淺談交叉分析與卡方檢定

長榮大學會計資訊學系◆翁耀臨 副教授

依照普通高級中學選修科目「選修數學（I）」課程暫行綱要 95 暫綱內容規定，第一章第三小節「交叉分析」需介紹定性變數之列聯表以及學會判斷兩定性變數是否有關聯等觀念。由於僅止於概念性的介紹，本小節在編寫上的難度頗高。尤其在如何確認兩定性變數的關聯性方面，翰林版編者利用條件機率的性質，一一檢驗兩定性變數是否為獨立事件作為判斷它們是否有關聯的依據。然而課文中出現了兩次「到底要接近到什麼程度才算有關聯？」之敘述，其實本節所根據的理論基礎乃統計學裡的「卡方檢定」。本文將從重要的統計推論——「假設檢定」出發，進而介紹出卡方檢定的基本概念，最後將課本內文的例題利用卡方檢定統計量判斷其關聯性，提供讀者對於「交叉分析」有更深一層的認識。

一 假設檢定

㈠理論基礎

統計學可區分為敘述統計與推論統計，而敘述統計已於翰林版高中數學第四冊第三章介紹過。至於推論統計部分，依據其推論的目的可區分為「估計」與「假設檢定」兩大主題，這也是統計學非常重要的理論基礎。由於往往欲分析的母體其原始資料筆數過於龐大，以致於無法以普查的方式取得母體參數，因此需透過抽樣的方法利用樣本統計量推估母體參數。這就建構出推論統計的概念。

所謂母體參數即為描述母體的特性之測量值。常見的有母體平均數、母體變異數，以及母體比例等。在翰林版高中數學第四冊第三章第六小節已介紹如何利用樣本統計量進行「區間估計」來推估母體參數，也就是讓我們了解有多少可能母體參數落在樣本統計量所建構的一個特定區間。翰林版高中「選修數學（I）」第一章第三小節將繼續探討推論統計的另一主題——「假設檢定」，亦即如何利用樣本統計量來判定是否拒絕研究者對母體參數假設之方法。舉例來說，日常生活中有可能有人會對某些產品的內容物是否有標示不實的情況產生質疑，也有研究單位想調查消費者對某公司提供的服務是否感到滿意等等。諸如這類例子，在統計學上提供了一項很有效的科學方法來解釋，也就是「假設檢定」。

所謂「假設檢定」（Hypothesis Testing），是指研究者先對母體參數做出一適當的暫時性假設，然後根據隨機抽樣的樣本，利用樣本統計量之抽樣分配（註①）決定是否支持該假設的過程。例如：消費者懷疑某品牌鮮奶平均容量不足一公升，則研究者以該品牌鮮奶平均容量大於或等於一公升作為暫時性假設。建立假設後，研究者再進一步蒐集樣本資料，如果證據越充分，推翻該暫時性假設的可能性就越高。在資料蒐集後，研究者必須根據樣本統計量的抽樣分配訂定一個拒絕假設的標準。如果樣本資料之檢定統計量落在拒絕的範圍，則研究者可以推翻原先建立之

暫時性假設，否則必須接受暫時性假設。假設檢定的主要精神在於：除非有足夠的證據可以拒絕暫時性假設，否則必須接受暫時性假設為真的事實。

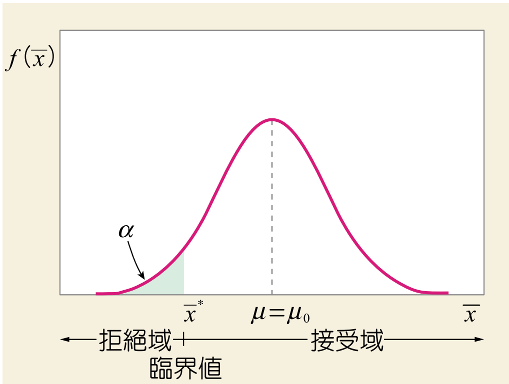
上述的暫時性假設稱為虛無假設（null hypothesis），通常以 H_0 表示，是指研究者欲推翻的統計假設。以上述例子為例，可建立虛無假設 H_0 為「鮮奶平均容量大於或等於一公升」，而虛無假設的反面敘述稱為對立假設（alternative hypothesis），通常以 H_1 表示。以上述例子為例，可建立虛無假設 H_1 為「鮮奶平均容量小於一公升」。

㈡檢定程序

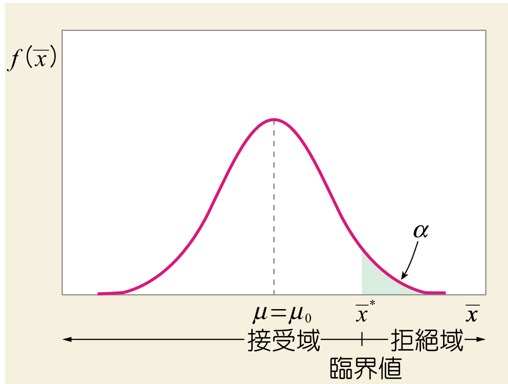
接下來我們將介紹如何利用樣本資料來決定拒絕或者接受虛無假設。以一組樣本統計量來估計母體參數時，樣本統計量的估計容許一定的誤差，這個誤差取決於樣本大小與研究者給定之信賴水準（註②）。研究者要求的程度不同，容許造成決策錯誤的機率也隨之不同。這類型犯錯的機率指的是下列兩種可能：一是當虛無假設 H_0 為真而拒絕 H_0，這種錯誤稱為型 I 錯誤（Type I Error），通常以 α 表示。二是當虛無假設 H_0 為偽而接受 H_0，這種錯誤稱為型 II 錯誤（Type II Error），通常以 β 表示。而容許型 I 錯誤之最大機率 α 值稱為顯著水準（level of significance），也就是若事實上虛無假設 H_0 是正確的，研究者拒絕 H_0 願意冒的最大風險。

以上述鮮奶平均容量為例，H_0：$\mu\geq 1$，H_1：$\mu<1$，其中 μ 表示鮮奶平均容量。則 α 以條件機率可表示為 $\alpha=P$（拒絕 H_0｜H_0 為真）$=P$（$\mu<1$｜$\mu\geq 1$）。而 β 以條件機率可表示為 $\beta=P$（接受 H_0｜H_0 為偽）$=P$（$\mu\geq 1$｜$\mu<1$）。這兩種類型的錯誤又以型 I 錯誤較為嚴重。在上例中如果經檢定後造成型 I 錯誤，亦即該公司的鮮奶平均容量事實上均有一公升以上，但是檢定的結果卻指控該公司的鮮奶平均容量不到一公升，有標示不實之嫌，導致影響該公司的信譽以及造成銷售量下降的嚴重損失。為求得較準確的檢定結果，理應同時降低所有類型的錯誤。不過在統計學的理論中，樣本數不變的條件下，無法同時降低 α 與 β 值。因此在僅可選擇型 I 或型 II 錯誤的條件下，統計之假設檢定選擇避免造成型 I 錯誤為主要考量，畢竟型 I 錯誤造成的影響遠比型 II 錯誤來得大。一般而言，在假設檢定中通常事先不控制 β 值，而先決定容許犯型 I 錯誤的最大機率 α 值，也就是所謂的顯著水準。一般均將 α 設得很小，常見的顯著水準為 0.01, 0.05 以及 0.1。

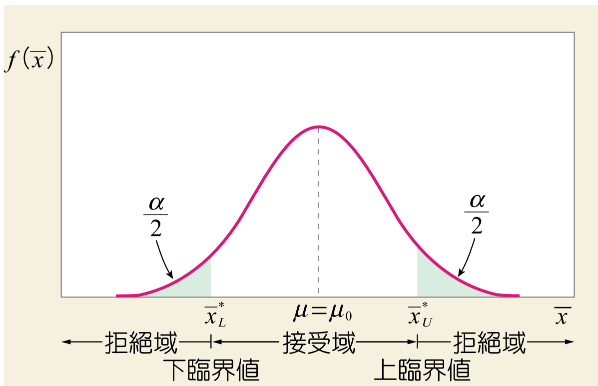
研究者先給定顯著水準 α 值後，可視其要求程度決定一個拒絕虛無假設的範圍，這個範圍稱為拒絕域（rejection region），拒絕域的端點稱為臨界值（critical value），臨界值的選擇完全取決於顯著水準 α 值。研究者決定拒絕域後，再依所選取的樣本計算其檢定統計量（test statistic）（註③），並判斷該檢定統計量是否落在拒絕域中。如果落在拒絕域，則拒絕虛無假設 H_0，否則只好接受 H_0。此即為假設檢定的判定過程。



