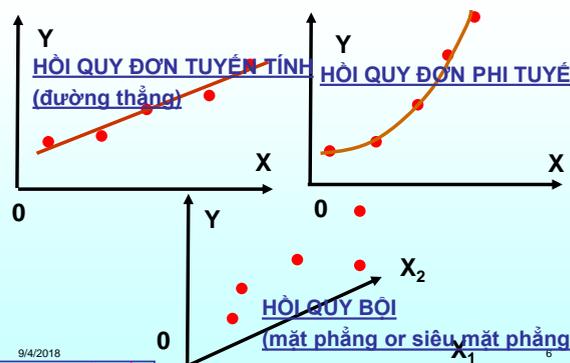
## TỔNG QUÁT

**HỒI QUY TUYẾN TÍNH**  
**HỒI QUY PHI TUYẾN**




9/4/2018 PGS. Dr. Nguyễn Thông 5

PHÂN TÍCH DỮ LIỆU  
Chương 8: Hồi quy tuyến tính



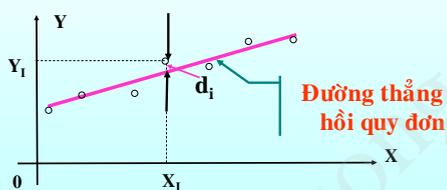
9/4/2018 PGS. Dr. Nguyễn Thông 6

## HỒI QUY TUYẾN TÍNH

- Hồi quy tuyến tính đơn
- Hồi quy tuyến tính bội

## HỒI QUY TUYẾN TÍNH ĐƠN

## HỒI QUY TUYẾN TÍNH ĐƠN



Mô hình được mô tả bởi phương trình tuyến tính có dạng sau:

$$y_t = a_0 + a_1 X_{1,t} + \varepsilon_t$$

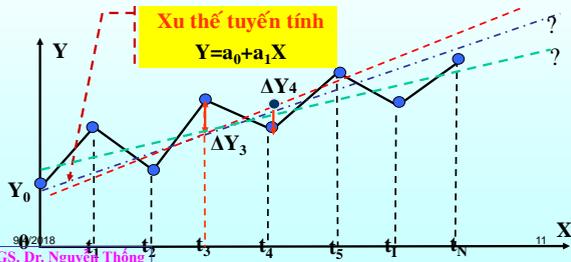
$t=1 \rightarrow N$  (số lượng quan trắc)

$y_t$  giá biến cần nghiên cứu tại thời điểm  $t$ ,  $x_{1,t}$  giá trị biến giải thích thứ  $i$  tại thời điểm  $t$ .

$N$  số lượng quan trắc các biến.

$a_0, a_1$ : tham số mô hình sẽ được xác định từ số liệu quan sát  
 $\varepsilon_t$ : sai số của mô hình.

## PHƯƠNG TRÌNH HỒI QUY TUYẾN TÍNH ĐƠN



Gọi  $\Delta y_i$  là khoảng cách thẳng đứng từ giá trị quan sát  $(x_i, y_i)$  đến đường thẳng cần xác định. Ta định nghĩa hàm mục tiêu:

$$D = \sum_{i=1}^N \Delta Y_i^2 = \sum_{i=1}^N [Y_i - (a_0 + a_1 X_i)]^2 \Rightarrow \min$$

Quan sát

Hồi quy

Đây là một hàm 2 biến  $a_0$  và  $a_1$ , để cho D cực trị (với ý nghĩa vật lý của bài toán ta biết đó là cực tiểu) ta phải có:

$$\begin{cases} \frac{\partial D}{\partial a_0} = 0 & (1) \\ \frac{\partial D}{\partial a_1} = 0 & (2) \end{cases}$$

Từ đó:  $\sum_i 2[y_i - (a_0 + a_1 x_i)] = 0$  [1]

$$\sum_i 2[y_i - (a_0 + a_1 x_i)] x_i = 0$$
 [2]

Giải hệ phương trình trên ta có:

$$a_1 = \frac{\sum_i x_i y_i - N \bar{x}_i \bar{y}_i}{\sum_i x_i^2 - N \bar{x}_i^2}$$

$$a_0 = \bar{y}_i - a_1 \bar{x}_i$$

### HÀM TÍNH a, b TRONG EXCEL

→ Intercept(Y,X) → tính tung độ gốc b của quan hệ tuyến tính.

→ Linest(Y,X) → tính hệ số a (độ dốc của quan hệ tuyến tính giữa  $Y_i$  &  $X_i$ ).

### PHƯƠNG TRÌNH PHƯƠNG SAI

$$SST = \sum_{i=1}^N (y_i - \bar{y})^2 \rightarrow \text{Phương sai tổng (bản chất số liệu)}$$

Sum of Square Total

$$SSE = \sum_{i=1}^N (\hat{y}_i - \bar{y})^2 \rightarrow \text{Phương sai mô hình (chất lượng mô hình)}$$

Sum of Square Explicative

$$SSR = \sum_{i=1}^N (y_i - \hat{y}_i)^2 \rightarrow \text{Phương sai thặng dư (sai số mô hình)}$$

Sum of Square Residual

$$SST = SSE + SSR$$
 [1]

[1] → phương trình phương sai

$y_i$ : giá trị quan sát,  $\hat{y}_i$ : giá trị tính từ mô hình

$\bar{y}$ : giá trị trung bình biến nghiên cứu

## KIỂM ĐỊNH & ĐÁNH GIÁ CHẤT LƯỢNG MÔ HÌNH HỒI QUY



## CHẤT LƯỢNG

### Hệ số xác định $R^2$

$$R^2 = \frac{SSE}{SST} = 1 - \frac{SSR}{SST} = \frac{\sum_i (y_i - \hat{y})^2}{\sum_i (y_i - \bar{y})^2} \in [0 \rightarrow 1]$$

$R^2 \rightarrow 1 \rightarrow$  mô hình càng **TỐT**

### Hệ số xác định $R^2$ hiệu chỉnh:

Khi số liệu quan sát  $n$  nhỏ, giá trị  $R^2$  được hiệu chỉnh như sau ( $k=1$ : hồi quy đơn):

$$\bar{R}^2 = 1 - \frac{n-1}{n-k-1} (1 - R^2)$$

Chú ý: Khi  $n \rightarrow$  lớn  $\rightarrow$

$$\bar{R}^2 \cong R^2$$

$$\bar{R}^2 \leq R^2$$

## KIỂM ĐỊNH HỆ SỐ $a_i$ MÔ HÌNH

$a_i$  là các giá trị trung bình thống kê  $\rightarrow$

Kiểm định  $a_i$  là khác 0 hay bằng 0 theo quan điểm thống kê.

