



# 十六、無母數統計法

## Chapter 16 Nonparametric Statistics

# 學習目標

## ■ 知識(認知)

- ◆ 可以清楚陳述無母數統計的意涵。
- ◆ 可以清楚陳述無母數統計中兩個配對母體變數分布相等性檢定的意涵。
- ◆ 可以說明無母數統計中，適合度檢定的程序和標準。
- ◆ 可以說明無母數統計中，三個母體以上分布相同性假設檢定的程序和標準。
- ◆ 評價無母數統計中，各種適合度假設檢定和相同性假設檢定的使用價值。

## ■ 技能

- ◆ 能夠計算各種情境下的檢定統計值。
- ◆ 能夠利用檢定統計值與臨界值的比較，提出統計推論。
- ◆ 綜合所學，能夠於實務領域中，依據特定情境的需求進行假設檢定程序。

## ■ 態度(情意)

- ◆ 意識到在日常生活或未來工作環境中，無母數統計法的重要性與應用價值。
- ◆ 在各種情境下，依循假設檢定的程序，接受統計推論所傳達的意涵。

# 章節內容

## ■ 十六、無母數統計

- ◆ 16.1 無母數統計特性
- ◆ 16.2 符號檢定
- ◆ 16.3 魏克森符號等級檢定
- ◆ 16.4 魏克森等級和檢定與曼-惠特尼檢定
- ◆ 16.5 克拉斯卡-瓦歷斯檢定
- ◆ 16.6 隨機性檢定或連檢定
- ◆ 16.7 等級相關檢定或Spearman等級相關檢定

# 學習內容

- 單一母體中位數檢定
- 兩配對母體分布是否相同檢定
- 魏克森符號等級檢定
- 魏克森等級和檢定
- 曼-惠特尼檢定
- 克拉斯卡-瓦歷斯檢定
- 隨機性檢定或連檢定
- 等級相關檢定
- Spearman等級相關檢定

# 學習建議

【開始上課】下  
載PDF講義預習



依據各單元順序，  
瀏覽數位教材內  
容



針對數位教材中  
練習題目(來賓考  
試)演練



學習評量



相關學習主題討  
論



同步上課學習

# 母數統計與無母數統計

## ■ 母數統計

- ◆ 前面各章節的統計假設與推論過程中，皆須假設研究變數分布屬於常態分布，方能利用 $Z$ 、 $t$ 、 $F$ 和 $\chi^2$ 進行顯著性檢定。

## ■ 無母數統計

- ◆ 在資料數量不多的情況下，對名目尺度和順序尺度的資料進行假設檢定時，可以稱為無母數檢定 (non-parametric tests) 或自由分布檢定 (distribution-free tests)。

# 提示舊經驗：激發學習動機

餐廳、旅行社或旅館經營管理

抽樣調查  
產品品質、服務品質、消費者滿意度

非常態樣  
本少推論  
研究假設

信賴水準(推論成功機率)

無母數統計法

採取因應作為

# 16.1 無母數統計特性

## ■ 母體分布的型態不限定

- ◆ 無母數統計特別適合於質化資料的分析，類別和順序資料通常不屬於常態分布，而是屬於多項分布型態，故無法獲得平均值、變異數等母體參數，不適合於有母數統計分析，適合於無母數統計分析。

## ■ 統計推論的對象不限於特定母體參數

- ◆ 無母數統計可檢定資料的分布是否合乎特定假設分布或兩母體的分布是否相同。

## ■ 無母數統計方法的假設條件較少

- ◆ 無母數統計不需要對母體分布或母體參數進行假設。
- ◆ 無母數統計亦不需設定樣本數量大小。

# 無母數統計優缺點比較

## ■ 無母數統計方法的優點

- ◆ 母體分布未知時，可以進行統計推論比較。
- ◆ 樣本數量 $n$ 不多時，計算過程相當簡單。
- ◆ 母體屬於常態分布時，較有母數分析不易得到顯著結果；母體不是常態分布時，無母數分析之檢定力較有母數分析高。
- ◆ 具有穩健特性。

## ■ 無母數統計方法的缺點

- ◆ 母體分布已知時，與母數統計法相比，其統計分析效益較差。
- ◆ 缺乏相關機率表格可以使用。
- ◆ 針對常態分布資料若進行無母數分析時，將使檢定力降低。

# 16.2 符號檢定

- 符號檢定(sign test)是無母數統計中最常使用與最基礎的一種方法
  - 可以使用於檢定
  - ◆ 16.2.1 單一母體中央趨勢(中位數)的假設。
    - 在母數統計的區間估計與假設檢定中，若樣本數量太少時，母體分布又未知，無法進行信賴區間估計與假設檢定，在無母數統計中，可以使用符號檢定中位數。
  - ◆ 16.2.2 成對母體分布是否相同的假設檢定。
    - 在無母數統計中，特別適合於母體分布未知，且樣本數量少的情況。

## 16.2.1 單一母體中位數檢定

- 母體分布屬於偏態分布時，在評量其中央趨勢時，以選用中位數 (median) 為評量指標較為適當。
  - ◆ 當母體分布事先未知，一般以檢定其中位數等於特定數值為其假設，而非以檢定其平均值等於特定數值為其假設，較為保守。
  - ◆ 在母體分布未知的情況下，運用符號檢定以了解母體的中央趨勢。
- 當母體分布未知或已知其為偏態分布時，欲檢定母體特定變數中位數  $M_d$  是否等於特定數值  $m_0$ 。
- 從單一母體隨機抽出  $n$  個基本單位為樣本，評量其特定變數的觀測值分別為  $x_1, x_2, x_3, \dots, x_n$ 。
- 分別將觀測值  $x_i$  與特定中位數數值  $m_0$  (推測值) 比較，兩者差異為
  - ◆  $D_i = x_i - m_0$

# 符號計算

- $D_i = x_i - m_0 > 0$  獲得一個正號(+);  $D_i = x_i - m_0 < 0$  獲得一個負號(-);  $D_i = x_i - m_0 = 0$  忽略不計。
  - ◆ 分別計算在  $n$  個觀測值獲得的正號(+)和負號(-)數量。
- 當母體中位數  $M_d$  等於  $m_0$  時，獲得正號(+)和負號(-)數量應相等或非常接近。
- 若發現獲得正號(+)和負號(-)數量明顯差異很大時，可推估該母體中位數  $M_d$  不等於特定中位數數值  $m_0$  (推測值)。
- 若母體特定變數中位數  $M_d$  剛好等於特定中位數數值  $m_0$  (推測值) 時，獲得正號(+)和負號(-)機率皆為 0.5。
  - ◆ 在  $n$  個觀測值中獲得正號(+)的次數屬於二項分布，運用符號檢定即是檢測獲得正號(+)的機率  $p$  是否等於 0.50。

# 單一母體中位數檢定程序

- A. 設定顯著水準 $\alpha$
- B. 設立虛無假設和對立假設
- C. 計算獲得正號(+)的次數或負號(-)的次數
- D. 決定拒絕區域與接受區域
- E. 計算檢定統計值 $S_0$ ：獲得正號(+)的次數
- F. 統計推論

單一母體中位數檢定

樣本數量小( $n \leq 20$ )

樣本數量大( $n > 20$ )

# 樣本數量小符號檢定

- 符號檢定統計值  $S_0$ ：獲得正號 (+) 的次數，屬於二項分布， $S_0 \sim B(n_r, p = 0.5)$ 。  $n_r$  是扣除  $D_i = x_i - m_0 = 0 (x_i = m_0)$  的樣本數量。
- 臨界值法統計判斷法則
- 機率  $p$  值法統計判斷法則

# 臨界值法統計判斷法則

## ■ 雙尾檢定

- ◆  $S_L$ (臨界值)  $\leq$  檢定統計值  $S_0 \leq S_H$ (臨界值)，接受虛無假設  $H_0$
- ◆ 檢定統計值  $S_0 < S_L$ (臨界值) 或 檢定統計值  $S_0 > S_H$ (臨界值)，接受對立假設  $H_1$

## ■ 左尾檢定

- ◆ 檢定統計值  $S_0 \geq S_L$ (臨界值)，接受虛無假設  $H_0$
- ◆ 檢定統計值  $S_0 < S_L$ (臨界值)，接受對立假設  $H_1$

## ■ 右尾檢定

- ◆ 檢定統計值  $S_0 \leq S_H$ (臨界值)，接受虛無假設  $H_0$
- ◆ 檢定統計值  $S_0 > S_H$ (臨界值)，接受對立假設  $H_1$

# 臨界值法中位數檢定範例(1/4)

**範例19.1** 觀光系小蘭同學欲知道學生每週餐飲費用(假設為非常態分布)的中位數是否為新台幣1000元，以知曉其自己的花費是高於中位數還是低於中位數。小蘭隨機抽出10位學生，其每週餐飲費用如下表所示，以顯著水準 $\alpha = 0.05$ 進行符號統計檢定。

學生	1	2	3	4	5	6	7	8	9	10
費用(元)	1200	950	1230	650	780	1350	1425	1620	1000	1304

題解：假設學生每週餐飲費用為非常態分布，故採用符號檢定方式。學生餐飲費用有可能高於新台幣1000元，亦有可能低於新台幣1000元，故此例屬於雙尾檢定

A. 設定顯著水準 $\alpha = 0.05$

B. 虛無假設 $H_0$ :每週餐飲費用中位數 = 1000元 或  $H_0$ :每週餐飲費用高於1000元的機率 $p = 0.50$

C. 對立假設 $H_1$ :每週餐飲費用中位數  $\neq$  1000元 或  $H_1$ :每週餐飲費用高於1000元的機率 $p \neq 0.50$

# 臨界值法中位數檢定範例(2/4)

**範例19.1** 觀光系小蘭同學欲知道學生每週餐飲費用(假設為非常態分布)的中位數是否為新台幣1000元，以知曉其自己的花費是高於中位數還是低於中位數。小蘭隨機抽出10位學生，其每週餐飲費用如下表所示，以顯著水準  $\alpha = 0.05$  進行符號統計檢定。

## D. 計算檢定統計值 $S_0$

學生	1	2	3	4	5	6	7	8	9	10
費用(元)	1200	950	1230	650	780	1350	1425	1620	1000	1304
符號	+	-	+	-	-	+	+	+	0	+

獲得正號(+)的次數為6，故  $S_0 = 6$ ， $n_r = 10$ (樣本數) - 1(觀測值  $x_i =$  檢定中位數  $m_0$  的樣本數) = 9。檢定統計值  $S_0$  屬於二項分布  $S_0 \sim B(9, 0.50)$ 。