(1)左尾檢定



(2)右尾檢定

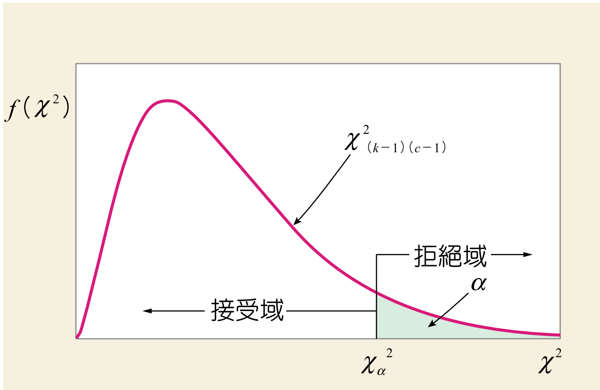


(3)雙尾檢定

圖一　單一母體平均數三種假設檢定形式的拒絕域與接受域

二 卡方檢定

卡方檢定主要是用於類別資料的資料分析，例如：不同品牌產品之市場佔有率是否有顯著差異、性別與產品喜好程度之關聯性、教育程度與薪資待遇之關聯性等。其檢定之主要方法是比較樣本資料之「觀察次數」（Observed Frequency）與當虛無假設為真的條件下之「期望次數」（Expected Frequency）的接近程度，然後依據卡方分配（註④）之假設檢定來判定接受或拒絕虛無假設。當「觀察次數」與「期望次數」之差異越大，檢定統計量 χ^2 值（註⑤）落在拒絕域的機率越高，越有可能拒絕虛無假設。



圖二　卡方檢定的拒絕域與接受域

卡方檢定主要可用來作以下三種檢定：(1)適合度檢定：檢定母體是否為某一特定分配之檢定方法。(2)獨立性檢定：檢定兩定性變數是否相互獨立之檢定方法。(3)齊一性檢定：檢定不同母體是否具有相同的分配或相同的比例之檢定方法。翰林版高中「選修數學（I）」第一章第三小節「交叉分析」所探討的內容即屬「獨立性檢定」。以下步驟是獨立性檢定的過程：

(1)將觀察到的資料表示為列聯表。

(2)依據假設檢定建立的基本原則建立虛無假設與對立假設

H_0：兩個定性變數無關聯

H_1：兩個定性變數有關聯

並計算第一類定性變數第 i 個等級與第二類定性變數第 j 個等級之期望次數：

$$E_{ij}=n\times P(r_i\cap c_j)=n\times P(r_i)\times P(c_j)$$

$$=\frac{(\text{第 } i \text{ 列樣本總數})\times(\text{第 } j \text{ 列樣本總數})}{\text{樣本總數}},$$

其中 $P(r_i)$ 表示列聯表中第 i 列（以 r 表示第一類定性變數，則 r_i 也就是第一類定性變數第 i 個等級）發生的機率，而 $P(c_j)$ 表示列聯表中第 j 行（也就是第二類定性變數第 j 個等級）發生的機率。至於 $P(r_i\cap c_j)=P(r_i)\times P(c_j)$，原因是 H_0 為真的情形下，r_i 和 c_j 是相互獨立的。

(3)計算卡方檢定的檢定統計量，其定義為：$\chi^2=\sum_i\sum_j\frac{(O_{ij}-E_{ij})^2}{E_{ij}}$，其中 O_{ij} 代表樣本資料之「觀察次數」，而 E_{ij} 代表樣本資料之「期望次數」。

(4)依據獨立性檢定的理論基礎：在 H_0 為真的情形下，$\chi^2=\sum_i\sum_j\frac{(O_{ij}-E_{ij})^2}{E_{ij}}$ 服從自由度為 $(r-1)(c-1)$ 的卡方分配（註⑥），其中每一個定性變數的期望次數皆須大於或等於 5。

(5)獨立性檢定決策法則：若 χ^2 大於 χ_α^2，則拒絕 H_0，其中 χ_α^2 為顯著水準 α 的拒絕域之臨界值。

三 課本例題說明

以下就翰林版高中「選修數學（I）」第一章第三小節之課文例題利用卡方檢定判斷兩定性變數之關聯性。假設給定的顯著水準都是 0.05。

(1) P46 表五　應考人員“性別”與“錄取與否”的列聯表

	錄取	未被錄取	合計
男	175	200	375
女	150	475	625
合計	325	675	1000

勞工局如何分析這表格的數據，以判定這次招考是否有性別歧視？

Step 1. 上述列聯表所代表的就是觀察次數 O_{ij}，亦即

$O_{11}=175$，$O_{12}=200$

$O_{21}=150$，$O_{22}=475$

Step 2. 建立假設。在本例題中

H_0：“性別”與“錄取與否”無關聯

H_1：“性別”與“錄取與否”有關聯

Step 3. 計算期望次數 E_{ij}

$$E_{11}=\frac{375\times325}{1000}=121.875，E_{12}=\frac{375\times675}{1000}=253.125$$

$$E_{21}=\frac{625\times325}{1000}=203.125，E_{22}=\frac{625\times675}{1000}=421.875$$

Step 4. 計算卡方檢定統計量

$$\chi^2=\frac{(175-121.875)^2}{121.875}+\frac{(200-253.125)^2}{253.125}+\frac{(150-203.125)^2}{203.125}+\frac{(475-421.875)^2}{421.875}$$

≈54.89

Step 5. 檢定結果

經查表得知自由度為（2－1）×（2－1）＝1 的 χ^2 值 ${\chi_{0.05}}^2=3.841$。因為 $54.89>3.841$，得到的結論是拒絕 H_0，也就是在顯著水準 $\alpha=0.05$ 的情形下，“性別”與“錄取與否”是有關聯的。

(2) P47 表八　應考人員“性別”與“錄取與否”的列聯表

	錄取	未被錄取	合計
男	25	50	75
女	75	150	225
合計	100	200	300

Step 1. 上述列聯表所代表的就是觀察次數 O_{ij}，亦即

$O_{11}=25$，$O_{12}=50$

$O_{21}=75$，$O_{22}=150$

Step 2. 建立假設。在本例題中

H_0：“及格與否”與“性別”無關聯

H_1：“及格與否”與“性別”有關聯

Step 3. 計算期望次數 E_{ij}

$$E_{11}=\frac{75\times100}{300}=25，E_{12}=\frac{75\times200}{300}=50$$

$$E_{21}=\frac{225\times100}{300}=75，E_{22}=\frac{225\times200}{300}=150$$

Step 4. 計算卡方檢定統計量

由於在本例中，對於所有的 i, j, $O_{ij}=E_{ij}$，因此 $\chi^2=0$

Step 5. 檢定結果

經查表得知自由度為（2－1）×（2－1）＝1 的 χ^2 值 $\chi_{0.05}{}^2=3.841$

因為 $0<3.841$，得到的結論是不拒絕 H_0，也就是在顯著水準 $\alpha=0.05$ 的情形下，“性別”與“錄取與否”是無關聯的。

(3) P48 例題 2　試分析“及格與否”與“性別”是否有關聯

	及格	不及格	合計
男	125	55	180
女	82	38	120
合計	207	93	300

Step 1. 上述列聯表所代表的就是觀察次數 O_{ij}，亦即

$O_{11}=125$，$O_{12}=55$

$O_{21}=82$，$O_{22}=38$

Step 2. 建立假設。在本例題中

H_0：“及格與否”與“性別”無關聯

H_1：“及格與否”與“性別”有關聯

Step 3. 計算期望次數 E_{ij}

$$E_{11}=\frac{180\times207}{300}=124.2，E_{12}=\frac{180\times93}{300}=55.8$$

$$E_{21}=\frac{120\times207}{1000}=82.8，E_{22}=\frac{120\times93}{300}=37.2$$

Step 4. 計算卡方檢定統計量

$$\chi^2=\frac{(125-124.2)^2}{124.2}+\frac{(55-55.8)^2}{55.8}+\frac{(82-82.8)^2}{82.8}+\frac{(38-37.2)^2}{37.2}\approx0.035$$