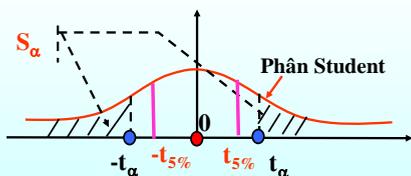
## KIỂM ĐỊNH HỆ SỐ $a_i$ MÔ HÌNH

$a_i \neq 0 \Rightarrow$  biến độc lập liên kết  $X_i$  có nghĩa trong phương trình hồi quy.

$a_i = 0 \Rightarrow$  biến  $X_i$  KHÔNG CÓ trong phương trình hồi quy.

$\rightarrow$  Kiểm định Student

## GIẢI THÍCH KIỂM ĐỊNH STUDENT (kiểm định các hệ số $a_i$ của mô hình và 0) BẰNG ĐỒ THỊ



- Nếu  $S_\alpha > 5\% \rightarrow t_\alpha = 0$  theo quan điểm thống kê.
- Nếu  $S_\alpha < 5\% \rightarrow t_\alpha$  khác 0 theo quan điểm thống kê.

$\rightarrow$  Gọi  $\sigma_{a1}$   $\rightarrow$  giá trị độ lệch chuẩn của hệ số  $a1$ :

Gọi:

$$A_1 = \frac{SSR}{(n-k-1)} \quad \text{Số quan sát}$$

$$A_2 = \sum_i X_i^2 - n\bar{X}^2 \quad \text{Số biến độc lập (hồi quy đơn } k=1)$$

$$\sigma_{a1} = \sqrt{\frac{A_1}{A_2}} \quad \text{Giá trị TB của } X_i$$

Goi:

$\sigma_{a0}$  → giá trị độ lệch chuẩn của hệ số  $a_0$ :

$$\sigma_{a0} = \sqrt{\frac{\sum X_i^2}{n} * \frac{A_1}{A_2}}$$

25

## VÍ DỤ

ANOVA				
	df	SS	MS	F
Regression	1	67.10113636	67.10114	12.8228
Residual	3	15.69886364	5.232955	
Total	4	82.8		
Coefficients				
	Coefficients	Standard Error	t Stat	P-value
Intercept	-13.0227273	8.390889392	-1.552008	0.21846
X	13.8068182	3.855690358	3.580894	0.03726

9/4/2018

PGS. Dr. Nguyễn Thông

## TÍNH TỰ TƯƠNG QUAN

$$\varepsilon_i(t) = \alpha\varepsilon_{i-1}(t) + \beta\varepsilon_{i-2}(t) + \gamma \quad \text{with } \alpha, \beta \neq 0$$

→  $\varepsilon$  có tính TỰ TƯƠNG QUAN BẬC 2

$\varepsilon_{i-1}$  Chuỗi trễ bậc 1 của  $\varepsilon_i$

i	0	1	2	3
$\varepsilon_i$	4	6	7	3
$\varepsilon_{i-1}$		4	6	7
$\varepsilon_{i-2}$			4	6

9/4/2018

PGS. Dr. Nguyễn Thông

27

## TỰ TƯƠNG QUAN

Tính tự tương quan bậc 1 →

$$\varepsilon_i = \alpha\varepsilon_{i-1} + \gamma$$

Với  $\alpha \rightarrow$  khác 0 → chuỗi  $\varepsilon_i$  có tính TỰ TƯƠNG QUAN BẬC 1 → MÔ HÌNH CHƯA TỐT → THÊM (THAY BIẾN) GIẢI THÍCH ?

9/4/2018

PGS. Dr. Nguyễn Thông

28

Xét hồi quy tuyến tính bội bậc k:

$$Y_i = a_1X_{1,i} + a_2X_{2,i} + \dots + a_kX_{k,i} + \varepsilon_i$$

Biến cần giải thích

X → Biến giải thích

Sai số mô hình

29

**KIỂM ĐỊNH CÁC THẶNG DƯ CÓ BI TÍNH TỰ TƯƠNG QUAN → Kiểm định Durbin-Watson, DW**

Thặng dư →

$$e_i = y_i - \hat{y}_i$$

Giá trị quan sát

Giá trị từ mô hình

9/4/2018

PGS. Dr. Nguyễn Thông

30

## CHÚ Ý

$$Y_t = a_0 + a_1 X_{1,t} + \varepsilon_t$$

→ Mô hình hồi quy tuyến tính đơn sẽ TỐT khi  $\varepsilon_t$  là một “NHIỀU TRẮNG” (không còn chứa thông tin) → chuỗi KHÔNG CÓ tính tự tương quan. Ok.

9/4/2018

31

PGS. Dr. Nguyễn Thống

Gọi  $\varepsilon_i$  là chuỗi thặng dư của mô hình hồi quy:

Tham số DW:

$$DW = \frac{\sum_{i=2}^n (\varepsilon_i - \varepsilon_{i-1})^2}{\sum_{i=1}^n \varepsilon_i^2}$$

DW = 2 → Chuỗi KHÔNG có tính tự tương quan

9/4/2018

32

PGS. Dr. Nguyễn Thống

## Kiểm định Durbin-Watson

Trong thực hành có thể dùng kiểm định Durbin-Watson để đánh giá (Tham số DW):

→ DW → 2 → Ok

→ DW KHÁC giá trị 2 càng nhiều → càng có nguy cơ tính TỰ TƯƠNG QUAN CỦA  $\varepsilon_i$ .

9/4/2018

33

PGS. Dr. Nguyễn Thống

## Xem SPSS

TKUD \ Durbin Watson Exampe 1.sav

9/4/2018

34

PGS. Dr. Nguyễn Thống

**Ví dụ:** Một thí nghiệm xác định góc nội ma sát và lực dính của tầng địa chất nghiên cứu. Kết quả của 12 thí nghiệm cho các ứng suất tiếp  $\tau$  và ứng suất pháp  $\sigma$  được trình bày trong bảng sau. Xác định  $\phi$  và  $c$  trong quan hệ.

$$\tau = \sigma \tan \phi + c$$

Hệ số m/hình hqтт

Mẫu	1	2	3	4	5	6
$\tau(\text{kg/cm}^2)$	3.2	3.7	3.7	4.2	5.2	5.3
$\sigma(\text{kg/cm}^2)$	3.6	4.2	4.3	4.9	6.2	6.5
Mẫu	7	8	9	10	11	12
$\tau(\text{kg/cm}^2)$	5.5	5.8	6.1	6.6	8.3	9.0
$\sigma(\text{kg/cm}^2)$	6.6	7.0	7.5	8.1	10.4	11.4

