



## 卡方檢定 Kiểm định Chi Square

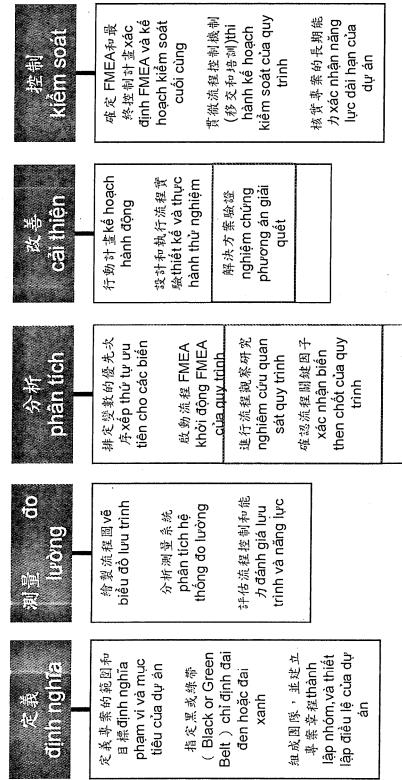
### 主題 chủ đề

- 介紹卡方 (Chi-Square) 檢定的基本概念 - 檢定獨立性 giới thiệu về khái niệm cơ bản của kiểm định Chi-Square- kiểm định tính độc lập.
- 使用 Minitab 執行檢定 sử dụng Minitab thực hiện kiểm định.
- 練習這個技巧 thực hiện kỹ năng này.

Six Sigma

2

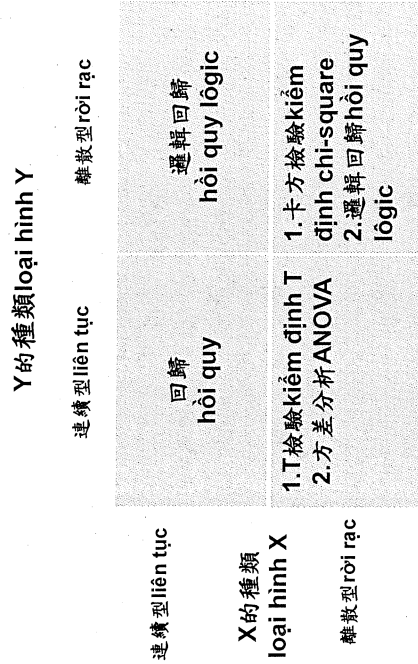
### 流程改善路徑圖 sơ đồ cải thiện lưu trình



Six Sigma

3

### 分析路徑圖：單個 Y 和單個 X sơ đồ phân tích: một X và một Y



Six Sigma

4

範例 thí dụ

品質部門想要知道產品品質 (好或壞) 和生產線 (A 或 B) 是否有關連  
bộ phận chất lượng muốn biết sản phẩm chất lượng(tốt hoặc hỏng)  
và dây chuyền (A hoặc B) có phải liên quan

使用什麼工具?  
sử dụng công cụ nào?

不連續的 Y  
不連續的 X  
Y không liên tục  
X không liên tục  
卡方檢定  
Kiểm định Kalsquare

連續型 liên tục  
X 的種類  
loại hình X  
離散型 rời rạc

Y 的種類 loại hình Y  
連續型 liên tục  
離散型 rời rạc

回歸  
hồi quy

1.T檢驗 kiểm định T  
2.方差分析 ANOVA

邏輯回歸  
hồi quy logic

1.卡方檢驗 kiểm định chi-square  
2.邏輯回歸 hồi quy logic

Six Sigma

5

卡方檢定: 介紹簡介 thiệu về kiểm định Chi Square

■ 此方法是先將資料分類排列在可能性表格中 (contingency table)  
phương pháp này là sắp xếp phân loại dữ liệu phân loại trong  
bảng biểu khả năng.

■ 目標是要決定來自同一離散型數據群體的兩種分類在統計上是否  
為互相獨立的 mục tiêu là phải xác định hai phạm trù dữ liệu loại  
rời rạc rút từ cùng một quần thể có phải độc lập với nhau.

■ 亦能被解釋為比較兩種或兩種以上的群體 cũng có thể coi là so  
sánh hai quần thể hoặc hai trở lên.

■ 應用的檢定分配即為卡方分配 phân bố được sử dụng để kiểm  
định là phân bố Chi Square.