Step 5. 檢定結果

經查表得知自由度為（2－1）×（2－1）＝1 的 χ^2 值 $\chi_{0.05}{}^2=3.841$

因為 $0.035<3.841$，得到的結論是不拒絕 H_0，也就是在顯著水準 $\alpha=0.05$ 的情形下，

　“性別”與“錄取與否”是無關聯的。

(4) P49 例題 3　試分析“抽到 A”與“花色是紅”的關聯性

	紅	黑	合計
A	2	2	4
不是 A	24	24	48
合計	26	26	52

Step 1. 上述列聯表所代表的就是觀察次數 O_{ij}, 亦即

$O_{11}=2$，$O_{12}=2$

$O_{21}=24$，$O_{22}=24$

Step 2. 建立假設。在本例題中

H_0：“及格與否”與“性別”無關聯

H_1：“及格與否”與“性別”有關聯

Step 3. 計算期望次數 E_{ij}

$$E_{11}=\frac{4\times26}{52}=2，E_{12}=\frac{4\times26}{52}=2$$

$$E_{21}=\frac{48\times26}{52}=24，E_{22}=\frac{48\times26}{52}=24$$

Step 4. 計算卡方檢定統計量

由於在本例中，對於所有的 i, j, $O_{ij}=E_{ij}$，因此 $\chi^2=0$

Step 5. 檢定結果

經查表得知自由度為（2－1）×（2－1）＝1 的 χ^2 值 ${\chi_{0.05}}^2=3.841$

因為 $0<3.841$，得到的結論是不拒絕 H_0，也就是在顯著水準 $\alpha=0.05$ 的情形下，
“抽到 A”與“花色是紅”是無關聯的。

(5) P50 表九　應考人員“性別”與“加薪與否”的列聯表

	加薪	不加薪	合計
男	75	225	300
女	50	150	200
合計	125	375	500

Step 1. 上述列聯表所代表的就是觀察次數 O_{ij}，亦即

$O_{11}=75$，$O_{12}=225$

$O_{21}=50$，$O_{22}=150$

Step 2. 建立假設。在本例題中

H_0：“性別”與“加薪與否”無關聯

H_1：“性別”與“加薪與否”有關聯

Step 3. 計算期望次數 E_{ij}

$$E_{11}=\frac{300\times 125}{500}=75，E_{12}=\frac{300\times 375}{500}=225$$

$$E_{21}=\frac{200\times 125}{500}=50，E_{22}=\frac{200\times 375}{500}=150$$

Step 4. 計算卡方檢定統計量

由於在本例中，對於所有的 i, j, $O_{ij}=E_{ij}$，因此 $\chi^2=0$

Step 5. 檢定結果

經查表得知自由度為（2－1）×（2－1）＝1 的 χ^2 值 $\chi_{0.05}{}^2=3.841$

因為 $0<3.841$，得到的結論是不拒絕 H_0，也就是在顯著水準 $\alpha=0.05$ 的情形下，"性別" 與 "加薪與否" 是無關聯的

(6) P52 例題 4　試分析 "發生事故" 與 "是否有手機" 的關聯性

	事故	無事故	合計
有手機	12	288	300
無手機	20	480	500
合　計	32	768	800

Step 1. 上述列聯表所代表的就是觀察次數 O_{ij}，亦即

$O_{11}=12$，$O_{12}=288$

$O_{21}=20$，$O_{22}=480$

Step 2. 建立假設。在本例題中

H_0："發生事故" 與 "是否有手機" 無關聯

H_1："發生事故" 與 "是否有手機" 有關聯

Step 3. 計算期望次數 E_{ij}

$$E_{11}=\frac{300\times 32}{800}=12，E_{12}=\frac{300\times 768}{800}=288$$

$$E_{21}=\frac{500\times 32}{800}=20，E_{22}=\frac{500\times 768}{800}=480$$

Step 4. 計算卡方檢定統計量

由於在本例中，對於所有的 i, j, $O_{ij}=E_{ij}$，因此 $\chi^2=0$

Step 5. 檢定結果

經查表得知自由度為（2－1）×（2－1）＝1 的 χ^2 值 $\chi_{0.05}{}^2=3.841$

因為 $0<3.841$，得到的結論是不拒絕 H_0，也就是在顯著水準 $\alpha=0.05$ 的情形下，"發生事故" 與 "是否有手機" 是無關聯的。

四 結論

翰林版高中「選修數學（I）」第一章第三小節課本內文提及：「在兩個定性變數的列聯表中，固定一個定性變數的值，讓另一個定性變數變動，若所得到的比率都接近於一個固定比率時，則稱兩種定性變數無關聯。」在第(2)、(4)、(5)、(6)例子的列聯表，我們可以很清楚的看出這一個固定比率。透過本文的介紹，讀者可以利用卡方檢定的理論了解課本內文的這段敘述，因為這一個固定比率導致 $O_{ij}=E_{ij}$，而得出 $\chi^2=0$ 的結果，因而作出不拒絕 H_0 的結論（註⑦）。

另一方面在第(3)例子中，雖然無法確定出該固定比率，但是及格的男女比例為 125：82 近似 1.52：1，而不及格的男女比例為 55：38 近似 1.45：1，兩比例相當接近，因此造成 χ^2 值不大，也得到不拒絕 H_0 的結論。

最後在第(1)例子中，錄取的男女比例為 175：150 近似 1.17：1，而未錄取的男女比例為 200：475 近似 0.42：1，兩比例差距甚遠，因此造成 χ^2 值變大，得到拒絕 H_0 的結論。

綜上所述可以回應課本內文的另一段敘述：「實際的資料不像上面這麼精準，只要近似就可以了。至於到底要差異到什麼程度才算有關聯？同學可以在大學的統計課學到。」透過本文的介紹，讀者可以清楚的清楚的看出：

(1)當樣本資料的觀察次數（O_{ij}）與期望次數（E_{ij}）相當接近時，會導致 χ^2 值趨近於 0，得出兩定性變數無關聯的結論。

(2)而當樣本資料的觀察次數（O_{ij}）與期望次數（E_{ij}）之間的差異太大以致於 χ^2 值超過拒絕域之臨界值 χ_α^2，得出兩定性變數有關聯的結論。

註：

①抽樣分配係指樣本統計量之機率分配。

②該名詞已於翰林版高中數學第四冊第三章介紹過。

③檢定統計量乃一用來量測介於虛無假設和樣本資料之間適合性之數值。

④卡方分配（*Chi- Square Distribution*）是指檢定統計量 χ^2 之抽樣分配。

⑤又稱卡方值。其一般定義為 $\chi^2=\sum_i \frac{(O_i-E_i)^2}{E_i}$，其中 O_i 代表樣本資料之「觀察次數」，而 E_i 代表樣本資料之「期望次數」。

⑥自由度（*degree of freedom*）係指檢定統計量中可以自由變動的變數個數。例如在 r 列 c 行的 $r\times c$ 列聯表中，每一橫列或每一縱行的最後一筆次數資料受限於橫列縱行的總合次數，故無法自由決定，因此失去 $r+c-1$ 個自由度。故自由度為 $rc-(r+c-1)=(r-1)(c-1)$。

⑦依據卡方檢定統計量 χ^2 的定義，χ^2 為一個大於或等於 0 的數，因此 $\chi^2=0$的卡方值一定不會落在拒絕域。

參考資料：

1. *Statistics for Business and Economics* （*A Practical Approach*）, by D.R.Anderson, D.J.Sweeney, T.A. Williams, and J–C.Chen.
2. *Statistics for management and Economics*, by Keller & Warrack.
3. *The Practice of Business Statistics Using Data for Decisions*, by D. S. Moore, G. P. McCabe, W. M. Duckworth, and L. C. Alwan.
4. 統計學導論，方世榮著。
5. 應用統計學，林惠玲、陳正倉合著。

第三章 變異數分析

變異數分析（ANOVA）是試驗設計資料的基本統計分析工具，是以 F-test 來檢測差異的顯著性，是依據二獨立的樣本變異數之比呈現一 F-分佈，將二不同情況下的變異數代入分子及分母，會得到檢測的 F-值，及相對的 p-值。

F-分佈

設有二常態母體，變異數分別為σ_1^2及σ_2^2，
二隨機樣本，樣本變異數 S_1^2，S_2^2，自由度為 v_1及 v_2，令

$$F = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2},$$

則 F 為一 F-分佈，自由度為 $(v_1，v_2)$，F-分佈的形狀由分子及分母的自由度決定。

一因子變異數分析 (1-way ANOVA)

<u>目標</u> : 比較 m 組均值 (母體均值 $\mu_1, \ldots, \mu_i, \ldots, \mu_m$)

<u>前提</u>：各組資料獨立且同質，誤差獨立，誤差服從常態分佈。

<u>Data</u> : *m* 組資料，每組 n 個觀測值。

y_{ij} --- 第 i 組資料之第 j 組觀測值

Group	1		*i*		*m*
	y_{11}	……	y_{i1}	……	y_{m1}
	y_{12}		y_{i2}		y_{m2}
	…		…		…
	y_{1n}		y_{in}		y_{mn}
Avg.					

【模式】 $y_{ij} = \mu_i + \varepsilon_{ij}$, $i= 1,\ldots, m$, $j= 1,\ldots, n$

ε_{ij} 試驗誤差 , 假設分佈為 $N(0, \sigma^2)$

註：此時假設 常態性，及同質性

【假說】H_0：$\mu_1= \mu_2= \ldots= \mu_m$

H_1：差異存在 (至少一對 i，j， $\mu_i\neq \mu_j$)

【平方和分解 】 $\boxed{SST = SS_t+ SSE}$

說明：整體變化可分解為組間變化及組內變化

【變異數，或稱為均方】

$MS_t = SS_t/(m-1)$ 量測各組間之變化，當 H_0 為真時，

期望值 = σ^2

$MSE = SSE /(N -m)$ 量測組內之變化 (假設各組同質)，

MSE 估計組間共同變異數， σ^2。

Q：欲檢定各組差異時，用何量作檢定？ 檢定量的分佈是什麼？ 在什麼情況下差異顯著 ？

【檢定量】 F = MS_t / MS_E

說明：1.在 H_0 為真之情況下，F 呈現一F分布，自由度是（m-1, N-m)。

2. 若差異存在，MS_t 值偏大，將產生一較大的 F 值。

故在 F 值大時，傾向於差異顯著。可視為一單尾檢定。

【檢定法】p-value = P(F-值 > F)，

當 p-value<α ，組間差異顯著，或 treatment 之效力顯著。

p-value 愈小，差異愈顯著。

說明：1、treatment 指試驗或取樣各組的處理方法。

2、因為是以二變異數之比為檢定標準，故稱為變異數分析。

變異數分析可彙編成一 ANOVA Table 如下：

ANOVA table for 1-way model					
Source	SS	df	MS	F-值	p-值
Treatment	SSt	m-1	MSt	F =MSt/MSE	
Error	SSE	N-m	MSE		
Total	SST	N-1			