9/4/2018

35

PGS. Dr. Nguyễn Thống

## SỬ DỤNG EXCEL HOẶC SPSS ĐỂ ỨNG DỤNG TÍNH HỒI QUY TUYẾN TÍNH

9/4/2018

36

PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 8: Hồi quy tuyến tính**

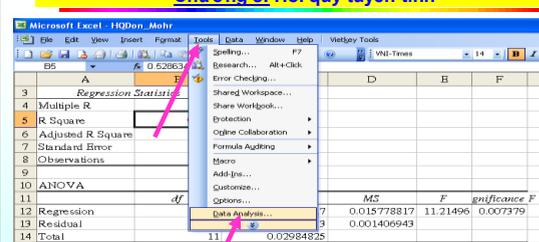


# EXCEL

37

PGS. Dr. Nguyễn Thông

**PHÂN TÍCH DỮ LIỆU**  
**Chương 8: Hồi quy tuyến tính**



	df	MS	F	Significance F
12 Regression	7	0.015778817	11.21496	0.007379
13 Residual	3	0.001409943		
14 Total	11	0.02984825		

	Coefficients	Standard Error	t Stat	P-value	Lower 95% Upper
17 Intercept	0.13266667	0.02652304	5.001940395	0.000536	0.07357 0.19
18 Sigma (kg/cm**2)	0.06486667	0.01936969	3.348874805	0.007379	0.021708 0.10

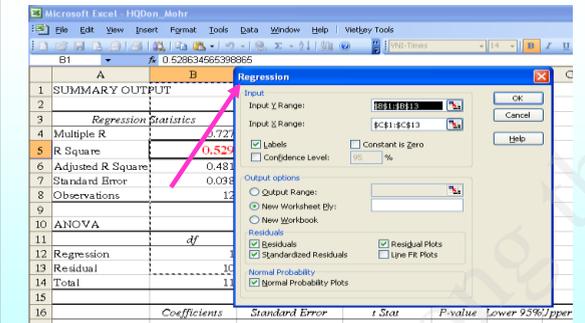
RESIDUAL OUTPUT

Observation	Input Variable	Residuals	Standard Residuals	Percentile
24	0.1651	0.0249	0.696237461	4.166667

9/4/2018

PGS. Dr. Nguyễn Thông

**PHÂN TÍCH DỮ LIỆU**  
**Chương 8: Hồi quy tuyến tính**



	Coefficients	Standard Error	t Stat	P-value	Lower 95% Upper
17 Intercept	0.13266667	0.02652304	5.001940395	0.000536	0.07357 0.19
18 Sigma (kg/cm**2)	0.06486667	0.01936969	3.348874805	0.007379	0.021708 0.10

9/4/2018

PGS. Dr. Nguyễn Thông

**PHÂN TÍCH DỮ LIỆU**  
**Chương 8: Hồi quy tuyến tính**

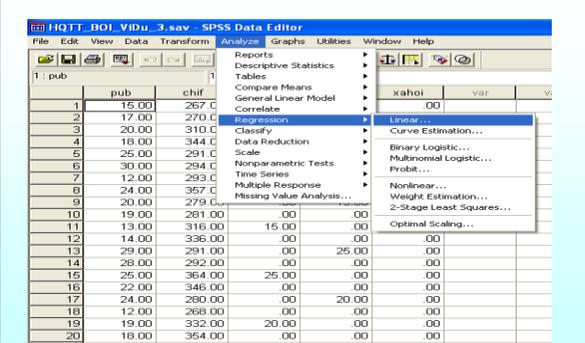
# SPSS

→ Analyze → Regression  
→ Linear

9/4/2018

PGS. Dr. Nguyễn Thông

**PHÂN TÍCH DỮ LIỆU**  
**Chương 8: Hồi quy tuyến tính**

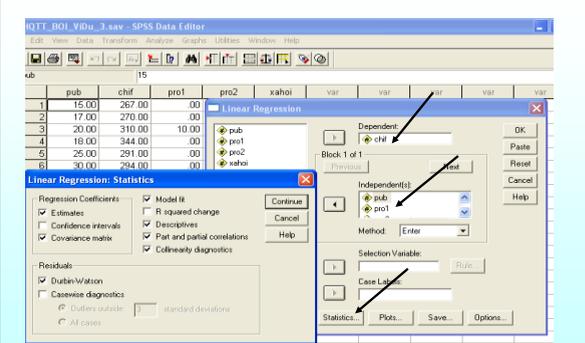


pub	chif	pro1	pro2	xahoi	var	var	var	var
1	15.00	267.00	.00					
2	17.00	270.00	.00					
3	20.00	310.00	10.00					
4	18.00	344.00	.00					
5	25.00	291.00	.00					
6	30.00	294.00	.00					
7	12.00	293.00	.00					
8	24.00	357.00	.00					
9	20.00	279.00	.00					
10	19.00	261.00	.00					
11	13.00	316.00	15.00	.00				
12	14.00	336.00	.00	.00				
13	29.00	291.00	.00	25.00	.00			
14	28.00	292.00	.00	.00				
15	25.00	364.00	25.00	.00	.00			
16	22.00	346.00	.00	.00	.00			
17	24.00	260.00	.00	20.00	.00			
18	12.00	268.00	.00	.00	.00			
19	19.00	332.00	20.00	.00	.00			
20	18.00	354.00	.00	.00	.00			

9/4/2018

PGS. Dr. Nguyễn Thông

**PHÂN TÍCH DỮ LIỆU**  
**Chương 8: Hồi quy tuyến tính**



Linear Regression Statistics

Regression Coefficients:  Estimates,  Confidence intervals,  Covariance matrix

Model fit:  R squared change,  Descriptives,  Part and partial correlations,  Collinearity diagnostics

Residuals:  Durbin-Watson,  Casewise diagnostics,  All cases,  standard deviations

9/4/2018

PGS. Dr. Nguyễn Thông

**PHÂN TÍCH DỮ LIỆU**  
**Chương 8: Hồi quy tuyến tính**

9/4/2018 PGS. Dr. Nguyễn Thống 43

**PHÂN TÍCH DỮ LIỆU**  
**Chương 8: Hồi quy tuyến tính**

9/4/2018 PGS. Dr. Nguyễn Thống 44

**PHÂN TÍCH DỮ LIỆU**  
**Chương 8: Hồi quy tuyến tính**

**HỒI QUY**  
**TUYẾN TÍNH BỘI**

9/4/2018 PGS. Dr. Nguyễn Thống 45

**PHÂN TÍCH DỮ LIỆU**  
**Chương 8: Hồi quy tuyến tính**

Trong trường hợp số biến giải thích > 1 → hồi quy tuyến tính bội:

Biến cần được giải thích

$$Y = a_0 + a_1 X_1 + a_2 X_2 + \dots + a_k X_k + \varepsilon$$

Ví dụ: Hồi quy tuyến tính bội bậc k.