Six Sigma

6

卡方分配 kiểm định Chi Square

■ 卡方值是根據自由度與期望及觀測頻率計算出 giá trị chi square là  
căn cứ bậc tự do và tần số kì vọng cùng với tần số quan trắc  
tính ta

$$\chi^2 = \frac{\sum (f_o - f_e)^2}{f_e}$$

where:

f<sub>o</sub> = Observed Frequency  
tần số quan trắc

f<sub>e</sub> = Expected Frequency  
tần số kì vọng

Six Sigma

7

假設 gia thuyết

卡方獨立性檢定 Chi Square kiểm định tính độc lập

統計學家假定大部分的變數都是互相獨立的，因此：nhà thống  
kê giả sử hầu hết các biến là độc lập với nhau, cho nên

H<sub>0</sub>: 數據是互相獨立的 (不相關) số liệu là độc lập với  
nhau(không liên quan)

H<sub>a</sub>: 數據是相依的 (相關) số liệu là dựa vào nhau(liên  
quan)

如果 P 值 < .05，就拒絕 H<sub>0</sub>  
Nếu giá trị P < .05, thì bác bỏ H<sub>0</sub>

Six Sigma

8

9 - 2

### 範例1thí dụ 1

■ 假如我們擲一個有正反兩面的錢幣100次，觀察到63次正面，37次反面。這個正反面的比例是碰巧發生的呢？還是這個錢幣被做假了？Giả sử chúng ta ném đồng xu có mặt trước và sau hai mặt 100 lần, quan trắc thấy 63 lần mặt trước, 37 lần mặt sau. Tỷ lệ mặt trước và mặt sau có phải tình cờ xảy ra? Hoặc là tiền xu đã bị làm giả?

觀察頻率 quan trắc số kỳ vọng

	$(f_o)$	$(f_e)$
正面 mặt chính	63	50
反面 mặt trái	37	50

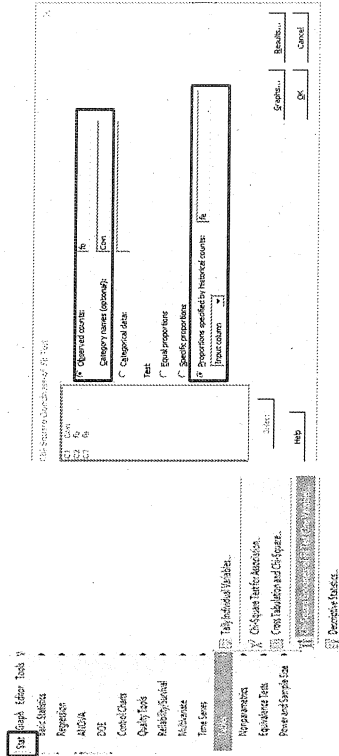
$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e} = 6.76$$

正反面的比例是碰巧發生的呢？還是這個錢幣被做假了？  
Tỷ lệ mặt trước và mặt sau có phải tình cờ xảy ra? Hoặc là tiền xu đã bị làm giả?

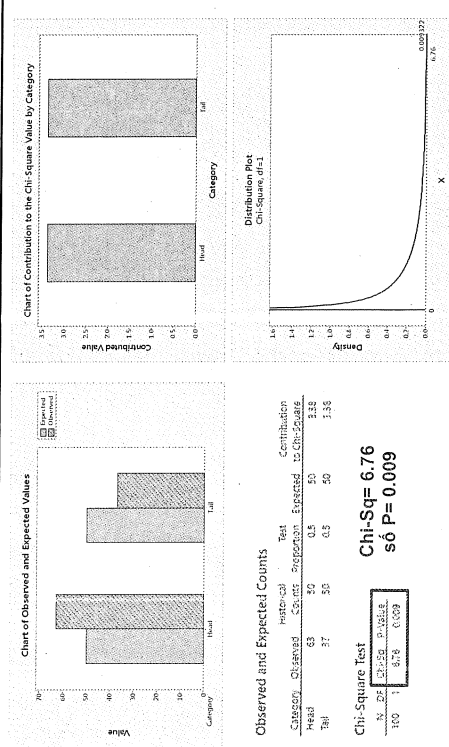
### 範例1：卡方擬合優度？(單變量)

#### Chi Square Goodness-of-Fit Test(One Variable)

◆ Stat> Tables> Chi Square Goodness-of-Fit Test(One Variable)