說明：1、F-值愈大，p-值愈小，差異愈顯著， F-值的判斷標準是隨著 df 而變，而 p-值是一機率值，與 df 無關，所以用 p-值 做為是否有顯差比較理想，通常以 p-值與 α-值比較，p-值 < α，則各組之間有顯著差異。

2、此處 p-值是m組平均值相等的可能機率

【例 3.1 】比較A、B、C三品種蛋白質含量之差異 (p42)

	1	2	3	4	¥ §
A	7	8	5	4	6
B	9	8	6	5	7
C	10	13	11	10	11

此試驗的處理是品種，有三個水準(level)，每組有4觀察值

方法 : 1-way ANOVA

ANOVA table					
Source	SS	df	MS	F-值	p-值
品種	56	2	28	9.69	0.00569
誤差	26	9	2.9		
整體	82	11			

p-value = .00569 < 0.05，結論此3組蛋白質均值之差異顯著。

事實上，變方分析是二組獨立資料比較的延伸，可由下列特例作解釋

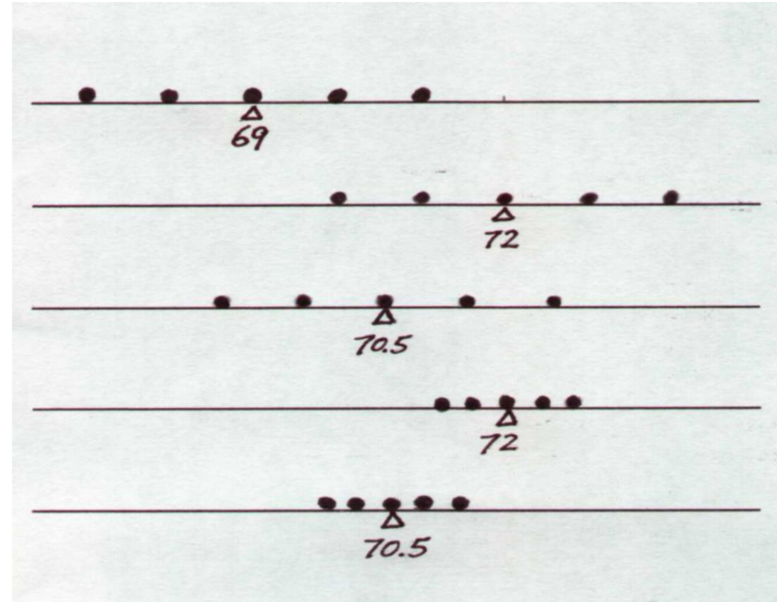
資料A: *67*，*68*，*69*，*70*，*71*

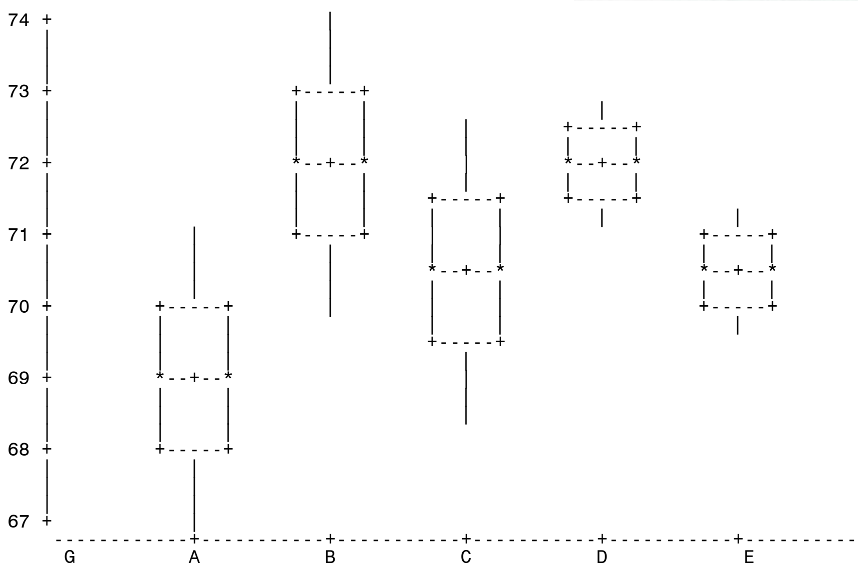
資料B: *70*，*71*，*72*，*73*，*74*

資料C: *68.5*，*69.5*，*70.5*，*71.5*，*72.5*

資料D: *71.2*，*71.6*，*72*，*72.4*，*72.8*

資料E: 69.7，*70.1*，*70.5*，*70.9*，*71.3*





檢定：A vs. B:　p-value = 0.017, 差異顯著

B vs. C:　p-value = 0.172, 差異不顯著

D vs. E:　p-value = 0.0056, 差異顯著

<u>**觀察**</u>：

A、B 二組平均數差異大，導致 A、B 二組差異顯著；D、E 二組變異數小，導致 D、E 二組差異顯著。由此得到各組資料是否差異顯著，不僅決定於組間均值的差異大小，也要看組內資料是否集中。所以ANOVA中的 F-值 = (組間變異)/(組內變異).

對二組比較時，ANOVA中的 F-值 實際上是 t-檢定 的 t-值之平方。

$$t = \frac{\overline{X}-\overline{Y}}{\sqrt{\hat{\sigma}^2}} = \frac{\text{平均數之差異}}{\text{標準誤}}, \quad t^2 = \frac{(\text{平均數之差異})^2}{\hat{\sigma}^2}$$

比較A、B、C三組之差別

【資料1 】

	1	2	3	4	平均
A	7	3	10	4	6
B	4	10	6	8	7
C	10	14	9	11	11

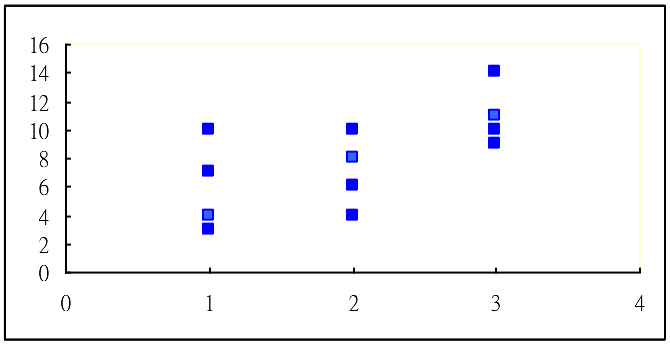
	A	B	C
mean ± SD	6±3.16	7±2.58	11±.2.22

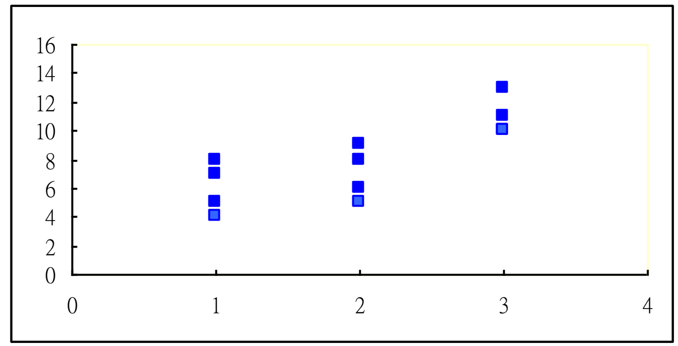
【資料2 】

	1	2	3	4	¥ §
A	7	8	5	4	6
B	9	8	6	5	7
C	10	13	11	10	11

	A	B	C
mean ± SD	6±1.83	7±1.83	11±.1.41

資料1 與資料2： 資料圖比較，三組平均值皆為 6, 7, 11





F-值 =3.49， p-value = .076 ，結論此3組均值之差異不顯著 (mse=7.19)

F-值 =9.69， p-value = .0057 ，結論此3組均值之差異顯著 (mse=2.89)

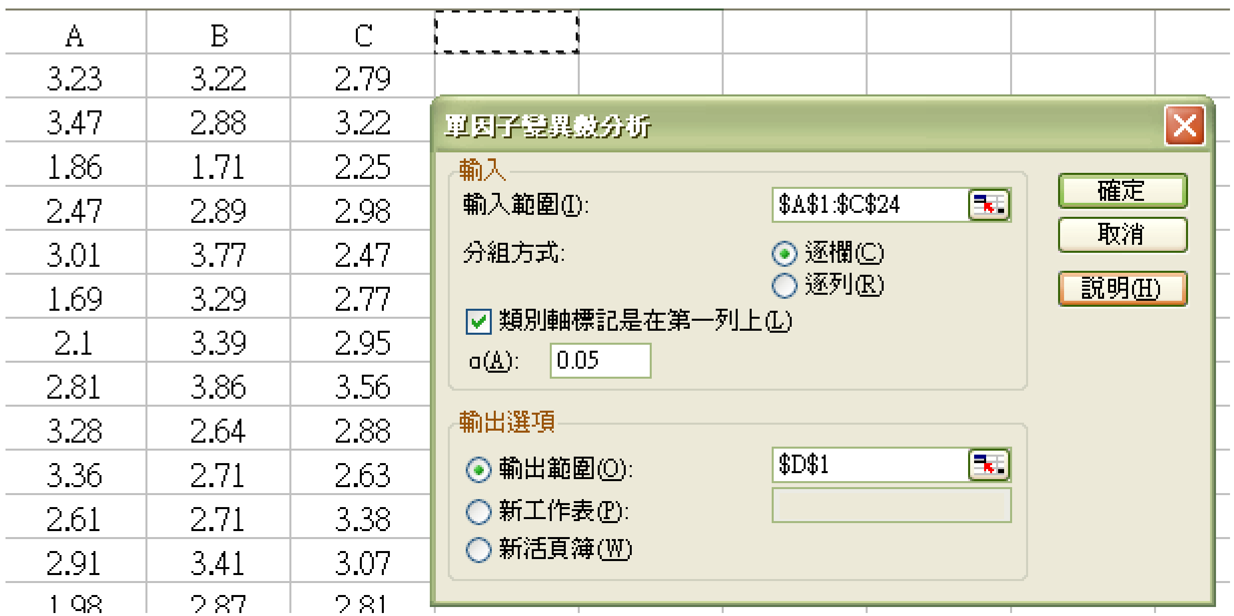
EXCEL

EXCEL:　輸入資料→

工具 →資料分析 → 單因子變異數分析

(EXCEL 只作變異數分析，不提供進一步比較，用於預檢資料)