Biến giải thích

9/4/2018 PGS. Dr. Nguyễn Thống 46

**PHÂN TÍCH DỮ LIỆU**  
**Chương 8: Hồi quy tuyến tính**

**GIẢ THIẾT**

**Giả thiết về ngẫu nhiên**

- H1 : các giá trị của  $x_{i,t}$  được quan trắc đúng (không có sai số).
- H2 :  $E(\varepsilon_t) = 0$ ; trung bình số học của các sai số là bằng 0.

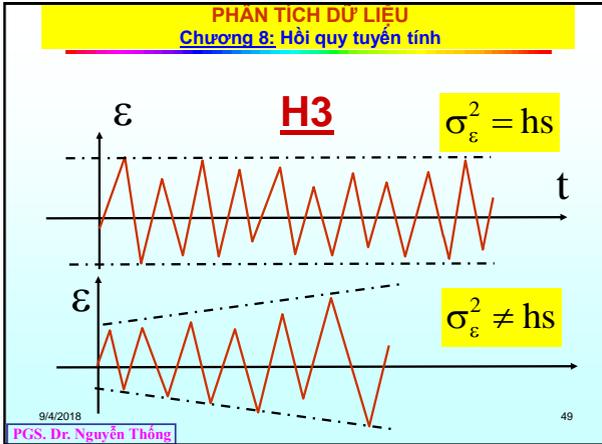
9/4/2018 PGS. Dr. Nguyễn Thống 47

**PHÂN TÍCH DỮ LIỆU**  
**Chương 8: Hồi quy tuyến tính**

- H3 :  $E(\varepsilon_t^2) = \sigma^2 = hs.$

Phương sai của sai số là hằng số với mọi t (biên độ dao động của  $\varepsilon$  quanh giá trị TB của nó là không bị phân kỳ).

9/4/2018 PGS. Dr. Nguyễn Thống 48



- PHÂN TÍCH DỮ LIỆU**  
Chương 8: Hồi quy tuyến tính
- H4 : các sai số thì độc lập với nhau.  
 $E(\varepsilon_t \cdot \varepsilon_{t+p}) = 0$  if  $p \neq 0$
  - H5 :  $Cov(x_{i,t}, \varepsilon_t) = 0$  : các sai số độc lập với biến giải thích.
- 9/4/2018 PGS. Dr. Nguyễn Thống 50

- PHÂN TÍCH DỮ LIỆU**  
Chương 8: Hồi quy tuyến tính
- Giả thiết về cấu trúc**
- H6 : các biến giải thích là độc lập tuyến tính với nhau, điều này cho phép ma trận  $[X'X]$  nghịch đảo được.  
(Tính đa cộng tuyến trong hồi quy bội - Cần loại bỏ)
- 9/4/2018 PGS. Dr. Nguyễn Thống 51

- PHÂN TÍCH DỮ LIỆU**  
Chương 8: Hồi quy tuyến tính
- Giả thiết về cấu trúc**
- H7 :  $X'X/n$  tiến về giới hạn hữu hạn  $\forall n$ .
  - H8 :  $n > (k+1)$  : số lượng quan trắc phải lớn hơn số lượng biến giải thích.
- 9/4/2018 PGS. Dr. Nguyễn Thống 52

**PHÂN TÍCH DỮ LIỆU**  
Chương 8: Hồi quy tuyến tính

**XÁC ĐỊNH HỆ SỐ MÔ HÌNH  $a_1$**

Dùng phương pháp BPTT (xem tài liệu Kinh tế lượng- Tg. Dr. Nguyễn Thống).

$$[Y]_{n,1} = [X]_{n,k+1} X [a]_{k+1,1} + [\varepsilon_t]_{n,1}$$

Sai số mô hình

9/4/2018 PGS. Dr. Nguyễn Thống 53

**PHÂN TÍCH DỮ LIỆU**  
Chương 8: Hồi quy tuyến tính

Với: Hệ số của mô hình hồi quy cần xác định

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}_{(n,1)} \quad X = \begin{pmatrix} 1 & X_{11} & \dots & X_{k1} \\ 1 & X_{12} & \dots & X_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1n} & \dots & X_{kn} \end{pmatrix}_{(n,k+1)} \quad a = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_k \end{pmatrix}_{(k+1,1)}$$

$n \rightarrow$  số quan trắc,  $k \rightarrow$  số biến độc lập

9/4/2018 PGS. Dr. Nguyễn Thống 54

## THIẾT LẬP PHƯƠNG TRÌNH ĐỂ XÁC ĐỊNH $a_i$

$$[Y]_{n,1} = [X]_{n,k+1} \times [a]_{k+1,1} + [\varepsilon_t]_{n,1}$$

→ Min

Dùng p/p bình phương tối thiểu:

$$\begin{aligned} \text{Min} \sum_i \varepsilon_t^2 &= \text{Min}(\varepsilon_t' \varepsilon_t) \\ &= \text{Min}[Y - Xa]' [Y - Xa] \\ &= \text{Min}(S) \end{aligned}$$

$X'$  → Ma trận chuyển vị của  $X$

## ÔN

Ví dụ: Tính module của vector  $U$

$$\vec{U} \begin{bmatrix} 2 \\ 3 \end{bmatrix} \Rightarrow \vec{U}' \equiv [2,3]$$

$$\vec{U}'\vec{U} = |U|^2 = 2*2 + 3*3 = 13$$

## MA TRẬN SUY BIẾN

Ví dụ: Xét ma trận vuông sau:

$$\vec{U} = \begin{bmatrix} 1 & 5 \\ 3 & 15 \end{bmatrix} \Rightarrow \det \vec{U} = 0$$

**Giải thích:** Xét theo hàng (hoặc cột), các vector là “phụ thuộc tuyến tính” → ma trận  $U$  suy biến →  $\det[U] = 0$ .

Vì  $S = f(a) \Rightarrow S$  cực trị (cực tiểu) → :

$$\frac{\partial S}{\partial a} = -\frac{\partial}{\partial a} [Y'Y - Y'Xa - a'X'Y + a'X'Xa] = 0$$

**Chú ý:**

$$\frac{\partial}{\partial a} [Y'Xa] = \frac{\partial}{\partial a} [aY'X]' = [Y'X]' = X'Y$$

Do đó:  $\frac{\partial S}{\partial a} = [-2X'Y + 2X'Xa] = 0$

$$\Rightarrow a \equiv [a] = [X'X]^{-1} [X'Y]$$

$[X']$  → ma trận chuyển vị  $[X]$

**CHÚ Ý**

Để ma trận :  $[X'X]^{-1}$

nghịch đảo được  $\rightarrow \det[X'X] \neq 0$

$\rightarrow$  Khi  $\det[X'X] \rightarrow 0 \rightarrow$  Hiện tượng **ĐA CỘNG TUYẾN** trong hồi quy tuyến tính bội.

