

Regression Analysis

W. M. Song 桑慧敏
Tsing Hua Univ. 清華大學

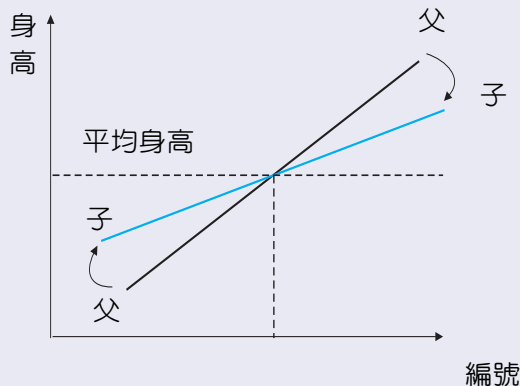
2015.12.16

1 History about Regression

2 Regression Application: 太陽系九大行星

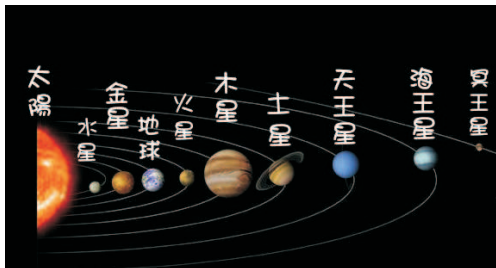
“Regression Towards Mean”

Regression towards the mean: Francis Galton (1877)



太陽系九大行星

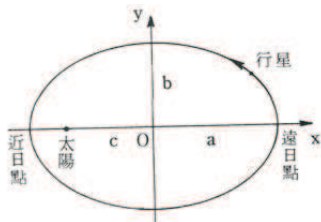
| 行星 | 公轉週期 (day) | 軌道的平均半徑 (10^9 cm) |
|----|------------|----------------------|
| 水 | 88 | 58 |
| 金 | 225 | 108 |
| 地 | 365 | 150 |
| 火 | 687 | 228 |
| 木 | 4339 | 778 |
| 土 | 10761 | 1427 |
| 天王 | 30690 | 2875 |
| 海王 | 60185 | 4505 |
| 冥王 | 90782 | 5914 |



- 這組數有何玄機?

太陽系九大行星: 繪圖

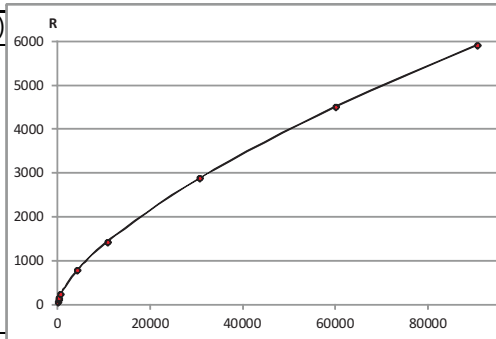
- 地球繞太陽的週期 (T): 1 年, 365 天 (Reference: 月球繞地球週期: 27.32 天)
- 地球繞太陽的平均軌道半徑 (R): 近日點距離和遠日點距離的平均值, 也等於橢圓的長軸半長, 也等於短軸頂點與太陽距離, 又稱為 1 A.U. (Astronomical Unit), 一天文單位, 1.5×10^{11} 公尺, 150×10^9 公分



- 何謂: 一天文單位

太陽系九大行星: 繪圖

| 行星 | 週期 (T: day) | 半徑 (R: 10^9 cm) |
|----|-------------|-------------------|
| 水 | 88 | 58 |
| 金 | 225 | 108 |
| 地 | 365 | 150 |
| 火 | 687 | 228 |
| 木 | 4339 | 778 |
| 土 | 10761 | 1427 |
| 天王 | 30690 | 2875 |
| 海王 | 60185 | 4505 |
| 冥王 | 90782 | 5914 |



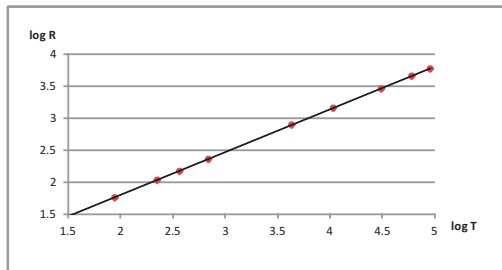
- 數據分析的第一步 — 整理數據 (如: 繪圖)
- 週期 (T) 與軌道 (R) 間有何函數關係?
- 模式一: 迴歸線 $R^2 = 0.98$, very high

模式二: $R = aT^b$. 估計 a, b 值

- $\log R = \log (aT^b)$
- $\log R = \log a + b \log T$
- $y = \alpha + \beta x$, where
 - $y = \log R$
 - $\alpha = \log a$
 - $\beta = \log b$
 - $x = \log T$

| T | R | | $\log T$ | $\log R$ |
|---------|-------|---|----------|----------|
| (88, | 58) | → | (1.944, | 1.763) |
| (225, | 108) | → | (2.352, | 2.033) |
| (365, | 150) | → | (2.562, | 2.176) |
| (687, | 228) | → | (2.837, | 2.358) |
| (4339, | 778) | → | (3.637, | 2.891) |
| (10761, | 1427) | → | (4.032, | 3.154) |
| (30690, | 2875) | → | (4.487, | 3.459) |
| (60185, | 4505) | → | (4.779, | 3.654) |
| (90782, | 5914) | → | (4.958, | 3.772) |

模式二: $R = aT^b$. 估計 a, b 值



$$\log R = 0.466 + 0.667$$

$$\log T$$

$$= 0.466 + (2/3) \log T$$

$$= 0.466 + \log T^{2/3}$$

$$R = 10^{0.466} T^{2/3}$$

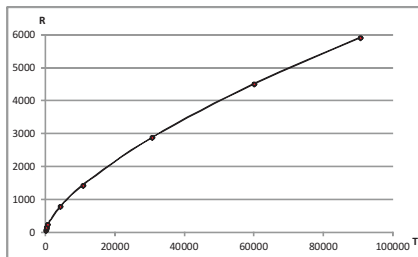
$$R^3 = 25T^2$$

- 上述數據作何解釋?
- $R^3/T^2 = 25$, 常數

Kepler (1571 -1630) 行星運動第三定律:

$$R^3/T^2 = 25, \text{ 常數}$$

| 行星 | 週期 (T: day) | 半徑 (R: 10^9 cm) |
|----|-------------|-------------------|
| 水 | 88 | 58 |
| 金 | 225 | 108 |
| 地 | 365 | 150 |
| 火 | 687 | 228 |
| 木 | 4339 | 778 |
| 土 | 10761 | 1427 |
| 天王 | 30690 | 2875 |
| 海王 | 60185 | 4505 |
| 冥王 | 90782 | 5914 |



Kepler (1571 -1630) 行星運動第三定律:

$$R^3/T^2 = 25, \text{ 常數}$$

- Q: 我們利用了那些邏輯?
- A: 資料收集
- A: Log function
- A: Regression Model
- Q: 只有電腦軟體, regression function, 是否辦的到?

Practice Regression via Minitab

行星運動第三定律