例：



A	B	C
3.23	3.22	2.79
3.47	2.88	3.22
1.86	1.71	2.25
2.47	2.89	2.98
3.01	3.77	2.47
1.69	3.29	2.77
2.1	3.39	2.95
2.81	3.86	3.56
3.28	2.64	2.88
3.36	2.71	2.63
2.61	2.71	3.38
2.91	3.41	3.07
1.98	2.87	2.81

第四章　均值比較測驗

- 進行 ANOVA 得到差異顯著的結果，表示至少有一對組的均值不等，有必要作進一步分析確切不等的狀況。
- 分析的方法因著 Treatment 是 quantitative (數量型的) 或 qualitative (特性型的) 而不同。

- 特性型的 treatment 一般選擇

  (I) 對對比較 (pairwise comparison)； 或，

  (II) 對比量比較 (constrast test) ；
- 數量型的 treatment 一般選擇

  (III) 相關性檢定，或尋求關係式 (迴歸分析)。

論文範例

Breast Cancer Risk in Rats Fed a Diet High in n-6 Polyunsaturated Fatty Acids During Pregnancy

Statistical Analyses

Statistical tests were performed by use of the SOLO software (BMDP Statistical Software, Los Angeles, CA). One- or two-way analysis of variance (ANOVA) (*36*) was used to analyze results for the body weight, food intake, and other parameters associated with pregnancy, hormonal data, and data for tumor size and latency. Where appropriate, between-group comparisons were performed using Fisher's least significant difference test. The logrank test was used to analyze tumor incidence for the two groups. In addition, the chi-squared test was used to analyze the difference in tumor incidence at the last week of tumor measurements (week 18). All statistical tests were two-sided.

I、對對比較 (pairwise comparison)

--- 同時對所有可能之對組作差異性的檢定，虛無假說如下，方法有許多，可任擇 2 ~ 3 法進行。

Test $H_0 : \mu_i = \mu_j$ for all $i \neq j$

a. Fisher’s Least significant Difference (LSD) method (例4.1)

$$\mathrm{LSD} = t_{\alpha/2;N-m}\sqrt{\mathrm{MS_E}\left(\frac{1}{n_i}+\frac{1}{n_j}\right)}$$

分界值

當 $\left|\bar{y}_{i.} - \bar{y}_{j.}\right| > \mathrm{LSD}$ 時，

判斷 μ_i 與 μ_j 的差異顯著

本法特點：計算簡單，易於理解，在差異顯著時，犯錯的機率偏大；當實驗的組數較多時，不準性愈高。

b. Duncan's Multiple range Test

分界值與序位差有關

$$D_p = r_{\alpha;(p,f)}\sqrt{MS_E / n_h}, \qquad n_h = \frac{m}{\Sigma(1/n_i)} \quad \text{(harmonic mean)}$$

當 $\left|\bar{y}_{i.} - \bar{y}_{j.}\right| > D_p$ 時，where the order difference of $\bar{y}_{i.}$ and $\bar{y}_{j.}$ = p-1

判斷 μ_i 與 μ_j 的差異顯著

特點：考慮均值的排序，對差異顯著要求較高

c. Tukey's HSD Test

分界值

$$T_\alpha = q_{\alpha;(m,f)} S_{\bar{y}_{i.}}$$

當 $\left|\bar{y}_{i.} - \bar{y}_{j.}\right| > T_\alpha$ 時， 判斷 μ_i 與 μ_j 的差異顯著。

d. Scheffe's test (例4.8)

用於同時檢定所有的線性組合，比其它方法保守，通常用於複合比較

e. <u>The Newman-Keuls Test</u>

類似 Duncan 法

f. Dunnett's test (例4.5)

用於資料中第一組為對照組時，執行試驗組與對照組的比較

Q: 如何選擇適當的比較方法

常用的方法有 LSD、Tukey's HSD 、Duncan 。

LSD法：判定差異顯著時的錯誤機率較高，容易分出差異，組數愈多，準確性愈差。

Scheffe法：錯誤率較低，是較保守的方法，不容易分出差異

以下列得到的檢定分界值為例，比較各方法 $(m=4,\ n_i = 4,\ \alpha=0.05)$

Bonferron： 分界值 = 8.66

Tukey HSD： 分界值 = 8.24

Duncan： 分界值 = 6.10，6.40，6.59

Fisher LSD： 分界值 = 6.10

Scheffe： 分界值 = 8.97

對對比較　SAS報表例一：

Tukey's Studentized Range (HSD) Test for y

Alpha	0.05
Error Degrees of Freedom	16
Error Mean Square	20.72125
Critical Value of Studentized Range	4.04609
Minimum Significant Difference	8.2368

分界值

Means with the same letter are not significantly different.

Tukey Grouping	Mean	N	group
A	100.020	5	C
B	86.560	5	D
C	70.920	5	B
C	70.060	5	A

字母相同之二組無顯著差異

對對比較　SAS報表例二：

有星號之二組差異顯著

Tukey's Studentized Range (HSD) Test for y

Comparisons significant at the 0.05 level are indicated by *.**				
group Comparison	**Difference Between Means**	**Simultaneous 95% Confidence Limits**		
C - D	13.460	5.435	21.485	***
C - B	29.100	21.075	37.125	***
C - A	29.470	21.787	37.153	***
D - C	-13.460	-21.485	-5.435	***
D - B	15.640	7.615	23.665	***
D - A	16.010	8.327	23.693	***
B - C	-29.100	-37.125	-21.075	***
B - D	-15.640	-23.665	-7.615	***
B - A	0.370	-7.313	8.053	

說明：在研究報告中，常以相同文字表示兩處理均值間差異不顯著，反之差異顯著，如下例：

表三、 不同蒸煮時間芋塊之質地剖面分析之參數

Table 3. Effect of steam-cooking time on texture parameter of cooked taro from texture profile analysis

	Primary parameter				Secondary parameter	
Steam -cooking time(min)	Firmness (g.force)	Springiness	Cohesiveness	Adhesiveness (g.force*mm)	Chewiness (g.force)	Gumminess (g.force)
20	n.d	n.d.	n.d.	n.d.	n.d.	n.d.
30	4327 a	0.949 a	0.328 a	1310 a	1356 a	1426 a
45	2920 b	0.948 a	0.280 b	1213 a	791 b	834 b
90	2733 b	0.931 a	0.300 ab	1005 b	773 b	816 b
120	2620 b	0.959 a	0.299 ab	817 c	728 b	784 b

[a,b,c]Means with different letters in the same column differ significantly ($p<0.05$)

Table 3. Mean arterial and portal plasma glucose concentrations, net portal-drained viscera (PDV) flux of glucose, and portal absorption of ileally digested glucose in pigs fed diets with 65% cornstarch (CS), 32.5% cornstarch + 32.5% potato starch (CPS), and 65% potato starch (PS)[a]

	Diet				
Item	CS	CPS	PS	SED[b]	P
Artery, mmol/L	5.87[c]	5.42[d]	5.45[d]	.10	.025
Portal vein, mmol/L	8.14[c]	6.94[d]	5.97[e]	.26	.003
Net portal flux, mmol/12 h	2,225.7[c]	1,310.2[d]	589.9[e]	130.1	.001
Absorption, %	88.9[c]	65.9[cd]	40.6[d]	10.2	.023

[a]Values are based on four observations.
[b]SED = SE of the difference.
[c,d,e]Means within a row with a different superscript are different at $P < .05$.