**ÔN**

Xét ma trận:

$$[A] = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} \Rightarrow \det[A] = 1*6 - 2*3 = 0$$

Nhận xét 2 vector cột là **PHỤ THUỘC TUYẾN TÍNH**:

$$\begin{bmatrix} 2 \\ 6 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 3 \end{bmatrix} \quad !!!$$

**CHÚ Ý**

$\det[X'X] \rightarrow 0 \rightarrow$  Khi các vector  $X_k$  (vector biến độc lập) có mang tính **PHỤ THUỘC TUYẾN TÍNH** với nhau.

**PTTT**  $\rightarrow$  Vector bất kỳ  $X_k$  là tổ hợp tuyến tính của các vector còn lại.

Phụ thuộc tuyến tính giữa  $X_k$

$$[X] = \begin{pmatrix} 1 & X_{11} & \dots & X_{k1} \\ 1 & X_{12} & \dots & X_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1n} & \dots & X_{kn} \end{pmatrix}_{(n,k+1)}$$

$n \rightarrow$  số quan trắc,  $k \rightarrow$  số biến độc lập

**Ví dụ:** Số liệu quan sát trong 6 năm về tiêu dùng (Y), thu nhập (X1) và tuổi trung bình (X2) nh sau. Dùng phần mềm SPSS xác định hồi quy tuyến tính.

- Xác định các hệ số của mô hình.
- Xác định độ lệch chuẩn của các hệ số.
- Kiểm định giả thiết các hệ số này so với 0.

$$Y = a_0 + a_1X_1 + a_2X_2 + \varepsilon$$

i	Y	X1	X2
1	140	200	45
2	155	220	34
3	180	250	28
4	195	270	24
5	270	400	36
6	330	500	40

$$Y = a_0 + a_1X_1 + a_2X_2 + \varepsilon$$

Xem SPSS  $\rightarrow$  TKUD \ HoiQuyBoi\_1.sav

PHÂN TÍCH DỮ LIỆU  
Chương 8: Hồi quy tuyến tính

# XÁC ĐỊNH HỆ SỐ MÔ HÌNH

9/4/2018 PGS. Dr. Nguyễn Thống 67

PHÂN TÍCH DỮ LIỆU  
Chương 8: Hồi quy tuyến tính

Mô hình  $\rightarrow Y = 37.22 + 0.629X_1 - 0.537X_2$

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
1 (Constant)	37.22	3.60		10.33	.002
X1	.629	.006	1.009	99.10	.000
X2	-.537	.097	-.056	-5.52	.012

9/4/2018 PGS. Dr. Nguyễn Thống 68

PHÂN TÍCH DỮ LIỆU  
Chương 8: Hồi quy tuyến tính

# ĐỘ LỆCH CHUẨN CỦA CÁC HỆ SỐ MÔ HÌNH

9/4/2018 PGS. Dr. Nguyễn Thống 69

PHÂN TÍCH DỮ LIỆU  
Chương 8: Hồi quy tuyến tính

# ĐỘ LỆCH CHUẨN

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
1 (Constant)	37.22	3.60		10.33	.002
X1	.629	.006	1.009	99.10	.000
X2	-.537	.097	-.056	-5.52	.012

9/4/2018 PGS. Dr. Nguyễn Thống 70

PHÂN TÍCH DỮ LIỆU  
Chương 8: Hồi quy tuyến tính

# KIỂM ĐỊNH CÁC HỆ SỐ CỦA $a_i$ MÔ HÌNH CÓ NGHĨA HAY KHÔNG (KHÁC 0 HAY BẰNG 0 THEO QUAN ĐIỂM THỐNG KÊ)

9/4/2018 PGS. Dr. Nguyễn Thống 71

PHÂN TÍCH DỮ LIỆU  
Chương 8: Hồi quy tuyến tính

$$t = B / \text{Std. Error} = \bar{a}_i / \sigma_{a_i}$$

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
1 (Constant)	37.22	3.60		10.33	.002
X1	.629	.006	1.009	99.10	.000
X2	-.537	.097	-.056	-5.52	.012

9/4/2018 PGS. Dr. Nguyễn Thống 72

PHÂN TÍCH DỮ LIỆU  
Chương 8: Hồi quy tuyến tính

Model	Unstandardized Coefficients		Standardized Coefficients		t	Sig.
	B	Std. Error	Beta			
1 (Constant)	37.22	3.60			10.33	.002
X1	.629	.006	1.009		99.10	.000
X2	-.537	.097	-.056		-5.52	.012

Nếu Sig < 5% → Hệ số a<sub>i</sub> KHÁC 0

9/4/2018

PGS. Dr. Nguyễn Thống

73

PHÂN TÍCH DỮ LIỆU  
Chương 8: Hồi quy tuyến tính

Tong tự ví dụ 1 với các số liệu nh sau:

N <sup>o</sup>	y	x1	x2	x3	x4	x5
1	147.4	2	5	20	8	51
2	187.3	4	4	22	18	55
3	139.9	5	9	24	11	58
4	147.1	6	12	22	12	60
5	105.3	8	14	30	10	62
6	131.1	9	12	29	12	55
7	124.7	10	10	27	8	47
8	170.7	7	9	21	13	65
9	172.6	10	8	20	10	45
10	144.8	9	12	22	9	58
11	189.3	9	15	24	19	55
12	165.5	11	13	28	18	50 <sub>4</sub>

9/4/2018

PGS. Dr. Nguyễn Thống

PHÂN TÍCH DỮ LIỆU  
Chương 8: Hồi quy tuyến tính

Hướng dẫn: Dùng SPSS với file  
→ TKUD\HoiQuyBoi\_2.sav

9/4/2018

PGS. Dr. Nguyễn Thống

75

PHÂN TÍCH DỮ LIỆU  
Chương 8: Hồi quy tuyến tính

Hướng dẫn: Dùng SPSS với file → TKUD\  
HoiQuyBoi\_2.sav

$$Y = 222.68 + 2.35X_1 - 4.94X_3 + 5.04X_4$$

Model		Unstandardized Coefficients		Standardized Coefficients		t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta				Tolerance	VIF
1	(Constant)	222.680	18.758			11.871	.000		
	X1	2.352	.824	.253		2.853	.029	.324	3.090
	X2	-1.072	.711	-.143		-1.507	.182	.282	3.552
	X3	-4.947	.461	-.687		-10.726	.000	.618	1.619
	X4	5.045	.333	.777		15.163	.000	.966	1.038
	X5	-.368	.288	-.086		-1.276	.249	.554	1.806