### 範例1：Minitab分析結果thí dụ 1:kết quả phân tích



### 範例2 thí dụ 2

■ 藥學的學生想要知道兩種頭痛藥是否會有不同療效。於是讓頭痛病患服用不同的頭痛藥並收集結果數據如下：sinh viên khoa dược phẩm muốn biết hai loại thuốc đau đầu có phải có hiệu quả khác nhau. Cho bệnh nhân đau đầu uống thuốc đau đầu khác nhau và thu số liệu kết quả như sau:

我們可以把表格轉90度角，結果不會改變 chúng ta có thể quay bảng biểu 90 độ góc, kết quả không thay đổi

兩個或多個互相獨立的欄位 hai cột hoặc nhiều cột độc lập với nhau

兩個或多個互相獨立的列 hai dòng hoặc nhiều dòng độc lập với nhau

	Pain	No Pain
Remedy 1	9	27
Remedy 2	5	48

此為可能性表格 đây là bảng biểu tính khả năng

數學上計算方法:觀測頻率

phương pháp tính toán tần số quan trắc

9, 27, 5, 和 48 為觀測到的頻率 (實際值)

9, 27, 5 và 48 là tần số quan trắc được (giá trị thực tế)

	Pain	No Pain
Remedy 1	9	27
Remedy 2	5	48

觀測頻率 ( $f_o$ ) tần số quan trắc ( $f_o$ )

Six Sigma

13

數學上計算方法:觀測頻率

phương pháp tính toán tần số quan trắc

使用各列與各欄的總值來估計期望頻率 (如果只有隨機變異存在的話, 我們應該觀測到的頻率) dùng tổng giá trị các hàng và cột để ước tính tần số kì vọng (nếu chỉ có biến thể ngẫu nhiên, tần số chúng ta nên quan trắc được)

	Pain	No Pain	Row Totals
Remedy 1	9	27	36
Remedy 2	5	48	53
Column Totals	14	75	89

Six Sigma

14

計算期望頻率 tính tần số kỳ vọng

將 X 設為期望值 đạt X là giá trị kỳ vọng

用計算時, 請注意四捨五入所造成的誤差 khi tính bằng tay, xin chú ý sự chênh lệch do bốn bỏ năm thêm tạ gây ra.

	Pain	No Pain	Row Total
Remedy 1	X		36
Remedy 2			53
Column Totals	14	75	89

(X 除以該欄總數) 應等於 (X 所在橫列總數除以全部總數) (X chia tổng số cột này) nên bằng (tổng số hàng ngang X chia tổng số toàn bộ)

$$\frac{X}{14} = \frac{36}{89} \rightarrow \text{所以 nên } X = 5.663$$

Six Sigma

15

填入期望頻率 điền vào tần số kì vọng

	Pain	No Pain	Row Total
Remedy 1	5.663	30.337	36
Remedy 2	8.337	44.663	53
Column Totals	14	75	89

填入其他的值 điền vào giá trị khác

Six Sigma

16

計算 (觀測頻率 - 期望頻率)

tính toán(tần số quan trắc- tần số kỳ vọng)

	Pain	No Pain	Row Total
Remedy 1	3.337	-3.337	36
Remedy 2	-3.337	3.337	53
Column Totals	14	75	89

將 (觀測頻率 - 期望頻率) 填入適當空格 (quan trắc- tần số kỳ vọng) điền vào chỗ trống thích đáng

將計算結果填入適當空格中 kết quả tính toán điền vào chỗ trống thích đáng

(observed-expected)<sup>2</sup>

expected

Six Sigma17

接著將四個計算值加總

tiếp theo cho bốn số gia công

	Pain	No Pain	Row Total
Remedy 1	1.966	0.367	36
Remedy 2	1.336	0.249	53
Column Totals	14	75	89

接下來可參考卡方參數表。對於 2 列 x 2 欄的情況，如果我們的總值 (卡方值) 大於 3.84，表示結果是偶然發生的機率小於 .05，因此拒絕總資料為互相獨立 (不相關) 的無效假設。我們接受兩種藥品所產生的療效是不一樣的。Tiếp theo có thể tham khảo bảng tham số Chi Square. Với tính hình 2 hàng x 2 cột, nếu tổng số của chúng ta (giá trị Chi Square) lớn hơn 3.84, chúng ta kết quả là xảy ra bất ngờ xác suất nhỏ hơn .05, vậy bác bỏ giả thuyết vô hiệu tức là dữ liệu là độc lập với nhau(không liên quan). Chúng ta chấp nhận hiệu quả của hai loại thuốc là khác nhau.