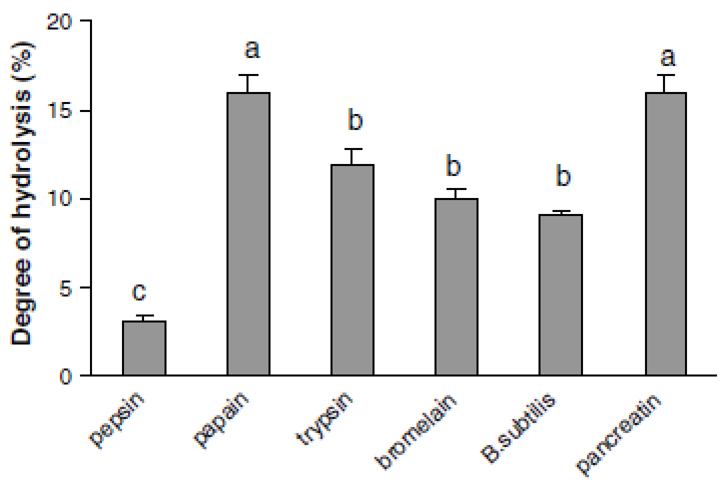


Fig. 1. Degree of hydrolysis of cell proteins in ethanol extracted *Chlorella* biomass with different proteases. A 10% suspension in water of extracted biomass was hydrolysed (20 U/g) at 37 °C for 4 h, accompanied by continuos stirring. The pH was adjusted to the optimum for each enzymatic preparation. Each value is the mean of three experiments. Means without the same letter above bars are significantly different at the 5% level according to the Student–Newman–Keuls test.

II. 對比量 (Contrast) 檢定

--- 二個特殊平均值組合的比較。

Contrast 定義： (for balanced data)

δ is a contrast of the treatment means, if

$$\delta = \sum_{i=1}^{m} c_i \mu_i, \quad \text{where} \quad \Sigma c_i = 0.$$

檢定 $H_0 : \delta = 0$ (i.e. $\Sigma c_i \mu_i = 0$)

$$F = MSc / MS_E ,$$

Where $MS_C = SS_C / 1, \quad SS_C = \dfrac{(\Sigma c_i Y_{i.})^2}{n^{-1} \cdot \Sigma c_i^2}$,

If p-value <α，$\delta \neq 0$ is significant. (差異顯著)

註：若 5 組中前 2 組與後 3 組比較，contrast 的係數為 (3, 3, -2, -2, -2)

【**Exp**】 5 種廣告 (A, B, C, D, E) 效果的比較

E 是一般廣告，A、B強調價格便宜，C、D強調品質優異

1. 廣告 A, B 與 C, D 之比較，constrast 係數為 (1, 1, -1, -1, 0)
2. 廣告 A, B, C, D 與 E 之比較，constrast 係數為 (1, 1, 1, 1, -4)

SPSS報表

對比係數

對比	AA 1	2	3	4	5
1	1	1	-1	-1	0
2	1	1	1	1	-4

對比檢定

		對比	對比值	標準誤	t	自由度	顯著性 (雙尾
PRICE	假設變異數相等	1	1.1667	.1401	8.326	25	.000
		2	.4200	.3133	1.341	25	.192
	未假設變異數相等	1	1.1667	.1442	8.090	18.031	.000
		2	.4200	.2841	1.478	8.780	.174

III、<u>Regression Analysis</u> (Chap 18)

對數量型的 treatment，對對比較的結果有時並無意義，可以迴歸分析，尋找一因子對反應值的估計關係式，此關係式可用來預測反應值，或估計極值；下為一例。

【EXC 3.2】劑量對血壓的影響

¾ ¶			
5mg	10mg	15mg	20mg
98	89	85	86
92	90	90	87
90	88	89	84
93	91	92	83
92	94	93	85
95	92	96	86

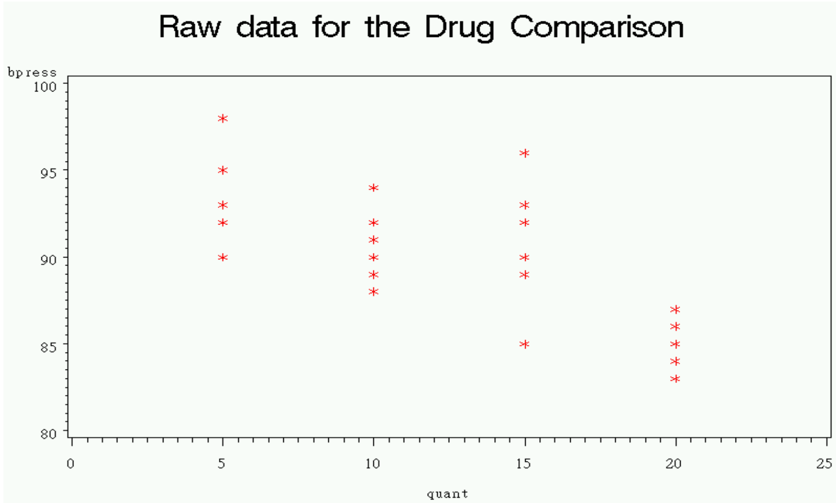
數量型的 treatment

SAS報表： ANOVA及對對比較

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	213.6666667	71.2222222	9.87	0.0003
Error	20	144.3333333	7.2166667		
Corrected Total	23	358.0000000			

Means with the same letter are not significantly different.			
t Grouping	**Mean**	**N**	**dosage**
A	93.333	6	5
A	90.833	6	15
A	90.667	6	10
B	85.167	6	20

有下降趨勢



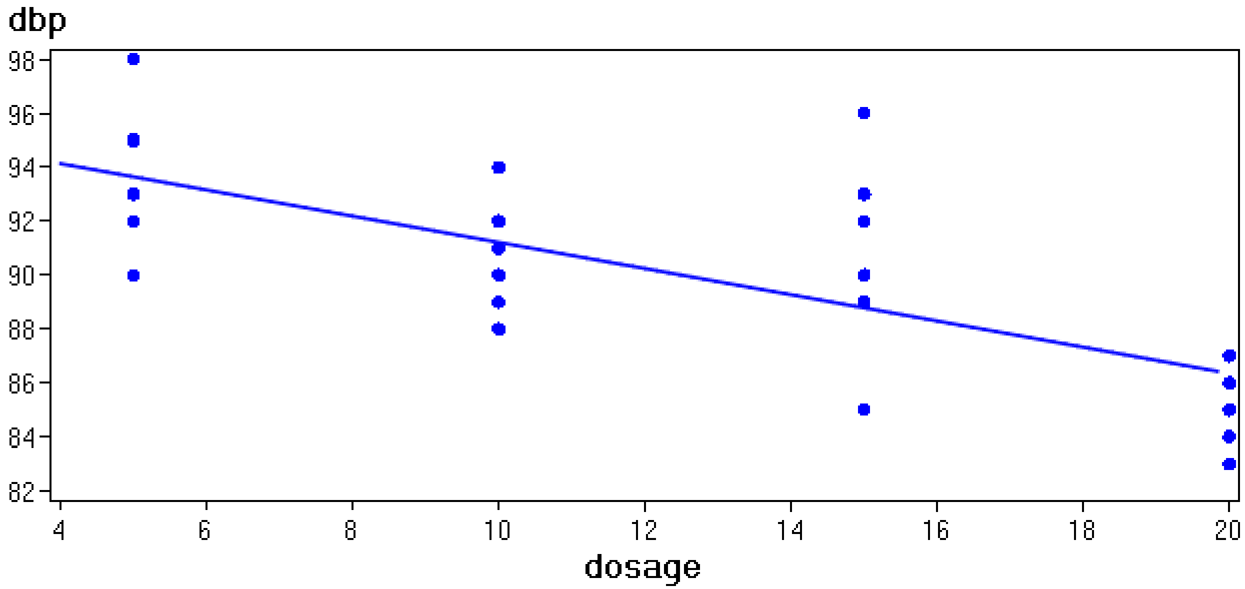
由ANOVA得到：劑量對老鼠血壓的影響是顯著的 (p=0.0003)，對對比較結果是劑量為20時與其它情況下有顯著差異，實際上劑量對血壓是一種漸進的影響；由資料分散圖猜測其影響可能是線性的或是二次關係的，

SAS報表： Regression Analysis

Root MSE	2.86330	**R-Square**	0.4962
Dependent Mean	90.00000	**Adj R-Sq**	0.4733

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	**1**	96.08333	1.43165	67.11	<.0001
dosage	dosage	**1**	-0.48667	0.10455	-4.65	0.0001



劑量對血壓的影響是顯著的

經由迴歸分析 (regression analysis) 得到直線關係式為：

$$血壓 = 96.1 - 0.487\ (劑量)\ ， \qquad R^2 = 0.496$$

若劑量增加 1 單位，血壓平均減少 0.487

此關係式可解釋 49.6 % 血壓的變量

第五章　資料轉換

使用變異數分析 (ANOVA) 之前提是：

(1)、累加性 -- 處理效應與環境效應間互相獨立。

(2)、獨立性 – 實驗誤差間互相獨立。

(3)、同質性(homogeneity) -- 各處理之試驗誤差變異數相同。

(4)、常態性 --實驗誤差的分佈符合常態分配。

接受變異數分析為結果之前　，　應先檢查資料是否符合上述前提。

(1) 適當設計之試驗，通常都符合累加性。

(2) 利用<u>隨機原則</u>設計的試驗，可促使誤差滿足獨立性。

(3) 變異數同質性檢定

$H_0：\sigma_1^2 = \ldots = \sigma_m^2$

變異數同質性檢定的方法有：

- Hartley’s F-max test (1950)
- Bartlett’s test (1949)
- Box test
- Levene test (1960)

當資料分佈是常態時， Bartlett法是很好的方法，但是它對分佈型態很敏感，依據 Conover etc. (1981), Levene檢定法不受限於資料之分佈，是較適當的方法。