# 臨界值法中位數檢定範例(3/4)

**範例19.1** 觀光系小蘭同學欲知道學生每週餐飲費用(假設為非常態分布)的中位數是否為新台幣1000元，以知曉其自己的花費是高於中位數還是低於中位數。小蘭隨機抽出10位學生，其每週餐飲費用如下表所示，以顯著水準  $\alpha = 0.05$  進行符號統計檢定。

## E. 決定接受區域與拒絕區域

在雙尾檢定中顯著水準  $\alpha = 0.05$ ，故左右兩側的拒絕虛無假設的機率為  $\frac{\alpha}{2} = 0.025$

二項分布  $S \sim B(9, 0.50)$  累積機率分布  $P(x \leq t) = \sum_{x=0}^t \binom{n}{x} \times p^x \times q^{n-x}$

$n$	$t$	累積機率	雙尾機率
9	0	0.0020	
9	1	0.0195	$\frac{\alpha}{2} = 0.025 \downarrow$
9	2	0.0898	
9	3	0.2539	
9	4	0.5000	
9	5	0.7461	
9	6	0.9102	
9	7	0.9805	
9	8	0.9980	$\frac{\alpha}{2} = 0.025 \uparrow$
9	9	1.0000	

$P(S \leq 1) = 0.0195 < \frac{\alpha}{2} = 0.025$  且  $P(S \leq 2) = 0.0898 > \frac{\alpha}{2} = 0.025$ ，故左側(較低)臨界值  $S_L = 2$ 。

$P(S \leq 8) = 0.9980$ ， $P(S = 9) = 1 - P(S \leq 8) = 1 - 0.9980 = 0.0020 < \frac{\alpha}{2} = 0.025$  且  $P(S \leq 7) = 0.9805$ ， $P(S = 8) + P(S = 9) = 1 - P(S \leq 7) = 1 -$

$0.9805 = 0.0195 < \frac{\alpha}{2} = 0.025$  並  $P(S \leq 6) = 0.9102$ ， $P(S = 7) + P(S = 8) +$

$P(S = 9) = 1 - P(S \leq 6) = 1 - 0.9102 = 0.0898 > \frac{\alpha}{2} = 0.025$ ，右側(較高)臨界值  $S_H = 6$ 。

# 臨界值法中位數檢定範例(4/4)

**範例19.1** 觀光系小蘭同學欲知道學生每週餐飲費用(假設為非常態分布)的中位數是否為新台幣1000元，以知曉其自己的花費是高於中位數還是低於中位數。小蘭隨機抽出10位學生，其每週餐飲費用如下表所示，以顯著水準  $\alpha = 0.05$  進行符號統計檢定。

F. 統計推論： $S_L = 2$ (臨界值)  $\leq$  檢定統計值  $S_0 = 6 \leq S_H = 6$ (臨界值)，接受虛無假設  $H_0$ ：每週餐飲費用中位數 = 1000元。依據小蘭同學的統計推論發現每週餐飲費用的中位數為新台幣1000元。

輕鬆練習一下

此投影片播放後3分鐘內，沒有進入該練習卷者，老師即關閉  
沒有進入者的練習卷(本練習8分鐘內完成)

**上課練習112ch16\_1**

# 機率 $p$ 值法統計判斷法則

- 運用臨界值法在顯著水準 $\alpha = 0.05$ ，進行臨界值上下限的運算時比較繁雜一點。
- 運用機率 $p$ 值法比較簡便一些。
- 雙尾檢定
  - ◆ 檢定統計值 $S_0 \geq \frac{n_r}{2}$  [ $n_r$ 是扣除 $D_i = 0(x_i = m_0)$ 的樣本數量]時，機率 $p = 2 \times P(S \geq S_0 | n_r, p = 0.50)$
  - ◆ 檢定統計值 $S_0 < \frac{n_r}{2}$  [ $n_r$ 是扣除 $D_i = 0(x_i = m_0)$ 的樣本數量]時，機率 $p = 2 \times P(S \leq S_0 | n_r, p = 0.50)$
- 左尾檢定
  - ◆ 機率 $p = P(S \leq S_0 | n, p = 0.50)$
- 右尾檢定
  - ◆ 機率 $p = P(S \geq S_0 | n, p = 0.50)$

# 機率 $p$ 值法中位數檢定範例(1/3)

**範例19.2** 觀光系小蘭同學欲知道學生上週餐飲費用(假設為非常態分布)的中位數是否為新台幣1000元，以知曉其自己的花費是高於中位數還是低於中位數。小蘭隨機抽出10位學生，其上週餐飲費用如下表所示，以顯著水準 $\alpha = 0.05$ 進行統計檢定。

學生	1	2	3	4	5	6	7	8	9	10
費用(元)	1200	950	1230	650	780	1350	1425	1620	1000	1304

題解：假設學生上週餐飲費用為非常態分布，故採用符號檢定方式。學生餐飲費用有可能高於新台幣1000元，亦有可能低於新台幣1000元，故此例屬於雙尾檢定

A. 設定顯著水準 $\alpha = 0.05$

B. 虛無假設 $H_0$ : 上週餐飲費用中位數 = 1000元 或  $H_0$ : 上週餐飲費用高於1000元的機率 $p = 0.50$

C. 對立假設 $H_1$ : 上週餐飲費用中位數  $\neq$  1000元 或  $H_1$ : 上週餐飲費用高於1000元的機率 $p \neq 0.50$

# 機率 $p$ 值法中位數檢定範例(2/3)

**範例19.2** 觀光系小蘭同學欲知道學生上週餐飲費用(假設為非常態分布)的中位數是否為新台幣1000元，以知曉其自己的花費是高於中位數還是低於中位數。小蘭隨機抽出10位學生，其上週餐飲費用如下表所示，以顯著水準 $\alpha = 0.05$ 進行統計檢定。

## D. 計算檢定統計值 $S_0$

學生	1	2	3	4	5	6	7	8	9	10
費用(元)	1200	950	1230	650	780	1350	1425	1620	1000	1304
符號	+	-	+	-	-	+	+	+	0	+

獲得正號(+)的次數為6，故 $S_0 = 6$ ， $n_r = 10$ (樣本數) - 1(觀測值 $x_i =$  檢定中位數 $m_0$ 的樣本數) = 9。檢定統計值 $S_0$ 屬於二項分布 $S_0 \sim B(9, 0.50)$ 。

# 機率 $p$ 值法中位數檢定範例(3/3)

**範例19.2** 觀光系小蘭同學欲知道學生上週餐飲費用(假設為非常態分布)的中位數是否為新台幣1000元，以知曉其自己的花費是高於中位數還是低於中位數。小蘭隨機抽出10位學生，其上週餐飲費用如下表所示，以顯著水準 $\alpha = 0.05$ 進行統計檢定。

**E.** 計算機率 $p$ 值

$$S_0 = 6 > \frac{n_r}{2} = \frac{9}{2} = 4.5$$

$$\text{機率 } p = 2 \times P(S \geq S_0 | n_r, p = 0.50) = 2 \times P(S \geq 6 | 9, p = 0.50) = 2 \times [1 - P(S \leq 5)] = 2 \times [1 - 0.7461(\text{BINOM.DIST})] = 2 \times 0.2539 = 0.5078$$

**F.** 機率 $p = 0.5078 >$  顯著水準 $\alpha = 0.05$ ，接受虛無假設  $H_0$ : 上週餐飲費用中位數 = 1000元。小蘭同學的統計推論發現上週餐飲費用的中位數為新台幣1000元。

輕鬆練習一下

此投影片播放後3分鐘內，沒有進入該練習卷者，老師即關閉  
沒有進入者的練習卷(本練習8分鐘內完成)

**上課練習112ch16\_2**

# 樣本數量大符號檢定

■ 當樣本數量增加( $n \geq 20$ )時，欲檢定母體中位數 $M_d$ 等於特定值 $m_0$ 之二項分布會趨近於常態分布，可以運用標準化 $z$ 值分布進行母體中位數 $M_d$ 等於特定中位數特定值 $m_0$ (推測值)之檢定。

## ■ 假設檢定

◆ 虛無假設 $H_0: M_d = m_0$  或  $H_0: P(M_d > m_0) = 0.50$

◆ 對立假設 $H_1: M_d \neq m_0$  或  $H_0: P(M_d > m_0) \neq 0.50$

■ 在樣本數量增加( $n \geq 20$ )時，欲檢定母體中位數 $M_d$ 等於特定值 $m_0$ 之二項分布會趨近於常態分布，獲得正號(+)  
次數的平均值與標準差分別為：

◆ 平均值 $\mu_S = 0.50 \times n$

◆ 標準差 $\sigma_S = \sqrt{0.50 \times 0.50 \times n} = 0.50 \times \sqrt{n}$

■ 樣本數量增加( $n \geq 20$ )，獲得正號(+)  
次數之二項分布會趨近於常態分布： $S_0 \sim N(0.50 \times n, 0.50 \times 0.50 \times n)$

◆ 檢定統計值  $z = \frac{S - E(S)}{\sqrt{\sigma_S^2}} = \frac{S - 0.50 \times n}{0.50 \times \sqrt{n}}$  (請參閱前面章節)

## 16.2.2 檢定兩配對母體分布相同

- 符號檢定可使用於檢定兩個配對母體分布是否相同。
  - ◆ 例如檢定相同消費者對兩家連鎖速食餐廳的滿意度是否相同；相同消費者對兩個生態旅遊行程的體驗價值是否相同；相同消費者對兩種咖啡豆所烹製咖啡飲料的喜愛程度是否相同。
- 符號檢定利用兩個配對樣本的差，屬於正值或負值，轉換為正號與負號的數量，以檢定兩個配對母體的分布是否相同。
  - ◆ 假設  $(x_{1i}, x_{2i})$  為配對的樣本觀測值，兩者之間的差  $d_i$  為
    - $d_i = x_{1i} - x_{2i}$