PGS. Dr. Nguyễn Thống

PHÂN TÍCH DỮ LIỆU  
Chương 8: Hồi quy tuyến tính

Bài tập: Người ta muốn kiểm tra sự quan hệ tuyến tính nếu có giữa số ngày nghỉ việc không phép của công nhân trong năm (Y) và các yếu tố thâm niên (x1) và tuổi tác (X2). Xem số liệu sau.

$$Y = a_0 + a_1X_1 + a_2X_2 + \varepsilon$$

9/4/2018

PGS. Dr. Nguyễn Thống

77

PHÂN TÍCH DỮ LIỆU  
Chương 8: Hồi quy tuyến tính

TT	Số ngày vắng (Y)	Thâm niên (X1)	Tuổi (X2)
1	5	5	30
2	4	15	45
3	2	10	42
4	6	6	30
5	8	8	32
6	6	7	35
7	5	10	40
8	3	2	28
9	7	5	50
10	2	19	54

9/4/2018

PGS. Dr. Nguyễn Thống

78

PHÂN TÍCH DỮ LIỆU  
Chương 8: Hồi quy tuyến tính

**Bài tập:** Người ta muốn kiểm tra sự quan hệ tuyến tính nếu có giữa lượng rác thải hàng ngày và nhiệt độ trong ngày (so với nhiệt độ chuẩn 65°F) và độ ẩm (so với độ ẩm chuẩn 55%). Xem số liệu sau.

$$Y = a_0 + a_1X_1 + a_2X_2 + \varepsilon$$

9/4/2018

PGS. Dr. Nguyễn Thống

79

PHÂN TÍCH DỮ LIỆU  
Chương 8: Hồi quy tuyến tính

TT	Rác thải ngày (Y)	Nhiệt (X1)	Độ ẩm (X2)
1	14	-5	1
2	12	-8	1
3	16	1	4
4	20	6	6
5	24	8	7
6	29	12	14
7	24	3	17
8	14	-2	2
9	16	-7	1
10	10	-3	-1
11	14	-8	-2
12	8	-2	3

9/4/2018

PGS. Dr. Nguyễn Thống

80

PHÂN TÍCH DỮ LIỆU  
Chương 8: Hồi quy tuyến tính

**HIỆN TƯỢNG  
ĐA CÔNG TUYẾN  
TRONG HỒI QUY  
TUYẾN TÍNH BỘI**



81

PHÂN TÍCH DỮ LIỆU  
Chương 8: Hồi quy tuyến tính

Để  $X_1, X_2, \dots, X_i$  giữ vai trò là biến giải thích để nghiên cứu biến  $Y \rightarrow$  các biến  $X_1, X_2, X_i$  phải **ĐỘC LẬP TUYẾN TÍNH**.

$\rightarrow$  Đây là điều kiện « CẦN » phải có khi nghiên cứu hồi quy tuyến tính bội.

$\rightarrow$  nếu **KHÔNG**  $\rightarrow$  các hệ số của mô hình ( $a_i$ ) không xác định được !!!

9/4/2018

PGS. Dr. Nguyễn Thống

82

PHÂN TÍCH DỮ LIỆU  
Chương 8: Hồi quy tuyến tính

**Ghi chú**

Để  $X_1$  &  $X_2$  là **ĐỘC LẬP TUYẾN TÍNH**:

$\rightarrow [X_1].[X_2] \rightarrow 0$

Để  $X_1, X_2$  &  $X_3$  là **ĐỘC LẬP TUYẾN TÍNH**:

$\rightarrow X_1$  **KHÔNG** LÀ TỔ HỢP TUYẾN TÍNH CỦA X CÒN LẠI.

9/4/2018

PGS. Dr. Nguyễn Thống

83

PHÂN TÍCH DỮ LIỆU  
Chương 8: Hồi quy tuyến tính

**ĐÁNH GIÁ HIỆN TƯỢNG  
ĐA CÔNG TUYẾN  
TRONG HỒI QUY BỘI  
VỚI THAM SỐ VIF**



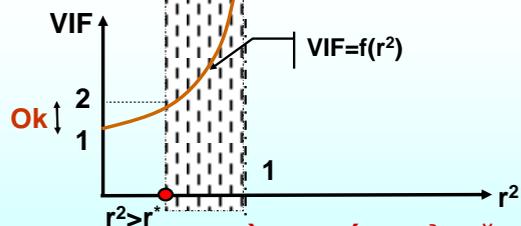
84

### VIF → Variance Inflation Factor

Xét 1 biến độc lập  $X_i$ :

$$VIF_i = \frac{1}{1 - r_i^2}$$

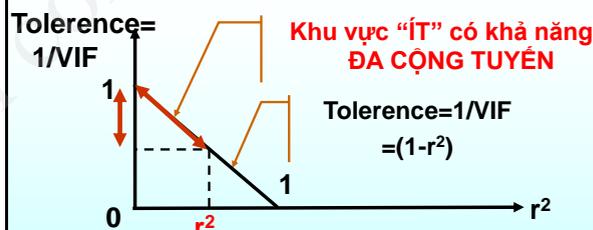
$r_i^2$  → hệ số xác định trong hồi quy tuyến tính của biến giải thích  $X_i$  theo tất cả các biến giải thích  $X_j$  còn lại (với  $i \neq j$ ).



**VIF càng lớn → CÀNG CÓ KHẢ NĂNG có hiện tượng ĐA CỘNG TUYẾN (vì  $r^2 \rightarrow 1$ )**

### CHÚ Ý

Phổ biến  $r^2 < 0,5 \rightarrow VIF < 2$ .  
→ Ok → KHÔNG CÓ HIỆN TƯỢNG về Đa cộng tuyến của các biến độc lập.



**ĐỒ THỊ ỨNG DỤNG TRONG PHÂN TÍCH ĐA CỘNG TUYẾN**

### CHÚ Ý

**Tolerance = 1/VIF**

Nếu Tolerance  $< (1-R^2)$  → CÓ THỂ có hiện tượng ĐA CỘNG TUYẾN.

Yêu cầu: Tolerance  $> (1-R^2)$

$R^2$  hệ số xác định (hiệu chỉnh) của mô hình.