1.966

0.367

1.336

0.249

3.918

Six Sigma19

計算 (觀測頻率 - 期望頻率) / 期望頻率

tính toán(tần số quan trắc- tần số kỳ vọng) / tần số kỳ vọng

	Pain	No Pain	Row Totals
Remedy 1	1.966	0.367	36
Remedy 2	1.336	0.249	53
Column Totals	14	75	89

將計算結果填入適當空格中 kết quả tính toán điền vào chỗ trống thích đáng

(observed-expected)<sup>2</sup>

expected

Six Sigma18

解釋結果 giải thích kết quả

Distribution Plot

Chi-Square, df=1

0

0.2

0.4

0.6

0.8

1.0

1.2

1.4

1.6

0

3.918

0.04777

X

Khi Số chi Square=3.918 ,

P-Value = 0.04777

Six Sigma20



不太好的消息 tin không tốt

■ 卡方檢定不具 “洞察力” 檢定 Chi Square không có “năng lực xét rõ”

■ 僅告訴我們橫列變量與欄位變量是否互相獨立 chỉ bảo cho chúng ta các biến số hàng ngang và biến số cột có phải độc lập với nhau

要注意樣本選擇 phải chú ý lựa chọn mẫu

Six Sigma

25

分析路徑圖 lô trình phân tích

卡方-獨立性檢定  
Chi Square-kiểm định tính độc lập

↓

收集資料 thu tập dữ liệu

↓

執行 Minitab Tables>Chi Square 指令 thực hiện Minitab Tables>Chi Square chỉ lệnh

↓

評估 P 值 ước tính P

↓

檢查可能性表格 kiểm tra bảng biểu tính khả năng

↓

做決策 làm quyết sách

Six Sigma

26

摘要 tóm lược

■ 卡方檢定用於計數之聲散型資料 kiểm định Chi Square phù hợp với dữ liệu loại hình tính số

■ 資料排列在一矩陣中，列間資料互相獨立，欄位元的資料亦互相獨立 dữ liệu trình bày trong một Matrix, dữ liệu của các hàng là độc lập với nhau, dữ liệu của các cột cũng độc lập với nhau

■ 如果觀測和期望頻率大致上是相等的，我們得到的結論便是只有隨機的變異存在 nếu tần số quan trắc và kỳ vọng độ chùng bằng nhau, kết luận chúng ta suy ra chỉ là biến thể ngẫu nhiên.

■ 如果觀測和期望頻率是大不相同的，我們得到的結論便是有除了隨機變異的效應存在 nếu tần số quan trắc và kỳ vọng là khác nhau rất lớn, kết luận chúng ta suy ra là có biến thể ngoài ngẫu nhiên.

■ 如果您只能收集到離散型的資料，卡方檢定是一個很不錯的檢定。如果 can thu tập dữ liệu rời rạc, kiểm định Chi Square là một kiểm định tốt nếu có thể lấy được dữ liệu liên tục, bạn có thể chọn phương pháp kiểm định mạnh mẽ hơn nữa

Six Sigma

27

練習 bài tập

開啟檔案 dữ liệu : Supplier Evaluation.MTW

■ 有三個金屬瓶裝壓縮氣體的可能供應商。由三個經銷商取得抽樣金屬瓶，並執行壓力測試。有些爆破了 (burst)，有些則沒有。試問是否和不同廠商有關? Có ba nhà cung ứng khả năng sản xuất bình kim loại đựng khí thể. Do ba nhà tiêu thụ lấy được mẫu bình kim loại, và thử nghiệm ép lực. Có một số bộc phá, có một số thì không. Hãy trả lời có phải liên quan với nhà cung ứng khác nhau.

	burst	no burst
supplier 1	7	61
supplier 2	6	82
supplier 3	5	92

■ 你能不能想出更有效率的測試方法以評估供貨商金屬瓶的耐壓性. Bạn có phải có thể nghĩ ra phương pháp kiểm định càng có hiệu quả hơn để đánh giá tính chịu ép của bình kim loại của nhà cung ứng.

Six Sigma

28

9 - 7