配對樣本之差

正值或負值

正號與負號數量

二項分布

統計推論

# 符號檢定標準

$$d_i = x_{1i} - x_{2i}$$

- 當  $d_i > 0$  為正值，獲得一個正號(+);  $d_i < 0$  為負值，獲得一個負號(-);  $d_i = 0$  代表  $x_{1i}$  與  $x_{2i}$  為兩個配對樣本觀測值相等，不計入正負符號數量。
  - ◆ 兩個配對母體的分佈相同，獲得正號和負號的數量應該相等或非常相近;
  - ◆ 兩個配對母體的分佈不相同，獲得正號(+)和負號(-)的數量應該不相等或差距非常大。
  - ◆ 當兩個配對母體分佈相同時，扣除  $x_{1i}$  與  $x_{2i}$  兩個配對樣本觀測值相等的觀測值數量，分別出現正號(+)與出現負號(-)的機率  $p$  應皆為 0.50;
- 運用符號檢定兩個配對母體分佈是否相同，如同檢定在二項分佈母體出現正號(+)或出現負號(-)的機率  $p$  是否等於 0.50。

# 檢定兩配對母體分布範例(1/3)

**範例19.3** 欲了解觀光系學生對A和B兩家連鎖速食餐廳的滿意度分布(假設為非常態分布)是否相同，今隨機抽取12位同學進行調查分析，調查結果如下表所示，在顯著水準 $\alpha = 0.05$ 下，進行學生對A和B兩家連鎖速食餐廳的滿意度分布是否相同之統計檢定。

學生	1	2	3	4	5	6	7	8	9	10	11	12
A餐廳滿意度	10	9	3	6	8	5	5	6	10	4	5	6
B餐廳滿意度	2	9	9	5	7	3	4	2	0	3	7	4

**題解：**學生對兩家連鎖速食餐廳的滿意度，不同學生的感受有可能是A餐廳高於B餐廳，亦有可能是B餐廳高於A餐廳，故此例屬於雙尾檢定

**A.** 設定顯著水準 $\alpha = 0.05$

**B.** 虛無假設 $H_0$ : 學生對兩家連鎖速食餐廳滿意度分布相同 或  $H_0$ : 學生對A餐廳的滿意度高於B餐廳的機率 $p = 0.50$

**C.** 對立假設 $H_1$ : 學生對兩家連鎖速食餐廳滿意度分布不相同 或  $H_1$ : 學生對A餐廳的滿意度高於B餐廳的機率 $p \neq 0.50$

# 檢定兩配對母體分布範例(2/3)

**範例19.3** 欲了解觀光系學生對A和B兩家連鎖速食餐廳的滿意度分布(假設為非常態分布)是否相同，今隨機抽取12位同學進行調查分析，調查結果如下表所示，在顯著水準 $\alpha = 0.05$ 下，進行學生對A和B兩家連鎖速食餐廳的滿意度分布是否相同之統計檢定。

## D. 計算檢定統計值 $S_0$

設學生滿意度A餐廳高於B餐廳者為 $d_i = x_{Ai} - x_{Bi} > 0$ 可以獲得一個正號(+); 學生滿意度A餐廳低於B餐廳者為 $d_i = x_{Ai} - x_{Bi} < 0$ 可以獲得一個負號(-); 學生滿意度A餐廳等於B餐廳者為 $d_i = x_{Ai} - x_{Bi} = 0$ 則忽略不計。

學生	1	2	3	4	5	6	7	8	9	10	11	12
A餐廳滿意度 $x_{Ai}$	10	9	3	6	8	5	5	6	10	4	5	6
B餐廳滿意度 $x_{Bi}$	2	9	9	5	7	3	4	2	0	3	7	4
符號	+		-	+	+	+	+	+	+	+	-	+

滿意度有差異的樣本數量 $n_r = 12 - 1 = 11$ ，獲得正號(+)數量為 $S_0 = 9$ 。檢定統計值 $S_0$ 屬於二項分布 $S_0 \sim B(11, 0.50)$ 。

# 檢定兩配對母體分布範例(3/3)

**範例19.3** 欲了解觀光系學生對A和B兩家連鎖速食餐廳的滿意度分布(假設為非常態分布)是否相同，今隨機抽取12位同學進行調查分析，調查結果如下表所示，在顯著水準 $\alpha = 0.05$ 下，進行學生對A和B兩家連鎖速食餐廳的滿意度分布是否相同之統計檢定。

**E.** 計算機率 $p$ 值

$$S_0 = 9 > \frac{n_r}{2} = \frac{11}{2} = 5.5$$

$$\begin{aligned} \text{機率 } p &= 2 \times P(S \geq S_0 | n_r, p = 0.50) = 2 \times P(S \geq 9 | 11, p = 0.50) = 2 \times [1 - P(S \leq 8)] = 2 \\ &\times [1 - 0.9673(\text{BINOM.DIST})] = 2 \times 0.0327 = 0.0654 \end{aligned}$$

**F.** 機率 $p = 0.0654 >$  顯著水準 $\alpha = 0.05$ ，接受虛無假設 $H_0$ ：學生對兩家連鎖速食餐廳滿意度分布相同。

輕鬆練習一下

此投影片播放後3分鐘內，沒有進入該練習卷者，老師即關閉  
沒有進入者的練習卷(本練習5分鐘內完成)

**上課練習112ch16\_3**

## 16.3 魏克森符號等級檢定

- 前述符號檢定僅利用所有配對組合中觀測值有差異的樣本數  $n$  與配對組合觀測值有差異時獲得正號(+)數量或負號(-)數量  $S_0$  兩個數值，未再利用配對組合中觀測值的差異量，浪費此數值資訊的參考價值。
- 魏克森符號等級檢定法針對符號檢定的缺失進行修正，同時考慮配對組合觀測值有差異時獲得正號(+)數量或負號(-)數量  $S_0$  與配對組合中觀測值的差異量的等級。
- 魏克森符號等級檢定法先計算配對樣本觀測值之差異量  $d_i = x_{Ai} - x_{Bi}$ ，取絕對值  $|d_i| = |x_{Ai} - x_{Bi}|$ ，再依據數值  $|d_i|$  大小，由小排到大，並給予等級。
  - ◆ 當差異量數值相等  $|d_i| = |d_j|$  時，利用其平均數作為等級。
  - ◆ 不計兩配對組合觀測值相同的樣本，不排序和等級。

# 檢定統計值

- 假設 $R^+$ 為正 $d_i$ 值之等級和； $R^-$ 為負 $d_i$ 值之等級和。
  - ◆ 兩配對母體分布型態相同， $R^+$ 與 $R^-$ 數值應相等或非常相近。
  - ◆  $R^+$ 與 $R^-$ 數值有明顯的差異量，兩配對母體分布型態不相同。
- 魏克森符號等級檢定法選擇 $R^+$ 與 $R^-$ 數值為檢定統計值。
  - ◆ 雙尾檢定
    - 檢定統計值選用 $R^+$ 與 $R^-$ 數值較小者，檢定統計值 $R = \min(R^+, R^-)$ 。便於查表與計算。
  - ◆ 右尾檢定
    - 檢定統計值 $R = R^-$ ，在右尾檢定時 $R^-$ 數值比較小。
  - ◆ 左尾檢定
    - 檢定統計值 $R = R^+$ ，在左尾檢定時 $R^+$ 數值比較小。
- 不論在右尾、左尾或雙尾中，檢定統計值 $R$ 數值愈小，顯示兩個母體分布型態愈不相同。

# 魏克森符號等級檢定法統計判斷法則

## ■ 雙尾檢定

- ◆ 檢定統計值  $R \leq$  臨界值  $R_{\frac{\alpha}{2}}$ ，接受虛無假設  $H_0$
- ◆ 檢定統計值  $R >$  臨界值  $R_{\frac{\alpha}{2}}$ ，拒絕虛無假設  $H_0$ ，接受對立假設  $H_1$

## ■ 左尾檢定

- ◆ 檢定統計值  $R \leq$  臨界值  $R_{\alpha}$ ，接受虛無假設  $H_0$
- ◆ 檢定統計值  $R >$  臨界值  $R_{\alpha}$ ，拒絕虛無假設  $H_0$ ，接受對立假設  $H_1$

## ■ 右尾檢定

- ◆ 檢定統計值  $R \leq$  臨界值  $R_{\alpha}$ ，接受虛無假設  $H_0$
- ◆ 檢定統計值  $R >$  臨界值  $R_{\alpha}$ ，拒絕虛無假設  $H_0$ ，接受對立假設  $H_1$

# 大量樣本數量

■ 在大量樣本數量( $n \geq 30$ )時，檢定統計值R分布會趨近於常態分布，其平均值和變異數分別為：

◆ 平均值  $E(R) = \bar{R} = \frac{n \times (n+1)}{4}$

◆ 變異數  $V(R) = \frac{n \times (n+1) \times (2 \times n+1)}{24}$

■ 在大量樣本數量( $n \geq 30$ )時，檢定統計值  $z = \frac{R - \frac{n \times (n+1)}{4}}{\sqrt{\frac{n \times (n+1) \times (2 \times n+1)}{24}}}$

# 大量樣本數量時，統計判斷法則

## ■ 雙尾檢定

- ◆ 檢定統計值  $z \geq$  臨界值  $-\frac{z_{\alpha}}{2}$ ，接受虛無假設  $H_0$
- ◆ 檢定統計值  $z < -\frac{z_{\alpha}}{2}$ ，拒絕虛無假設  $H_0$ ，接受對立假設  $H_1$

## ■ 左尾檢定

- ◆ 檢定統計值  $z \geq$  臨界值  $-z_{\alpha}$ ，接受虛無假設  $H_0$
- ◆ 檢定統計值  $z < -z_{\alpha}$ ，拒絕虛無假設  $H_0$ ，接受對立假設  $H_1$

## ■ 右尾檢定

- ◆ 檢定統計值  $z \geq$  臨界值  $-z_{\alpha}$ ，接受虛無假設  $H_0$
- ◆ 檢定統計值  $z < -z_{\alpha}$ ，拒絕虛無假設  $H_0$ ，接受對立假設  $H_1$

# 魏克森符號等級檢定法範例(1/4)

**範例19.5** 觀光系小蘭同學欲知道學生每週餐飲費用(假設為非常態分布)的中位數是否為新台幣1000元，以知曉其自己的花費是高於中位數還是低於中位數。小蘭隨機抽出10位學生，其每週餐飲費用如下表所示，以顯著水準 $\alpha = 0.05$ 利用魏克森符號等級檢定法進行統計檢定。

學生	1	2	3	4	5	6	7	8	9	10
費用(元)	1200	950	1230	650	780	1350	1425	1620	1000	1304

題解：學生餐飲費用有可能高於新台幣1000元，亦有可能低於新台幣1000元，故此例屬於雙尾檢定(a two-tailed test)