SPSS提供 Levene 檢定法；

SAS提供 Bartlett，Brown-Forsythe 及 Levene 檢定法。

例5.1 三種同質性檢定法結果 (SAS報表)

Levene's Test for Homogeneity of germ Variance
ANOVA of Squared Deviations from Group Means

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
concent	3	1367.2	455.7	2.38	0.1004
Error	20	3835.3	191.8		

Brown and Forsythe's Test for Homogeneity of germ Variance
ANOVA of Absolute Deviations from Group Medians

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
concent	3	34.5218	11.5073	2.69	0.0738
Error	20	85.5632	4.2782		

Bartlett's Test for Homogeneity of germ Variance

Source	DF	Chi-Square	Pr > ChiSq
concent	3	9.2077	0.0267

(4) 常態性檢定 --- 檢定誤差是否遵循一常態分佈。

H_0：誤差呈現一常態分佈

SPSS提供 Kolmogorov-Smirnov, Shapiro-Wilk 檢定法；

SAS提供此二法另加 Cramer-von Mises 及 Anderson-Darling 檢定法。

基本上以殘差為檢定對象

$\boxed{\text{殘差 (residual)}}$

$$e_{ij} = Y_{ij} - \hat{Y}_{ij}, \quad \text{or}, \quad e_{ij} = Y_{ij} - \bar{Y}_i$$

殘差，e_{ij}，可視為觀測的誤差，用於估計真實誤差

若模式適合，則殘差正確地反映出 ε_{ij} 的特性

例：

三品種收量表

					均值
V1	14	6	11	15	11.5
V2	20	22	18	20	20
V3	22	24	27	19	23

殘差

2.5	-5.5	-0.5	3.5
0	2	-2	0
-1	1	4	-4

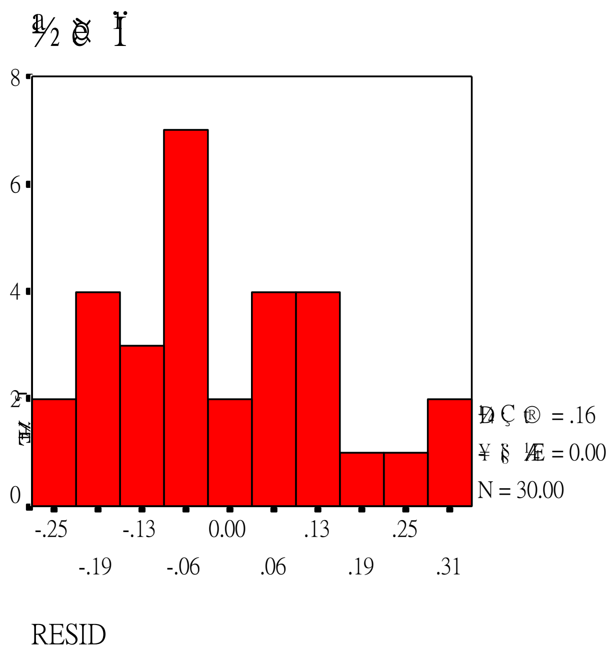
SPSS output

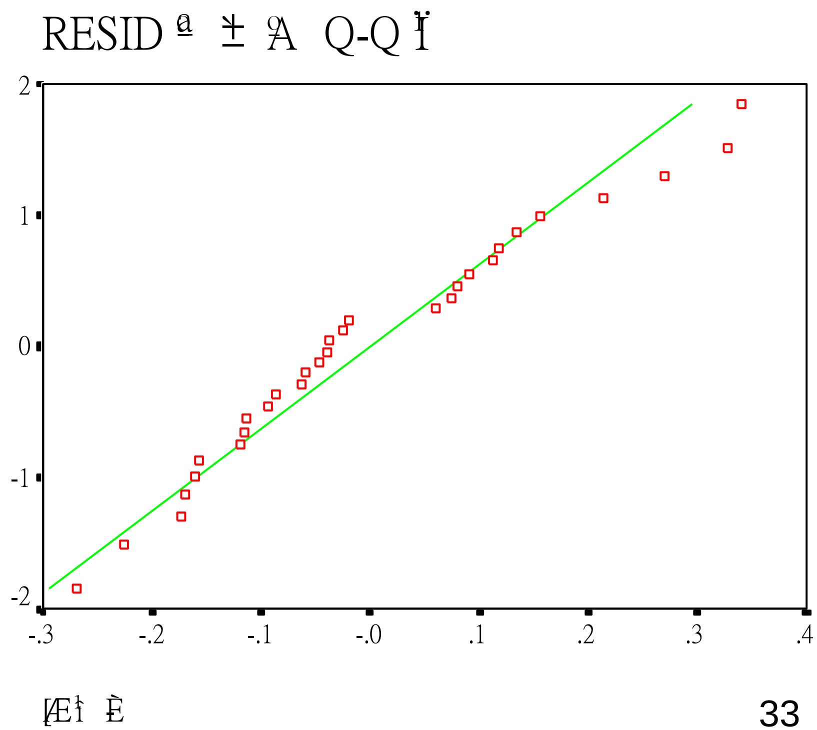
(先得到 residual, 再做資料預檢)

常態檢定

	Kolmogorov-Smirnov檢定[a]			Shapiro-Wilk 常態性檢定		
	統計量	自由度	顯著性	統計量	自由度	顯著性
RESID	.150	30	.083	.954	30	.317

a. Lilliefors 顯著性校正





將 e_i 標準化得到的值，稱為 t 化殘差

$$e_i^* = \frac{e_i}{\sqrt{MSE}}$$

如果資料未通過常態性檢定，可以檢查 t 化殘差，t 化殘差得到的值理論上應落在 ±3 之內，（$|e_i^*| < 3$ ），若有某筆資料在 ±3 之外，則此資料為離群值(outlier)，或異常值，可檢查資料取得過程是否有疑點。

例：探討餵食糖尿病腎病變SD大鼠富含polyunsaturated fatty acids (PUFAs)之飲食，是否可以進一步改善因腎臟脂肪囤積而引起的糖尿病腎病之病理過程。

Statistical analyses

Before assessing the different variables, we carried out a Kolmogorov-Smirnov test to check the normal distribution of the variables. Data that fit the normal distribution were compared by 1-way analysis of variance. The Levene test for homogeneity was used to test for equal variance between samples. When equal variance could be assumed, the Bonferroni post hoc test was used to identify significant differences between multiple test groups. When equal variance could not be assumed, the Games-Howell post hoc test was applied. A nonparametric Kruskal-Wallis test with post hoc Mann-Whitney U tests with the Bonferroni inequality was performed on the final body weight, kidney, retroperitoneal fat, plasma TG level, urinary Cr excretion, and renal SREBP-1 mRNA data, because it was in contrast with the normality hypothesis. The correlation coefficient was calculated using Spearman rank correlation coefficient. Data (NL and NR group, n = 6; DML group, n = 11; DMR group, n = 10) were presented as means ± SEM. The level of significance was set at P b .05. Analyses were performed using SPSS 11.0 for Microsoft Windows (SPSS, Chicago, Ill). Sample sizes were determined based on power analysis

[Exc3.1] （完整多組比較的分析法）

A, B, C, D四大豆品種進行產量比較 (p47)

Factor : 品種 **Observation** : 產量

4 treatments : A, B, C, D (qualitative)

分析步驟:

1. 整體分析 ：同質性檢定、常態性檢定、1-way ANOVA
2. 因子效應分析：對對比較
3. 估計各組平均值
4. 結論

1. 1-way ANOVA (SAS 報表)

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	39.83000000	13.27666667	23.67	0.0002
Error	8	4.48666667	0.56083333		
Corrected Total	11	44.31666667			

差異顯著

2. 變異數同質性檢定

Levene's Test for Homogeneity of yield Variance
ANOVA of Squared Deviations from Group Means

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
brand	3	1.4276	0.4759	2.45	0.1381
Error	8	1.5525	0.1941		

p-值大於0.05時，資料符合同質性

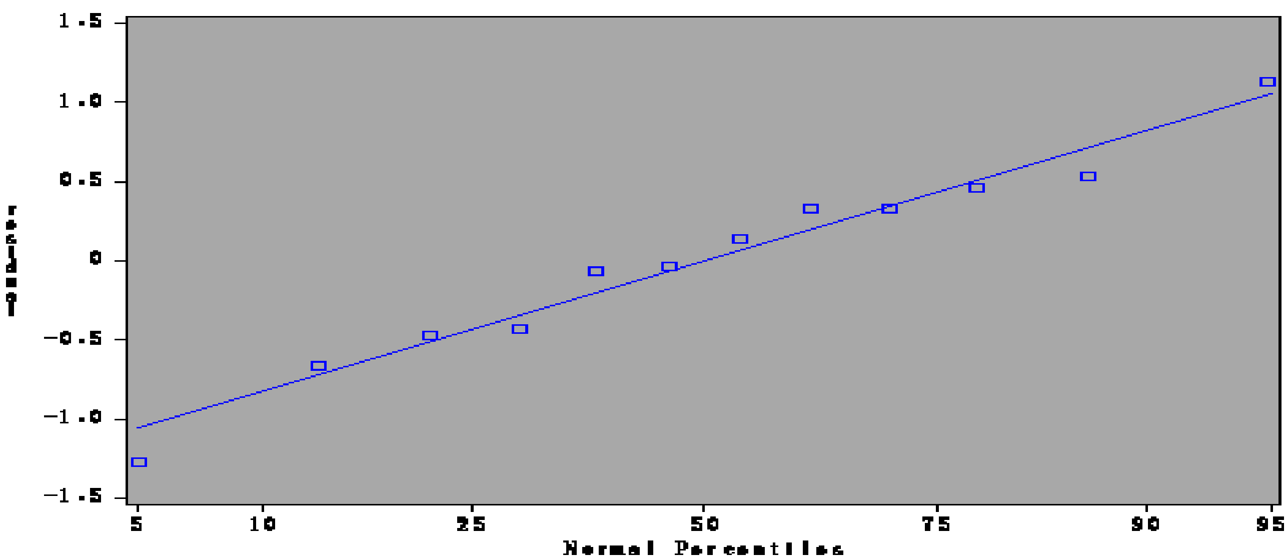
四組資料變異數相同

3. 檢測常態性 (先得到 residual, 再分析)