**SPSS NHẬN BIẾT ĐA CỘNG TUYẾN (xem SPSS với file HồiQuyBoi.sav)**

Ví dụ: Một nghiên cứu y học về quan hệ giữa Cholesterol (biến nghiên cứu Y) và các yếu tố Tuổi (Age), Trọng lượng (Weight), Chiều cao (height), Áp huyết, số liệu đã chuẩn hóa, của 14 đối tượng như sau:

PHÂN TÍCH DỮ LIỆU  
Chương 8: Hồi quy tuyến tính

$$\text{Choles.} = a_0 + a_1 \cdot \text{Age} + a_2 \cdot \text{Weight} + a_3 \cdot \text{Height} + a_4 \cdot \text{Tension}$$

9/4/2018

91

PGS. Dr. Nguyễn Thống

PHÂN TÍCH DỮ LIỆU  
Chương 8: Hồi quy tuyến tính

tt	Y	Tuổi X1	Trọng lượng X2	Cao X3	Áp huyết X4
1	-3.3	-1.6	-1	0.7	0
2	-0.2	-0.5	-1.2	-0.5	-0.2
3	5	1.3	-1	-0.9	-0.5
4	-1	0.8	0.8	1.5	1
5	0	0.4	0	0.3	0.2
6	-6	-0.3	0.2	1.9	1
7	-4	-1.4	-0.7	0.6	0.5

9/4/2018

92

PGS. Dr. Nguyễn Thống

PHÂN TÍCH DỮ LIỆU  
Chương 8: Hồi quy tuyến tính

tt	Tuổi	Trọng lượng	Cao	Áp huyết	
8	-1.5	-0.2	-0.2	0.3	0.2
9	3.0	1.7	0.3	0.2	0.5
10	0.6	0.5	0.0	0.8	0.2
11	0.4	0.5	0.0	0.3	0.0
12	7.0	0.4	-0.3	-1.9	-1.0
13	1.5	-1.1	2.3	-1.0	-2.0
14	2.0	-0.5	1.3	-1.0	-0.5

9/4/2018

93

PGS. Dr. Nguyễn Thống

PHÂN TÍCH DỮ LIỆU  
Chương 8: Hồi quy tuyến tính

→ SPSS

Coefficients <sup>a</sup>								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
		1	(Constant)	.420			.242	
	X1	2.032	.292	.575	6.961	.000	.699	1.431
	X2	.004	.308	.001	.012	.991	.659	1.517
	X3	-2.172	.546	-.657	-3.979	.003	.174	5.735
	X4	-.725	.789	-.166	-.920	.382	.146	6.855

→ Loại bỏ X4

Coefficients <sup>a</sup>								
		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
		1	(Constant)	.49			.23	
	X1	1.88	.24	.53	7.78	.000	1.00	1.002
	X2	.17	.25	.05	.67	.516	.99	1.008
	X3	-2.63	.23	-.80	-11.57	.000	.99	1.009

→ Loại bỏ X2

Coefficients <sup>a</sup>								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
		1	(Constant)	.50			.22	
	X1	1.88	.24	.53	7.96	.000	1.00	1.001
	X3	-2.64	.22	-.80	-11.97	.000	1.00	1.001

a. Dependent Variable: Y

### KẾT QUẢ HỒI QUY

$$\text{Choles.} = a_0 + a_1 \cdot \text{Age} + a_3 \cdot \text{Height}$$

9/4/2018

PGS. Dr. Nguyễn Thống

97

### TÓM TẮT

- Xem xét giá trị  $R^2$  (càng gần giá trị 1 → OK)
- Xem xét các hệ số mô hình có KHÁC 0 ? (sig < 5% → ok, nếu không → loại biến liên quan)
- Xem xét tính tự tương quan (D-W → 2 tốt nhất), nếu KHÔNG → thêm biến giải thích
- Kiểm tra tính ĐA CỘNG TUYẾN VỚI tham số VIF (< 2 → OK). Loại bỏ biến có VIF > 2.

9/4/2018

PGS. Dr. Nguyễn Thống

98

### BIẾN GIẢ (BIẾN CHỈ BÁO, DUMMY)



9/4/2018

PGS. Dr. Nguyễn Thống

99

### BIẾN CHỈ BÁO

Biến chỉ báo được sử dụng trong trường hợp biến giải thích có dạng là các biến cục bộ tác dụng lên biến nghiên cứu.

Ví dụ 3: Kết quả kinh doanh 10 năm cuối của Công ty Z nh sau với:

Dso : doanh thu ; Pub : chi phí quảng cáo ;  
Promo : tỷ lệ giảm giá ; Xhoi : yếu tố xã hội

9/4/2018

PGS. Dr. Nguyễn Thống

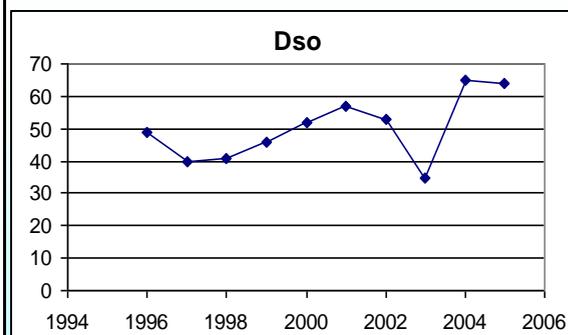
100

Năm	Dso (100tr.)	Pub (tr.)	Promo (%)	Xhoi
1996	49	41	10	0
1997	40	32	15	0
1998	41	38	10	0
1999	46	40	10	0
2000	52	40	5	0
2001	57	42	0	0
2002	53	44	0	0
2003	35	46	0	1
2004	65	50	5	0
2005	64	55	0	0

9/4/2018

PGS. Dr. Nguyễn Thống

101

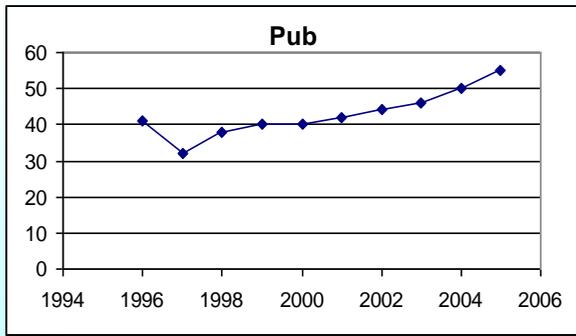


9/4/2018

PGS. Dr. Nguyễn Thống

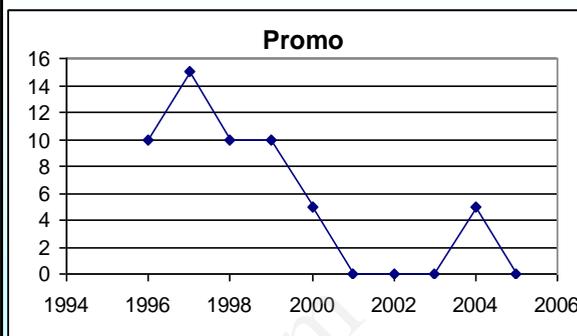
102

PHÂN TÍCH DỮ LIỆU  
Chương 8: Hồi quy tuyến tính



PGS. Dr. Nguyễn Thống

PHÂN TÍCH DỮ LIỆU  
Chương 8: Hồi quy tuyến tính



PGS. Dr. Nguyễn Thống

PHÂN TÍCH DỮ LIỆU  
Chương 8: Hồi quy tuyến tính

## BIẾN ĐỊNH TÍNH



PGS. Dr. Nguyễn Thống

105

PHÂN TÍCH DỮ LIỆU  
Chương 8: Hồi quy tuyến tính

## BIẾN ĐỊNH TÍNH

Có 2 loại biến trong nghiên cứu hồi quy:

- **Biến định lượng:** Thu nhập, Doanh thu, Số lượng sản phẩm tiêu thụ...
- **Biến định tính:** Nghề nghiệp, Trình độ văn hóa, Giới tính, Đồng ý không đồng ý, Màu sắc, ...

PGS. Dr. Nguyễn Thống

106

PHÂN TÍCH DỮ LIỆU  
Chương 8: Hồi quy tuyến tính

### BIẾN ĐỊNH TÍNH

i	Y (t/nhập)	Giới tính		Trình độ			Thâm niên
		Nam	Nu	<=PT	DH	SDH	
1	15000	1	0	1	0	0	8
2	25000	0	1	0	1	0	15
.	.	1	0	.	.	.	.
.	.	.	.	.	.	.	.
N	12000	0	1	0	1	0	5

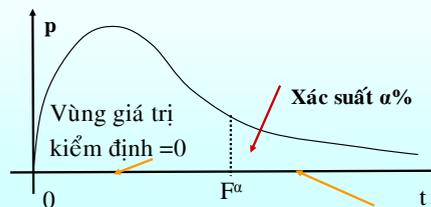
PGS. Dr. Nguyễn Thống

107

PHÂN TÍCH DỮ LIỆU  
Chương 8: Hồi quy tuyến tính

$$\Pr(F_{v_1;v_2} > F^\alpha) = \alpha\%$$

Có giá trị của  $\alpha, v_1, v_2$  xác định giá trị của  $F_\alpha$  bởi:



### PHÂN PHỐI FISHER

PGS. Dr. Nguyễn Thống

Vùng giá trị kiểm định khác 0

108

**PHÂN TÍCH DỮ LIỆU**  
**Chương 8: Hồi quy tuyến tính**

$v \setminus  v $	1	2	3	4	5	6	7	8	9
1	161	200	216	225	230	234	237	239	241
2	18.5	19	19.2	19.2	19.3	19.3	19.4	19.4	19.4
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.1
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
8	5.32	4.46	4.07	3.84	3.69	3.58	3.5	3.44	3.39
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
10	4.96	4.1	3.71	3.48	3.33	3.22	3.14	3.07	3.02
11	4.84	3.98	3.59	3.36	3.2	3.09	3.01	2.95	2.9
12	4.75	3.89	3.49	3.26	3.11	3	2.91	2.85	2.8
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
14	4.6	3.74	3.34	3.11	2.96	2.85	2.76	2.7	2.65

9/4/2018 109

PGS. Dr. Nguyễn Thống **Fisher với  $\alpha=5\%$**

**PHÂN TÍCH DỮ LIỆU**  
**Chương 8: Hồi quy tuyến tính**

15	4.54	3.68	3.29	3.06	2.9	2.79	2.71	2.64	2.59
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
17	4.45	3.59	3.2	2.96	2.81	2.7	2.61	2.55	2.48
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
19	4.38	3.52	3.13	2.9	2.74	2.63	2.54	2.48	2.42
20	4.35	3.49	3.1	2.87	2.71	2.6	2.51	2.45	2.39
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
22	4.3	3.44	3.05	2.82	2.66	2.55	2.46	2.4	2.34
23	4.28	3.42	3.03	2.8	2.64	2.53	2.44	2.37	2.32
24	4.26	3.4	3.01	2.78	2.62	2.51	2.42	2.36	2.3
25	4.24	3.39	2.99	2.76	2.6	2.49	2.4	2.34	2.28
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
60	4	3.15	2.76	2.53	2.37	2.25	2.17	2.1	2.04
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96
	3.84	3	2.6	2.37	2.21	2.1	2.01	1.94	1.88

9/4/2018 110

PGS. Dr. Nguyễn Thống **Fisher với  $\alpha=5\%$**

**PHÂN TÍCH DỮ LIỆU**  
**Chương 8: Hồi quy tuyến tính**

10	12	15	20	24	30	40	60	120	$\infty$
242	244	246	248	249	250	251	252	253	254
19.4	19.4	19.4	19.5	19.5	19.5	19.5	19.5	19.5	19.5
8.79	8.74	8.7	8.66	8.64	8.62	8.59	8.57	8.55	8.53
5.96	5.91	5.86	5.8	5.77	5.75	5.72	5.69	5.66	5.63
4.74	4.68	4.62	4.56	4.53	4.5	4.46	4.43	4.4	4.37
4.06	4	3.94	3.87	3.84	3.81	3.77	3.74	3.7	3.67
3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.3	3.27	3.23
3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
3.14	3.07	3.01	2.94	2.9	2.86	2.83	2.79	2.75	2.71
2.98	2.91	2.85	2.77	2.74	2.7	2.66	2.62	2.58	2.54
2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.4
2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.3
2.67	2.6	2.53	2.46	2.42	2.38	2.34	2.3	2.25	2.21
2.6	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
2.54	2.48	2.4	2.33	2.29	2.25	2.2	2.16	2.11	2.07

9/4/2018 111

PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 8: Hồi quy tuyến tính**

2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
2.45	2.38	2.31	2.23	2.19	2.15	2.1	2.06	2.01	1.96
2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
2.39	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
2.35	2.28	2.2	2.12	2.08	2.04	1.99	1.95	1.9	1.84
2.32	2.25	2.18	2.1	2.05	2.01	1.96	1.92	1.87	1.81
2.3	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
2.27	2.2	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
2.08	2	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
1.99	1.92	1.84	1.75	1.7	1.65	1.59	1.53	1.47	1.39
1.91	1.83	1.75	1.66	1.61	1.55	1.5	1.43	1.35	1.25
1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1

9/4/2018 112

PGS. Dr. Nguyễn Thống

**PHÂN TÍCH DỮ LIỆU**  
**Chương 8: Hồi quy tuyến tính**






# HẾT CHƯƠNG

## Xin cảm ơn!

9/4/2018 113

PGS. Dr. Nguyễn Thống