A. 設定顯著水準 $\alpha = 0.05$

B. 虛無假設 $H_0$ :每週餐飲費用中位數 = 1000元 或  $H_0$ :每週餐飲費用高於1000元的機率 $p = 0.50$

C. 對立假設 $H_1$ :每週餐飲費用中位數  $\neq$  1000元 或  $H_1$ :每週餐飲費用高於1000元的機率 $p \neq 0.50$

D. 計算檢定統計值R  $m_0 = 1000$ 元

# 魏克森符號等級檢定法範例(2/4)

**範例19.5** 觀光系小蘭同學欲知道學生每週餐飲費用(假設為非常態分布)的中位數是否為新台幣1000元，以知曉其自己的花費是高於中位數還是低於中位數。小蘭隨機抽出10位學生，其每週餐飲費用如下表所示，以顯著水準  $\alpha = 0.05$  利用魏克森符號等級檢定法進行統計檢定。

學生 $i$	1	2	3	4	5	6	7	8	9	10
費用(元) $x_i$	1200	950	1230	650	780	1350	1425	1620	1000	1304
$d_i = x_i - m_0$	200	-50	230	-350	-220	350	425	620	0	304
$ d_i $	200	50	230	350	220	350	425	620	0	304
序位等級	2	1	4	6.5	3	6.5	8	9		5
正值等級	2		4			6.5	8	9		5
負值等級		1		6.5	3					

觀測值與特定數值 $m_0$ (推測值)有差異的樣本數  $n_r = 10 - 1 = 9$

正的 $d_i$ 值之等級和  $R^+ = 2 + 4 + 6.5 + 8 + 9 + 5 = 34.5$

負的 $d_i$ 值之等級和  $R^- = 1 + 6.5 + 3 = 10.5$

檢定統計值  $R = \min(R^+, R^-) = \min(34.5, 10.5) = 10.5 = R^-$

# 魏克森符號等級檢定法範例(3/4)

**範例19.5** 觀光系小蘭同學欲知道學生每週餐飲費用(假設為非常態分布)的中位數是否為新台幣1000元，以知曉其自己的花費是高於中位數還是低於中位數。小蘭隨機抽出10位學生，其每週餐飲費用如下表所示，以顯著水準  $\alpha = 0.05$  利用魏克森符號等級檢定法進行統計檢定。

E. 決定臨界值  $R_{\frac{\alpha}{2}} = R_{\frac{0.05}{2}} = R_{0.025}$

Wilcoxon符號等級檢定臨界值表-配對母體						
單尾 $\alpha =$	雙尾 $\alpha =$	$n = 5$	$n = 6$	$n = 7$	$n = 8$	$n = 9$
0.05	0.10	1	2	4	6	8
0.025	0.05		1	2	4	6
0.01	0.02			0	2	3
0.005	0.01				0	2

$R_{\frac{\alpha}{2}} = R_{\frac{0.05}{2}} = R_{0.025} = 6$  (查Wilcoxon符號等級檢定臨界值表)

## 魏克森符號等級檢定法範例(4/4)

**範例19.5** 觀光系小蘭同學欲知道學生每週餐飲費用(假設為非常態分布)的中位數是否為新台幣1000元，以知曉其自己的花費是高於中位數還是低於中位數。小蘭隨機抽出10位學生，其每週餐飲費用如下表所示，以顯著水準 $\alpha = 0.05$ 利用魏克森符號等級檢定法進行統計檢定。

**F. 檢定統計值** $R = 10.5 >$  **臨界值** $R_{\frac{\alpha}{2}} = R_{\frac{0.05}{2}} = R_{0.025} = 6$ ，接受對立假設 $H_1$ :每週餐飲費用中位數  $\neq 1000$ 元。

# 魏克森符號等級檢定法檢定力較佳

- 使用魏克森符號等級檢定法時比符號檢定法多運用差異量大小的資訊，使得魏克森符號等級檢定法檢定力較佳，統計推論比較可靠。

輕鬆練習一下

此投影片播放後3分鐘內，沒有進入該練習卷者，老師即關閉  
沒有進入者的練習卷(本練習8分鐘內完成)

**上課練習112ch16\_4**

## 16.4 魏克森等級和檢定與曼-惠特尼檢定

- 欲檢定兩個獨立母體平均值之差，在小樣本( $n < 30$ )情況需要假設母體屬於常態分布，運用有母數統計法，當兩母體變異數已知時，採用標準化 $z$ 值分布；兩母體變異數未知時，採用 $t$ 值分布。
- 小樣本( $n < 30$ )情況下需要假設母體屬於非常態分布，就必須使用無母數統計法中的魏克森等級和檢定和曼-惠特尼 $U$ 檢定。
  - ◆ 當母體分布無法確定或無法獲得準確地(樣本)觀測值之數值，但可以明確的知道觀測值之數值大小排列順序時，即可使用魏克森等級和檢定和曼-惠特尼 $U$ 檢定(簡稱M-W檢定)，進行兩個獨立母體的平均值或分布是否相等的檢定。

## 16.4.1 魏克森等級和檢定

- 魏克森等級和檢定法適用於兩個獨立母體分布未知，同時隨機抽樣獲得的樣本數量少( $n < 30$ )之情況。
  - ◆ 兩個相互獨立母體，從母體1隨機抽出 $n_1$ 個樣本，從母體2隨機抽出 $n_2$ 個樣本，再將所有 $n = n_1 + n_2$ 個樣本混合，依據其觀測值大小，由小排到大排序，並給予序位等級。
- 魏克森等級和檢定法中檢定統計值為兩個獨立母體樣本的個別等級和，通常使用樣本數量 $n$ 較少的母體樣本之等級和為檢定統計值，
  - 母體1樣本數量 $n_1 <$  母體2樣本數量 $n_2$ ，使用母體1樣本的等級和 $W_1$ 為檢定統計值；
  - 母體1樣本數量 $n_1 =$  母體2樣本數量 $n_2$ ，兩個獨立母體任選一個樣本的等級和 $W_1$ 或 $W_2$ 為檢定統計值。
  - ◆ 兩個獨立母體樣本的等級和數值相等和非常接近，顯示兩個母體的分布沒有差異；
  - ◆ 兩個獨立母體樣本的等級和數值相距甚遠，顯示兩個母體的分布有差異。

# 魏克森等級和檢定法統計判斷法則

## ■ 雙尾檢定

- ◆  $W_L$ (臨界值)  $\leq$  檢定統計值  $W \leq W_H$ (臨界值)，接受虛無假設  $H_0$
- ◆ 檢定統計值  $W < W_L$ (臨界值) 或 檢定統計值  $W > W_H$ (臨界值)，拒絕虛無假設  $H_0$ ，接受對立假設  $H_1$ 
  - $W_L$  是在左側(較低)臨界值； $W_H$  是在右側(較高)臨界值，運用較低樣本數  $n_1$ 、較高樣本數  $n_2$  和顯著水準  $\alpha$  三個參數查詢魏克森等級和檢定表即可獲得此臨界值

# 單尾檢定

- 設母體1為樣本數量較少者，當母體1分布位於母體2的右側時，檢定統計值  $W_1 \leq W_H$  (臨界值)，接受虛無假設  $H_0$
- 設母體1為樣本數量較少者，當母體1分布位於母體2的右側時，檢定統計值  $W_1 > W_H$  (臨界值)，接受對立假設  $H_1$
- 設母體1為樣本數量較少者，當母體1分布位於母體2的左側時，檢定統計值  $W_1 \geq W_L$  (臨界值)，接受虛無假設  $H_0$
- 設母體1為樣本數量較少者，當母體1分布位於母體2的左側時，檢定統計值  $W_1 < W_L$  (臨界值)，接受對立假設  $H_1$

# 魏克森等級和檢定法範例(1/3)

**範例19.6** 欲了解觀光系學生對A和B兩家連鎖速食餐廳的滿意度分布是否相同，若滿意度分布屬於**非常態分布**，今分別隨機抽取10和12位同學進行調查分析，調查結果如下表所示，在顯著水準 $\alpha = 0.05$ 下，利用**魏克森等級和檢定法**，進行學生對A和B兩家連鎖速食餐廳的滿意度分布是否相同之統計檢定。

學生	1	2	3	4	5	6	7	8	9	10	11	12
A餐廳滿意度	10	9	3	6	8	5	5	6	10	4		
B餐廳滿意度	2	9	9	5	7	3	4	2	0	3	7	4

題解：學生對兩家連鎖速食餐廳的滿意度，不同學生的感受有可能是A餐廳高於B，亦有可能是B餐廳高於A，故此例屬於雙尾檢定

- A. 設定顯著水準 $\alpha = 0.05$
- B. 虛無假設 $H_0$ : 學生對兩家連鎖速食餐廳滿意度分布相同
- C. 對立假設 $H_1$ : 學生對兩家連鎖速食餐廳滿意度分布不相同
- D. 計算檢定統計值W

## 魏克森等級和檢定法範例(2/3)

**範例19.6** 欲了解觀光系學生對A和B兩家連鎖速食餐廳的滿意度分布是否相同，若滿意度分布屬於非常態分布，今分別隨機抽取10和12位同學進行調查分析，調查結果如下表所示，在顯著水準 $\alpha = 0.05$ 下，利用魏克森等級和檢定法，進行學生對A和B兩家連鎖速食餐廳的滿意度分布是否相同之統計檢定。

將兩家餐廳的樣本觀測值數值合併排列其序位等級，若兩觀測值數值相同時，取其平均值為其序位等級。

學生	1	2	3	4	5	6	7	8	9	10	11	12
A餐廳滿意度 $x_{Ai}$	10	9	3	6	8	5	5	6	10	4		
序位等級	21.5	19	5	13.5	17	11	11	13.5	21.5	8		
B餐廳滿意度 $x_{Bi}$	2	9	9	5	7	3	4	2	0	3	7	4
序位等級	2.5	19	19	11	15.5	5	8	2.5	1	5	15.5	8

A餐廳的樣本數 $n_A = 10 < B$ 餐廳的樣本數 $n_B = 12$ ，檢定統計值為A餐廳樣本的序位等級和 $W_A = 21.5 + 19 + 5 + 13.5 + 17 + 11 + 11 + 13.5 + 21.5 + 8 = 141$

## 魏克森等級和檢定法範例(3/3)

**範例19.6** 欲了解觀光系學生對A和B兩家連鎖速食餐廳的滿意度分布是否相同，若滿意度分布屬於非常態分布，今分別隨機抽取10和12位同學進行調查分析，調查結果如下表所示，在顯著水準 $\alpha = 0.05$ 下，利用魏克森等級和檢定法，進行學生對A和B兩家連鎖速食餐廳的滿意度分布是否相同之統計檢定。

E.  $W_L = 89$ (臨界值)  $\leq$  檢定統計值  $W = 141 \leq W_H = 141$ (臨界值)，接受虛無假設 $H_0$ : 學生對兩家連鎖速食餐廳滿意度分布相同。

輕鬆練習一下

此投影片播放後3分鐘內，沒有進入該練習卷者，老師即關閉  
沒有進入者的練習卷(本練習8分鐘內完成)

**上課練習112ch16\_5**

## 16.4.2曼-惠特尼U檢定

■ 曼-惠特尼U檢定(Mann-Whitney U test)與魏克森等級和檢定法類似，皆是使用於檢定兩個獨立母體分布是否相同，皆是運用兩個獨立母體個別樣本的等級和進行檢定。

### ■ 曼-惠特尼U檢定

- ◆ 小量樣本數( $n_1 \leq 10; n_2 \leq 10$ )
- ◆ 大量樣本數( $n_1 \geq 10; n_2 \geq 10$ )

# 小量樣本數

## ■ 檢定程序

◆ 母體1隨機抽出 $n_1$ 個樣本與母體2隨機抽出 $n_2$ 個樣本，混合排序，由小排到大，分別計算其序位等級和 $W_1$ 和 $W_2$ 。

◆ 計算兩個參數值 $U_1$ 和 $U_2$

$$\square U_1 = n_1 \times n_2 + \frac{n_1 \times (n_1 + 1)}{2} - W_1 \quad U_2 = n_1 \times n_2 + \frac{n_2 \times (n_2 + 1)}{2} - W_2$$

●  $n_1 \times n_2 + \frac{n_1 \times (n_1 + 1)}{2}$ 是母體1樣本的最大序位等級和； $n_1 \times n_2 + \frac{n_2 \times (n_2 + 1)}{2}$ 是母體2樣本的最大序位等級和。

●  $U_1 + U_2 = n_1 \times n_2$ 。

◆ 計算檢定統計值 $U$

□ 雙尾檢定  $U = \min(U_1, U_2)$

□ 單尾檢定，對立假設 $H_1$ ：母體1分布位於母體2的右側， $U = U_1$

□ 單尾檢定，對立假設 $H_1$ ：母體1分布位於母體2的左側， $U = U_2$

◆ 利用曼-惠特尼 $U$ 統計機率表計算機率 $p$ 值

# 曼-惠特尼U統計機率判斷法則

## ■ 雙尾檢定

◆ 機率  $p > \frac{\alpha}{2}$ ，接受虛無假設  $H_0$

◆ 機率  $p < \frac{\alpha}{2}$ ，接受對立假設  $H_1$

## ■ 單尾檢定

◆ 機率  $p > \alpha$ ，接受虛無假設  $H_0$

◆ 機率  $p < \alpha$ ，接受對立假設  $H_1$

# 曼-惠特尼U檢定法範例(1/3)

**範例19.7** 欲了解學生對A和B兩家連鎖速食餐廳的滿意度分布是否相同，滿意度屬於非常態分布，分別隨機抽取5和5位同學進行調查分析，調查結果如下表，在顯著水準 $\alpha = 0.05$ ，利用曼-惠特尼U檢定法，進行學生對A和B兩家連鎖速食餐廳的滿意度分布是否相同之統計檢定。

學生	1	2	3	4	5
A餐廳滿意度	10	9	3	6	8
B餐廳滿意度	2	9	9	5	7

**題解：**學生對兩家連鎖速食餐廳的滿意度，不同學生的感受有可能是A餐廳高於B，亦有可能是B餐廳高於A，故此例屬於雙尾檢定

- A.** 設定顯著水準 $\alpha = 0.05$
- B.** 虛無假設 $H_0$ ：學生對兩家連鎖速食餐廳滿意度分布相同
- C.** 對立假設 $H_1$ ：學生對兩家連鎖速食餐廳滿意度分布不相同
- D.** 計算檢定統計值U：將兩家餐廳的樣本觀測值數值合併排列其序位等級，若兩觀測值數值相同時，取其平均值為其序位等級。

# 曼-惠特尼U檢定法範例(2/3)

**範例19.7** 欲了解學生對A和B兩家連鎖速食餐廳的滿意度分布是否相同，滿意度屬於非常態分布，分別隨機抽取5和5位同學進行調查分析，調查結果如下表，在顯著水準 $\alpha = 0.05$ ，利用曼-惠特尼U檢定法，進行學生對A和B兩家連鎖速食餐廳的滿意度分布是否相同之統計檢定。

學生	1	2	3	4	5	序位等級和W
A餐廳滿意度 $x_{Ai}$	10	9	3	6	8	
序位等級	10	8	2	4	6	30
B餐廳滿意度 $x_{Bi}$	2	9	9	5	7	
序位等級	1	8	8	3	5	25

$$U_A = n_A \times n_B + \frac{n_A \times (n_A + 1)}{2} - W_A = 5 \times 5 + \frac{5 \times (5 + 1)}{2} - 30 = 25 + 15 - 30 = 10$$

$$U_B = n_A \times n_B + \frac{n_B \times (n_B + 1)}{2} - W_B = 5 \times 5 + \frac{5 \times (5 + 1)}{2} - 25 = 25 + 15 - 25 = 15$$

$$\text{雙尾檢定 } U = \min(U_A, U_B) = \min(10, 15) = 10$$

## 曼-惠特尼U檢定法範例(3/3)

**範例19.7** 欲了解學生對A和B兩家連鎖速食餐廳的滿意度分布是否相同，滿意度屬於非常態分布，分別隨機抽取5和5位同學進行調查分析，調查結果如下表，在顯著水準 $\alpha = 0.05$ ，利用曼-惠特尼U檢定法，進行學生對A和B兩家連鎖速食餐廳的滿意度分布是否相同之統計檢定。

**E.** 從附錄Mann-Whitney U統計機率表 $n_A = 5$ ， $n_B = 5$ ， $U = 10$ 查表獲得 $P(U \leq 10) = 0.3452 > \frac{\alpha}{2} = \frac{0.05}{2}$ ，接受虛無假設 $H_0$ ：學生對兩家連鎖速食餐廳滿意度分布相同。

輕鬆練習一下

此投影片播放後3分鐘內，沒有進入該練習卷者，老師即關閉  
沒有進入者的練習卷(本練習8分鐘內完成)

**上課練習112ch16\_6**

# 大量樣本數

■ 兩個母體樣本數量 $n$ 皆高於10時，若虛無假設為真，統計值 $U$ 屬於常態分布，可使用標準化 $Z$ 分布進行統計檢定。

◆ 統計值 $U$ 的期望值 $E(U) = \frac{n_1 \times n_2}{2}$

◆ 統計值 $U$ 的變異數 $V(U) = \sigma_U^2 = \frac{n_1 \times n_2 \times (n_1 + n_2 + 1)}{12}$

◆ 統計值 $U$ 的標準差 $\sigma_U = \sqrt{\sigma_U^2} = \sqrt{\frac{n_1 \times n_2 \times (n_1 + n_2 + 1)}{12}}$

■ 檢定統計值 $z = \frac{U - E(U)}{\sigma_U} = \frac{U - \frac{n_1 \times n_2}{2}}{\sqrt{\frac{n_1 \times n_2 \times (n_1 + n_2 + 1)}{12}}}$

# 統計判斷法則

## ■ 雙尾檢定

◆ 臨界值  $-\frac{z_{\alpha}}{2} \leq$  檢定統計值  $z \leq$  臨界值  $\frac{z_{\alpha}}{2}$ ，接受虛無假設  $H_0$

◆ 檢定統計值  $z < -\frac{z_{\alpha}}{2}$  或 檢定統計值  $z > \frac{z_{\alpha}}{2}$ ，接受對立假設  $H_1$

## ■ 左尾檢定

◆ 檢定統計值  $z \geq$  臨界值  $-z_{\alpha}$ ，接受虛無假設  $H_0$

◆ 檢定統計值  $z < -z_{\alpha}$ ，接受對立假設  $H_1$

## ■ 右尾檢定

◆ 檢定統計值  $z \leq$  臨界值  $z_{\alpha}$ ，接受虛無假設  $H_0$

◆ 檢定統計值  $z > z_{\alpha}$ ，接受對立假設  $H_1$

# 曼-惠特尼U檢定法範例(1/3)

**範例19.8** 欲了解學生對A和B兩家連鎖速食餐廳滿意度分布是否相同，該分布屬於非常態，分別隨機抽取10和12位同學進行調查分析，調查結果如下表所示，在顯著水準 $\alpha = 0.05$ 下，利用曼-惠特尼U檢定進行學生對A和B兩家連鎖速食餐廳的滿意度分布是否相同之統計檢定。

學生	1	2	3	4	5	6	7	8	9	10	11	12
A餐廳滿意度	10	9	3	6	8	5	5	6	10	4		
B餐廳滿意度	2	9	9	5	7	3	4	2	0	3	7	4

題解：學生對兩家連鎖速食餐廳的滿意度，不同學生感受有可能是A餐廳高於B，亦有可能是B餐廳高於A，屬於雙尾檢定

A. 設定顯著水準 $\alpha = 0.05$ ，左側臨界值 $-z_{\frac{\alpha}{2}} = -z_{\frac{0.05}{2}} = -1.96$ ；右側臨界值 $z_{\frac{\alpha}{2}} = z_{\frac{0.05}{2}} = 1.96$ (NORM.S.INV)。

B. 虛無假設 $H_0$ : 學生對兩家連鎖速食餐廳滿意度分布相同

C. 對立假設 $H_1$ : 學生對兩家連鎖速食餐廳滿意度分布不相同

D. 計算檢定統計值z

# 曼-惠特尼U檢定法範例(2/3)

**範例19.8** 欲了解學生對A和B兩家連鎖速食餐廳滿意度分布是否相同，該分布屬於非常態，分別隨機抽取10和12位同學進行調查分析，調查結果如下表所示，在顯著水準 $\alpha = 0.05$ 下，利用曼-惠特尼U檢定進行學生對A和B兩家連鎖速食餐廳的滿意度分布是否相同之統計檢定。

將兩家餐廳的樣本觀測值數值合併排列其序位等級，若兩觀測值數值相同時，取其平均值為其序位等級。

學生	1	2	3	4	5	6	7	8	9	10	11	12	等級和W
A餐廳滿意度 $x_{Ai}$	10	9	3	6	8	5	5	6	10	4			
序位等級	21.5	19	5	13.5	17	11	11	13.5	21.5	8			141
B餐廳滿意度 $x_{Bi}$	2	9	9	5	7	3	4	2	0	3	7	4	
序位等級	2.5	19	19	11	15.5	5	8	2.5	1	5	15.5	8	112

$$U_A = n_A \times n_B + \frac{n_A \times (n_A + 1)}{2} - W_A = 10 \times 12 + \frac{10 \times (10 + 1)}{2} - 141 = 120 + 55 - 141 = 34$$

$$U_B = n_A \times n_B + \frac{n_B \times (n_B + 1)}{2} - W_B = 10 \times 12 + \frac{12 \times (12 + 1)}{2} - 112 = 120 + 78 - 112 = 86$$

## 曼-惠特尼U檢定法範例(3/3)

**範例19.8** 欲了解學生對A和B兩家連鎖速食餐廳滿意度分布是否相同，該分布屬於非常態，分別隨機抽取10和12位同學進行調查分析，調查結果如下表所示，在顯著水準 $\alpha = 0.05$ 下，利用曼-惠特尼U檢定進行學生對A和B兩家連鎖速食餐廳的滿意度分布是否相同之統計檢定。

$$\text{雙尾檢定 } U = \min(U_A, U_B) = \min(34, 86) = 34$$

$$\text{統計值 } U \text{ 的期望值 } E(U) = \frac{n_A \times n_B}{2} = \frac{10 \times 12}{2} = 60$$

$$\text{統計值 } U \text{ 的變異數 } V(U) = \sigma_U^2 = \frac{n_A \times n_B \times (n_A + n_B + 1)}{12} = \frac{10 \times 12 \times (10 + 12 + 1)}{12} = 230$$

$$\text{檢定統計值 } z = \frac{U - E(U)}{\sigma_U} = \frac{34 - 60}{\sqrt{230}} = \frac{-26}{15.1658} = -1.7144$$

E. 左側臨界值  $-z_{\frac{\alpha}{2}} = -1.96 \leq \text{檢定統計值 } z = -1.7144 \leq \text{右側臨界值 } z_{\frac{\alpha}{2}} = 1.96$ ，接受虛無假

設  $H_0$ : 學生對兩家連鎖速食餐廳滿意度分布相同。

輕鬆練習一下

此投影片播放後3分鐘內，沒有進入該練習卷者，老師即關閉  
沒有進入者的練習卷(本練習5分鐘內完成)

**上課練習112ch16\_7**

## 16.5 克拉斯卡-瓦歷斯檢定

- 母數統計法中變異數分析，皆會假設不同的母體分布皆屬於常態分布，同時其變異數相等。
  - ◆ 變異數分析不適用於順序尺度的數值型態之分析。
- 運用克拉斯卡-瓦歷斯檢定法(Kruskal-Wallis test, K-W檢定)於順序尺度的數值型態，與母數統計法中的變異數分析目的雷同。
  - ◆ 克拉斯卡-瓦歷斯檢定法可以視為曼-惠特尼U檢定的一種延伸運用，以檢定三個或三個以上獨立母體分布是否相同。
  - ◆ 克拉斯卡-瓦歷斯檢定法使用於檢定 $k$ 組獨立樣本是否來自相同的一個母體或中位數相等的 $k$ 組母體。

# 克拉斯卡-瓦歷斯檢定程序

A. 設定顯著水準 $\alpha$ 。

B. 虛無假設 $H_0: M_{d1} = M_{d2} = M_{d3} = \dots = M_{dk}$ 或 $H_0: k$ 個母體分布皆相同。

C. 對立假設 $H_1: M_{di} \neq M_{dj}$ 或 $H_0: k$ 個母體中至少有兩個母體之間的分布不相同。

D. 計算檢定統計值：

$$\blacklozenge W = \frac{12}{n_t \times (n_t + 1)} \times \sum_{i=1}^k \frac{(\sum_{i=1}^k R_i)^2}{n_i} - 3 \times (n_t + 1)$$

E. 檢定統計值 $W \leq$  臨界值 $\chi_{\alpha, k-1}^2$ ，推論虛無假設 $H_0: M_{d1} = M_{d2} = M_{d3} = \dots = M_{dk}$ 成立。

F. 檢定統計值 $W >$  臨界值 $\chi_{\alpha, k-1}^2$ ，推論對立假設 $H_1: M_{di} \neq M_{dj}$ 成立。

# 事後比較公式

$$\blacksquare \text{CD} = z \frac{\alpha}{k \times (k-1)} \times \sqrt{\frac{N \times (N+1)}{12} \times \left( \frac{1}{n_i} + \frac{1}{n_j} \right)}$$

- $\bar{R}_i$ : 代表第*i*組獨立個別樣本在全部樣本中的順位(rank)之平均值
- $\bar{R}_j$ : 代表第*j*組獨立個別樣本在全部樣本中的順位(rank)之平均值
- $z \frac{\alpha}{k \times (k-1)}$ : 代表在機率為 $\frac{\alpha}{k \times (k-1)}$ 時的標準化z值。例：在 $\alpha = 0.05$ 和 $k = 3$ 條件， $z = 2.3940$ 。
- $N$ : 代表全部樣本數量
- $n_i$ : 代表第*i*組獨立樣本的樣本數量
- $k$ : 獨立樣本的組數

## 顯著性判斷

- ◆  $|\bar{R}_i - \bar{R}_j| < \text{CD}$  代表第*i*組和第*j*組中位數之間經事後比較後，發現沒有顯著性差異。
- ◆  $|\bar{R}_i - \bar{R}_j| > \text{CD}$  代表第*i*組和第*j*組中位數之間經事後比較後，發現達到顯著性差異水準。

# 克拉斯卡-瓦歷斯檢定法範例(1/3)

**範例19.9** 欲了解學生對A、B和C三家連鎖餐廳滿意度分布是否相同，滿意度屬於非常態，分別隨機抽取8、9和7位同學前往進行調查分析，結果如下表，在顯著水準 $\alpha = 0.05$ 下，利用克拉斯卡-瓦歷斯檢定法，進行學生對三家連鎖餐廳的滿意度分布是否相同之統計檢定。

學生	1	2	3	4	5	6	7	8	9
A餐廳滿意度	5	7	3	6	8	5	5	6	
B餐廳滿意度	2	6	9	5	7	3	4	2	0
C餐廳滿意度	10	9	9	8	7	6	8		

題解：

A. 設定顯著水準  $\alpha = 0.05$ ，臨界值  $\chi_{\alpha, k-1}^2 = \chi_{0.05, 3-1}^2 = \chi_{0.05, 2}^2 = 5.9915(\text{CHISQ.INV.RT})$ 。

B. 虛無假設  $H_0$ : 學生對三家連鎖速食餐廳滿意度分布相同

C. 對立假設  $H_1$ : 學生對三家連鎖速食餐廳滿意度分布不完全相同

D. 計算檢定統計值  $W$

# 克拉斯卡-瓦歷斯檢定法範例(2/3)

**範例19.9** 欲了解學生對A、B和C三家連鎖餐廳滿意度分布是否相同，滿意度屬於非常態，分別隨機抽取8、9和7位同學前往進行調查分析，結果如下表，在顯著水準 $\alpha = 0.05$ 下，利用克拉斯卡-瓦歷斯檢定法，進行學生對三家連鎖餐廳的滿意度分布是否相同之統計檢定。

將三家餐廳的樣本觀測值數值合併排列其序位等級，若兩觀測值數值相同時，取其平均值為其序位等級。

學生	1	2	3	4	5	6	7	8	9	等級和W
A餐廳滿意度 $x_{Ai}$	5	7	3	6	8	5	5	6		
序位等級	8.5	16	4.5	12.5	19	8.5	8.5	12.5		90
B餐廳滿意度 $x_{Bi}$	2	6	9	5	7	3	4	2	0	
序位等級	2.5	12.5	22	8.5	16	4.5	6	2.5	1	75.5
C餐廳滿意度 $x_{Ci}$	10	9	9	8	7	6	8			
序位等級	24	22	22	19	16	12.5	19			134.5

$$n_t = n_A + n_B + n_C = 8 + 9 + 7 = 24$$

# 克拉斯卡-瓦歷斯檢定法範例(3/3)

**範例19.9** 欲了解學生對A、B和C三家連鎖餐廳滿意度分布是否相同，滿意度屬於非常態，分別隨機抽取8、9和7位同學前往進行調查分析，結果如下表，在顯著水準 $\alpha = 0.05$ 下，利用克拉斯卡-瓦歷斯檢定法，進行學生對三家連鎖餐廳的滿意度分布是否相同之統計檢定。

$$\text{檢定統計值 } W = \frac{12}{n_t \times (n_t + 1)} \times \sum_{i=1}^k \frac{(\sum_{i=1}^k R_i)^2}{n_i} - 3 \times (n_t + 1) = \frac{12}{24 \times (24 + 1)} \times \left[ \frac{90^2}{8} + \frac{75.5^2}{9} + \frac{134.5^2}{7} \right] - 3 \times (24 + 1) = \frac{12}{600} \times (1012.5000 + 633.3611 + 2584.3214) - 75 = 84.6037 - 75 = 9.6037$$

**E.** 檢定統計值  $W = 9.6037 >$  臨界值  $\chi_{\alpha, k-1}^2 = 5.9915$ ，接受對立假設  $H_1$ ：學生對三家連鎖速食餐廳滿意度分布不完全相同。

# A和B兩家連鎖速食餐廳滿意度事後比較

$$\blacksquare \text{CD}_{AB} = Z_{\frac{\alpha}{k \times (k-1)}} \times \sqrt{\frac{N \times (N+1)}{12} \times \left( \frac{1}{n_A} + \frac{1}{n_B} \right)} = Z_{\frac{0.05}{3 \times (3-1)}} \times$$

$$\sqrt{\frac{24 \times (24+1)}{12} \times \left( \frac{1}{8} + \frac{1}{9} \right)} = Z_{0.00833} \times 3.4359 = 2.3940 \times 3.4359 = 8.2255$$

$$\blacklozenge \bar{R}_A = \frac{W_A}{n_A} = \frac{90}{8} = 11.25 ; \bar{R}_B = \frac{W_B}{n_B} = \frac{75.5}{9} = 8.3889$$

■  $|\bar{R}_A - \bar{R}_B| = 11.25 - 8.3889 = 2.8611 < \text{CD}_{AB} = 8.2255$ ，代表第A組和第B組中位數之間經事後比較後，發現沒有顯著性差異。

# A和C兩家連鎖速食餐廳滿意度事後比較

$$\blacksquare \text{CD}_{AC} = Z_{\frac{\alpha}{k \times (k-1)}} \times \sqrt{\frac{N \times (N+1)}{12} \times \left( \frac{1}{n_A} + \frac{1}{n_C} \right)} = Z_{\frac{0.05}{3 \times (3-1)}} \times$$

$$\sqrt{\frac{24 \times (24+1)}{12} \times \left( \frac{1}{8} + \frac{1}{7} \right)} = Z_{0.00833} \times 3.6596 = 2.3940 \times 3.6596 = 8.7611$$

$$\blacklozenge \bar{R}_A = \frac{W_A}{n_A} = \frac{90}{8} = 11.25 ; \bar{R}_C = \frac{W_C}{n_C} = \frac{134.5}{7} = 19.2143$$

■  $|\bar{R}_A - \bar{R}_C| = |11.25 - 19.2143| = 7.9643 < \text{CD}_{AC} = 8.7611$ ，代表第A組和第C組中位數之間經事後比較後，發現沒有達到顯著性差異水準。

# B和C兩家連鎖速食餐廳滿意度事後比較

$$\blacksquare CD_{BC} = Z_{\frac{\alpha}{k \times (k-1)}} \times \sqrt{\frac{N \times (N+1)}{12} \times \left( \frac{1}{n_B} + \frac{1}{n_C} \right)} = Z_{\frac{0.05}{3 \times (3-1)}} \times$$

$$\sqrt{\frac{24 \times (24+1)}{12} \times \left( \frac{1}{9} + \frac{1}{7} \right)} = Z_{0.00833} \times 3.5635 = 2.3940 \times 3.5635 = 8.5309$$

$$\blacklozenge \bar{R}_B = \frac{W_B}{n_B} = \frac{75.5}{9} = 8.3889 ; \bar{R}_C = \frac{W_C}{n_C} = \frac{134.5}{7} = 19.2143$$

■  $|\bar{R}_B - \bar{R}_C| = |8.3889 - 19.2143| = 10.8254 > CD_{BC} = 8.5309$ ，代表第B組和第C組中位數之間經事後比較後，發現有達到顯著性差異水準。

◆  $\bar{R}_B = 8.3889 < \bar{R}_C = 19.2143$ ，C連鎖速食餐廳滿意度有顯著性的比B連鎖速食餐廳滿意度高。

輕鬆練習一下

此投影片播放後3分鐘內，沒有進入該練習卷者，老師即關閉  
沒有進入者的練習卷(本練習10分鐘內完成)

**上課練習112ch16\_8**

## 16.6 隨機性檢定

- 進行統計推論時，採用簡單隨機抽樣獲得的樣本資料，可能產生的誤差相較於非隨機樣本為小。
  - ◆ 採用隨機抽樣獲得的樣本，對母體的代表性較高。
- 檢驗抽樣所獲得的樣本分布，其**是否**為**隨機抽樣**所獲得，就變得**非常重要**。
- 迴歸分析中觀測值與預測值之殘差，必須具備隨機特性，若非隨機特性，則不符合迴歸分析的基本假設，故檢驗觀測值與預測值之殘差是否為隨機亦非常重要。

# 連(runs)

■ 一排連續數列(集合)中，相鄰相同元素的最大子數列(子集合)，即成為一個連(run)。

◆ 例如在成功(S)與失敗(F)的範例中，產生下列結果的集合。

□ S S S F F F F S S S F S S S S

□ 前面連續三個相同成功S，構成S的最大子數列，接續連續四個失敗F，構成F的最大子數列，依序，前述15個成功與失敗元素中，一共發現5個連(runs)。

◆ 一個數列中，連的數量太多或太少都代表此數列出現成功與失敗的分布不隨機出現，隨機性檢定或連檢定皆屬於雙尾檢定。

■ 小量樣本數( $n_1 \leq 10$ 和 $n_2 \leq 10$ )

■ 大量樣本數( $n_1 \geq 10$ 和 $n_2 \geq 10$ )

# 小量樣本數檢定程序

A. 設定顯著水準  $\alpha = 0.05$

B. 虛無假設  $H_0$ : 樣本觀測值屬於隨機分布

C. 對立假設  $H_1$ : 樣本觀測值不屬於隨機分布

D. 計算檢定統計值—連數R

◆ 樣本觀測值大於中位數者以H符號標示；樣本觀測值小於中位數者以L符號標示。

◆ 分別計算H和L符號的出現次數，將L符號的出現個數假設為  $n_1$ ；H符號的出現個數假設為  $n_2$ 。

◆ 查連檢定累計機率表  $P(R \leq r)$  機率，其中  $r$  為連數，數值範圍大於等於2的正整數；統計值連數的觀測值。

◆ 計算機率  $p$  值

$$\square r \geq \frac{2 \times n_1 \times n_2}{n_1 + n_2} + 1, \text{ 機率 } p = 2 \times P(R \geq r | n_1, n_2)$$

$$\square r < \frac{2 \times n_1 \times n_2}{n_1 + n_2} + 1, \text{ 機率 } p = 2 \times P(R \leq r | n_1, n_2)$$

# 統計推論

## E. 決定統計推論

◆  $r \geq \frac{2 \times n_1 \times n_2}{n_1 + n_2} + 1$  , 機率  $p = 2 \times P(R \geq r | n_1, n_2)$

□ 機率  $p = 2 \times P(R \geq r | n_1, n_2) \geq$  顯著水準  $\alpha$  , 接受虛無假設  $H_0$

□ 機率  $p = 2 \times P(R \geq r | n_1, n_2) <$  顯著水準  $\alpha$  , 接受對立假設  $H_1$

◆  $r < \frac{2 \times n_1 \times n_2}{n_1 + n_2} + 1$  , 機率  $p = 2 \times P(R \leq r | n_1, n_2)$

□ 機率  $p = 2 \times P(R \leq r | n_1, n_2) \geq$  顯著水準  $\alpha$  , 接受虛無假設  $H_0$

□ 機率  $p = 2 \times P(R \leq r | n_1, n_2) <$  顯著水準  $\alpha$  , 接受對立假設  $H_1$

# 大量樣本數

■ 檢定統計值R的分布會趨近於常態分布，可用標準化Z值運算機率，在雙尾檢定方法進行統計推論。

## ■ 檢定程序

A. 設定顯著水準 $\alpha = 0.05$

B. 虛無假設 $H_0$ : 樣本觀測值屬於隨機分布

C. 對立假設 $H_1$ : 樣本觀測值不屬於隨機分布

D. 計算檢定統計值-z值

$$\square \text{檢定統計值 } z = \frac{R - E(R)}{\sigma_R} = \frac{R - \left(\frac{2 \times n_1 \times n_2}{n_1 + n_2} + 1\right)}{\sqrt{\frac{2 \times n_1 \times n_2 \times (2 \times n_1 \times n_2 - n_1 - n_2)}{(n_1 + n_2)^2 \times (n_1 + n_2 - 1)}}$$

E. 決定統計推論

□ 臨界值 $-\frac{z\alpha}{2} \leq \text{檢定統計值 } z \leq \text{臨界值 } \frac{z\alpha}{2}$ ，接受虛無假設 $H_0$

□ 檢定統計值 $z < \text{臨界值 } -\frac{z\alpha}{2}$ 或檢定統計值 $z > \text{臨界值 } \frac{z\alpha}{2}$ ，接受對立假設 $H_1$

# 隨機性檢定範例(1/3)

**範例19.10** 欲了解咖啡館販售卡布奇諾咖啡飲料數量(杯數)是否屬於隨機性，查詢過去兩週卡布奇諾咖啡飲料的販售數量依日期序如下所示，在顯著水準 $\alpha = 0.05$ 下利用隨機性檢定法，針對該咖啡館販售卡布奇諾咖啡飲料數量(杯數)進行隨機統計檢定。

15	24	31	24	14	26	37	28	19	45	7	9	23	18
----	----	----	----	----	----	----	----	----	----	---	---	----	----

題解：

- A. 設定顯著水準 $\alpha = 0.05$
- B. 虛無假設 $H_0$ : 每天販售卡布奇諾咖啡飲料數量屬於隨機分布
- C. 對立假設 $H_1$ : 每天販售卡布奇諾咖啡飲料數量不屬於隨機分布
- D. 計算檢定統計值-連數R

計算中位數，將前述樣本觀測值依據數值大小，從小排到大排列。樣本數量 $n = 14$ 屬於偶數，中位數為第 $\frac{n}{2}$ 和 $\frac{n}{2} + 1$ 個(順位)最大樣本數值(觀測值)的算術平均值。第 $\frac{n}{2} = \frac{14}{2} = 7$ 和 $\frac{n}{2} + 1 = \frac{14}{2} + 1 = 7 + 1 = 8$ 個(順位)，故中位數 $= \frac{23+24}{2} = 23.5$ 。

## 隨機性檢定範例(2/3)

**範例19.10** 欲了解咖啡館販售卡布奇諾咖啡飲料數量(杯數)是否屬於隨機性，查詢過去兩週卡布奇諾咖啡飲料的販售數量依日期序如下所示，在顯著水準 $\alpha = 0.05$ 下利用隨機性檢定法，針對該咖啡館販售卡布奇諾咖啡飲料數量(杯數)進行隨機統計檢定。

7	9	14	15	18	19	23	24	24	26	28	31	37	45
---	---	----	----	----	----	----	----	----	----	----	----	----	----

依據原來樣本的日期序位，將樣本觀測值與中位數進行比較，若樣本觀測值數值大於中位數則標示為H，樣本觀測值數值小於中位數則標示為L。當樣本觀測值數值等於中位數時，將該觀測值視為不計捨棄。

15	24	31	24	14	26	37	28	19	45	7	9	23	18
L	H	H	H	L	H	H	H	L	H	L	L	L	L
1	2		3	4		5	6	7					

連數 $R = 7$ 。L符號數量有7個  $= n_1 = 7$ ；H符號數量有7個  $= n_2 = 7$ 。

## 隨機性檢定範例(3/3)

**範例19.10** 欲了解咖啡館販售卡布奇諾咖啡飲料數量(杯數)是否屬於隨機性，查詢過去兩週卡布奇諾咖啡飲料的販售數量依日期序如下所示，在顯著水準 $\alpha = 0.05$ 下利用隨機性檢定法，針對該咖啡館販售卡布奇諾咖啡飲料數量(杯數)進行隨機統計檢定。

若 $r = 7 < \frac{2 \times n_1 \times n_2}{n_1 + n_2} + 1 = \frac{2 \times 7 \times 7}{7 + 7} + 1 = 7 + 1 = 8$ ，機率 $p = 2 \times P(R \leq r | n_1, n_2) = 2 \times P(R \leq 7 | 7, 7) = 2 \times 0.383$ (查連檢定累計機率表) = 0.766

**E.** 機率 $p = 0.766 >$  顯著水準 $\alpha = 0.05$ ，**接受**虛無假設 $H_0$ ：咖啡館每天販售卡布奇諾咖啡飲料數量屬於隨機分布

輕鬆練習一下

此投影片播放後3分鐘內，沒有進入該練習卷者，老師即關閉  
沒有進入者的練習卷(本練習5分鐘內完成)

**上課練習112ch16\_9**

# 16.7等級相關檢定

- 母數統計法中欲評量母體資料中變數 $X$ 和變數 $Y$ 兩個配對變數之間的相關程度
  - ◆ 例如：欲評量學生身高與體重兩個變數的相關程度，即使用Pearson相關係數 $\rho_{xy}$ 。
- 欲評量母體資料中兩個順序尺度(ordinal scale)變數的相關程度
  - ◆ 例如：欲評量學生國文的成績排名與英文的成績排名兩個配對變數的相關程度，即必須使用等級相關係數 $\rho_s$ 或Spearman等級相關係數(Spearman rank correlation coefficient)。
- 兩個隨機變數不屬於常態分布，不適合使用Pearson相關係數 $\rho_{xy}$ 評量兩個變數的相關程度，兩個隨機變數不屬於常態分布時，可以將原始數值型態為區間或比例尺度轉換為順序尺度的型態，再以等級相關係數 $\rho_s$ 評量兩個變數的相關程度。

# 等級相關係數

■ 利用**樣本資料**所得的等級相關係數，一般使用 $r_s$ 符號標示。

■ 
$$r_s = 1 - \frac{6 \times \sum_{i=1}^n d_i^2}{n \times (n^2 - 1)}$$

◆  $d_i$ ：兩個配對變數第 $i$ 項觀測值的序位等級之差

◆  $n$ ：配對樣本數量

# 等級相關係數數值範圍

- 等級相關係數 $r_s$ 的數值範圍與Pearson相關係數 $r_{xy}$ 的數值範圍相同，皆必須介於-1到+1之間。
- 等級相關係數 $r_s$ 的數值意涵亦與Pearson相關係數 $r_{xy}$ 相同。
  - ◆  $r_s = 1$ ，代表 $x_i$ 與 $y_i$ 兩個變數的等級順序完全相同。
  - ◆  $r_s = -1$ ，代表 $x_i$ 與 $y_i$ 兩個變數的等級順序完全相反。
  - ◆  $r_s = 0$ ，代表 $x_i$ 與 $y_i$ 兩個變數的等級順序完全沒有關係。
- 利用樣本觀測值即可估算樣本等級相關係數 $r_s$ ，運用樣本等級相關係數 $r_s$ 推論母體等級相關係數 $\rho_s$ ，以檢定母體等級相關係數 $\rho_s$ 是否等於0，即代表兩個順序尺度變數之間是否有關係存在。
  - ◆ 小樣本數量( $n < 30$ )
  - ◆ 大樣本數量( $n \geq 30$ )

# 小樣本數量

- 在母體等級相關係數  $\rho_s = 0$  的情況，樣本等級相關係數  $r_s$  屬於一對稱分布，其平均值為 0，變異數為  $\frac{1}{n-1}$ ，標示為
  - ◆ 樣本等級相關係數  $r_s \sim$  對稱分布  $(0, \frac{1}{n-1})$
- 在顯著水準  $\alpha$  與樣本數量  $n$  情況下，可利用 Spearman 等級相關臨界值表進行統計檢定，以了解母體等級相關係數  $\rho_s$  是否等於 0。

# 統計推論

## ■ 雙尾檢定

◆ 臨界值  $-r_{\frac{\alpha}{2}} \leq$  檢定統計值  $r_s \leq$  臨界值  $r_{\frac{\alpha}{2}}$ ，接受虛無假設  $H_0$

◆ 檢定統計值  $r_s < -r_{\frac{\alpha}{2}}$  或 檢定統計值  $r_s > r_{\frac{\alpha}{2}}$ ，接受對立假設  $H_1$

## ■ 左尾檢定

◆ 檢定統計值  $r_s \geq$  臨界值  $-r_{\alpha}$ ，接受虛無假設  $H_0$

◆ 檢定統計值  $r_s < -r_{\alpha}$ ，接受對立假設  $H_1$

## ■ 右尾檢定

◆ 檢定統計值  $r_s \leq$  臨界值  $r_{\alpha}$ ，接受虛無假設  $H_0$

◆ 檢定統計值  $r_s > r_{\alpha}$ ，接受對立假設  $H_1$

# 大樣本數量

■ 母體等級相關係數  $\rho_s = 0$  情況，樣本等級相關係數  $r_s$  屬於一常態分布，其平均值為 0，變異數為  $\frac{1}{n-1}$ ，標示為

◆ 樣本等級相關係數  $r_s \sim N(0, \frac{1}{n-1})$

■ 顯著水準  $\alpha$  與樣本數量  $n$  情況，可利用標準化 Z 值進行統計檢定，以了解母體等級相關係數  $\rho_s$  是否等於 0。

◆ 檢定統計值  $z = \frac{r_s - E(r_s)}{\sigma_{r_s}} = \frac{r_s - 0}{\sqrt{\frac{1}{n-1}}}$

# 統計推論

## ■ 雙尾檢定

◆ 臨界值  $-\frac{z_{\alpha}}{2} \leq$  檢定統計值  $z \leq$  臨界值  $\frac{z_{\alpha}}{2}$ ，接受虛無假設  $H_0$

◆ 檢定統計值  $z < -\frac{z_{\alpha}}{2}$  或 檢定統計值  $z > \frac{z_{\alpha}}{2}$ ，接受對立假設  $H_1$

## ■ 左尾檢定

◆ 檢定統計值  $z \geq$  臨界值  $-z_{\alpha}$ ，接受虛無假設  $H_0$

◆ 檢定統計值  $z < -z_{\alpha}$ ，接受對立假設  $H_1$

## ■ 右尾檢定

◆ 檢定統計值  $z \leq$  臨界值  $z_{\alpha}$ ，接受虛無假設  $H_0$

◆ 檢定統計值  $z > z_{\alpha}$ ，接受對立假設  $H_1$

# 等級相關檢定範例(1/3)

**範例19.14** 欲了解大學生國文成績排名與英文成績排名是否有關係存在，隨機抽取12位大二學生，分別蒐集其大一國文成績與英文成績，如下表所示，在顯著水準 $\alpha = 0.05$ 下，利用Spearman等級相關檢定法，分析國文成績排名與英文成績排名是否有關係存在。

學生	1	2	3	4	5	6	7	8	9	10	11	12
國文成績	85	45	88	81	80	90	76	81	60	51	77	82
英文成績	62	78	45	65	78	81	19	65	75	79	67	69

題解：

- A. 設定顯著水準 $\alpha = 0.05$
- B. 虛無假設 $H_0: \rho_s = 0$ (國文成績排名與英文成績排名沒有關係存在)
- C. 對立假設 $H_1: \rho_s \neq 0$ (國文成績排名與英文成績排名有關係存在)
- D. 計算檢定統計值—樣本等級相關係數 $r_s$

# 等級相關檢定範例(2/3)

**範例19.14** 欲了解大學生國文成績排名與英文成績排名是否有關係存在，隨機抽取12位大二學生，分別蒐集其大一國文成績與英文成績，如下表所示，在顯著水準 $\alpha = 0.05$ 下，利用Spearman等級相關檢定法，分析國文成績排名與英文成績排名是否有關係存在。

學生	1	2	3	4	5	6	7	8	9	10	11	12	合計
國文成績 $x_i$	85	45	88	81	80	90	76	81	60	51	77	82	
國文排名 $R_{x_i}$	3	12	2	5.5	7	1	9	5.5	10	11	8	4	
英文成績 $y_i$	62	78	45	65	78	81	19	65	75	79	67	69	
英文排名 $R_{y_i}$	10	3.5	11	8.5	3.5	1	12	8.5	5	2	7	6	
$d_i = R_{x_i} - R_{y_i}$	-7	8.5	-9	-3	3.5	0	-3	-3	5	9	1	-2	
$d_i^2$	49	72.25	81	9	12.25	0	9	9	25	81	1	4	352.5

樣本數量 $n = 12$

$$\text{檢定統計值 } r_s = 1 - \frac{6 \times \sum_{i=1}^n d_i^2}{n \times (n^2 - 1)} = 1 - \frac{6 \times 352.5}{12 \times (12^2 - 1)} = 1 - \frac{2115}{1716} = 1 - 1.2325 = -0.2325$$

## 等級相關檢定範例(3/3)

**範例19.14** 欲了解大學生國文成績排名與英文成績排名是否有關係存在，隨機抽取12位大二學生，分別蒐集其大一國文成績與英文成績，如下表所示，在顯著水準 $\alpha = 0.05$ 下，利用Spearman等級相關檢定法，分析國文成績排名與英文成績排名是否有關係存在。

E. 臨界值 $-r_{\frac{\alpha}{2}} = -r_{\frac{0.05}{2}} = -0.591 \leq$  檢定統計值 $r_s = -0.2325 \leq$  臨界值 $r_{\frac{\alpha}{2}} = r_{\frac{0.05}{2}} = 0.591$ (查詢Spearman等級相關臨界值表)，**接受**虛無假設 $H_0: \rho_s = 0$ (國文成績排名與英文成績排名沒有關係存在)

輕鬆練習一下

此投影片播放後3分鐘內，沒有進入該練習卷者，老師即關閉  
沒有進入者的練習卷(本練習5分鐘內完成)

**上課練習112ch16\_10**

先使用『來賓練習』熟習課程內涵

正式平常考每一題留下運算檔案

匡列足夠時間，一次登入，一口氣完成，一次繳交

關閉時間：D+3日2359

**平常考112ch16**

您辛苦了

第十六章課程結束