P-值大於0.05時，資料符合常態假設

Tests for Normality

Test	Statistic		p Value	
Shapiro-Wilk	**W**	0.978866	**Pr < W**	0.9787
Kolmogorov-Smirnov	**D**	0.124947	**Pr > D**	>0.1500
Cramer-von Mises	**W-Sq**	0.031349	**Pr > W-Sq**	>0.2500
Anderson-Darling	**A-Sq**	0.202274	**Pr > A-Sq**	>0.2500

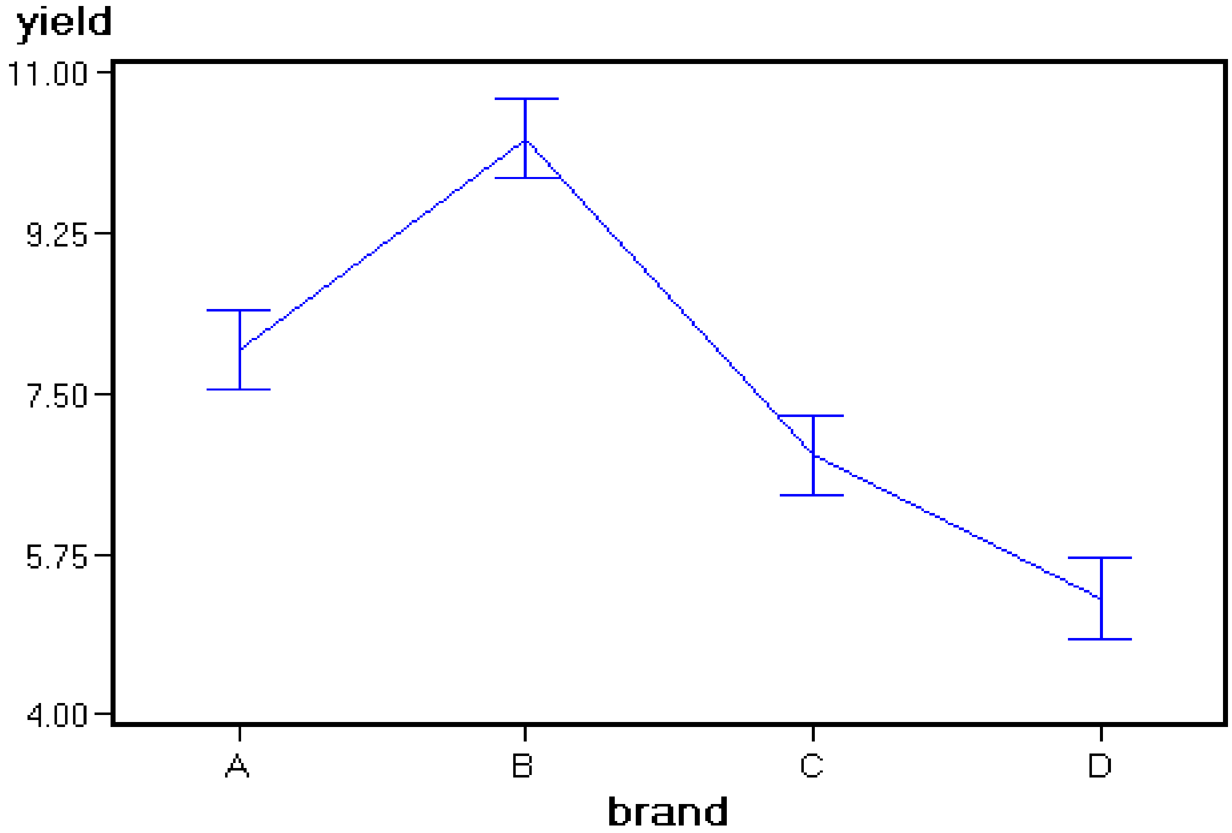


4. 對對比較

lsd Grouping	Mean	N	brand
A	10.2667	3	B
B	7.9667	3	A
B	6.8333	3	C
C	5.2667	3	D

A 與 C 無顯著差異，其餘各組間皆有顯著差異。

5. 平均值區間比較圖



總結

1、資料滿足同質性(Levene’s test, p=0.1381)及常態性(Shapiro-Wilk test, p=9787)。

2、由ANOVA得到四品種生產量的差異顯著(p=0.0002)。

3、由LSD法檢定出品種A 與 品種C 之間產量無顯著差異，其餘各組間皆有顯著差異。

4、品種B 的產量最高，品種D 的產量最低，各平均產量的估計及標準誤如下表：

品種	B	A	C	D
mean ± SEM	10.267±0.577[a]	7.967±0.503 [b]	6.833±.0.451 [b]	5.267±0.206 [c]
註：上標字母不同之各組間在α=0.05的水準下均值有顯著差異。				

[Example] （含離群值之例）

A, B, C三產品 protein 之比較

Factor : 品種　　　　　**Observation** : protein

3 treatments : A, B, C (qualitative)

三品種 protein

A	7	8	5	20
B	9	8	6	5
C	10	13	11	10

1. 1-way ANOVA

Source	DF	S S	M S	F Value	Pr > F
Model	2	34.6667	17.3333	1.01	0.4011
Error	9	154.000	17.1111		
Total	11	188.6667			

差異不顯著

2. 變異數同質性檢定

Levene's Test for Homogeneity of protein Variance
ANOVA of Squared Deviations from Group Means

Source	DF	S S	M S	F Value	Pr > F
BRAND	2	2818.7	1409.3	2.12	0.1759
Error	9	5979.0	664.3		

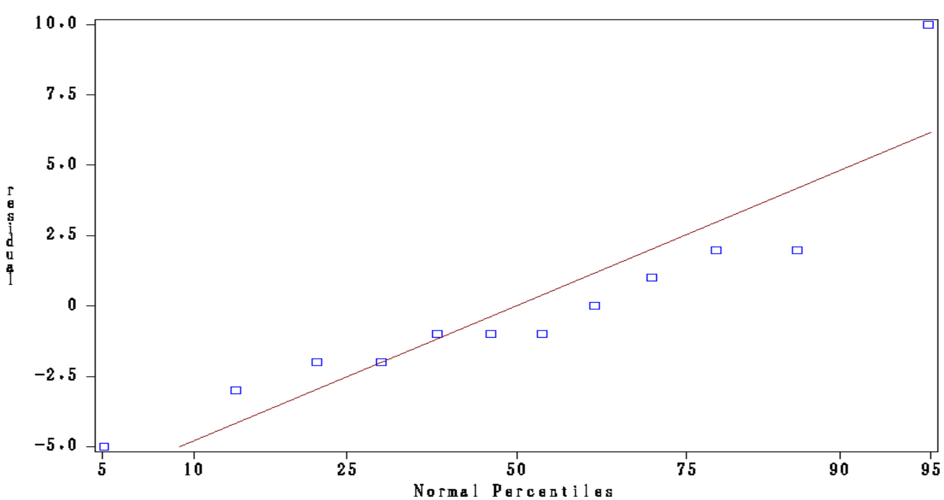
三組資料變異數相同。

3. 檢測常態性　(先得到 residual, 再分析)

P-值小於0.05時，資料不符合常態假設

Tests for Normality

Test	Statistic		p Value	
Shapiro-Wilk	W	0.841167	Pr < W	0.0286
Kolmogorov-Smirnov	D	0.213157	Pr > D	0.1346
Cramer-von Mises	W-Sq	0.115507	Pr > W-Sq	0.0629
Anderson-Darling	A-Sq	0.740192	Pr > A-Sq	0.0405



常態機率圖顯示有一離群資料

4. 對對比較

ANOVA 結論是三組無顯著差異，不需討論對對比較。

此資料得到無顯著差異的結果，且不滿足常態性，ANOVA的結果是不足以採信；但知有一離群個案造成無法滿足常態性，可以對此個案來源檢查，是記錄有誤，或確實是一特殊個案，是否可排除此個案，只對其它資料比較？
對不滿足常態性的資料，可改用無母數法，或是利用轉換的技巧，將資料換成常態資料後，再執行ANOVA。

資料轉換

當資料不符合常態性或同質性，可嘗試採用適當轉換法；
轉換需注意實質意義。

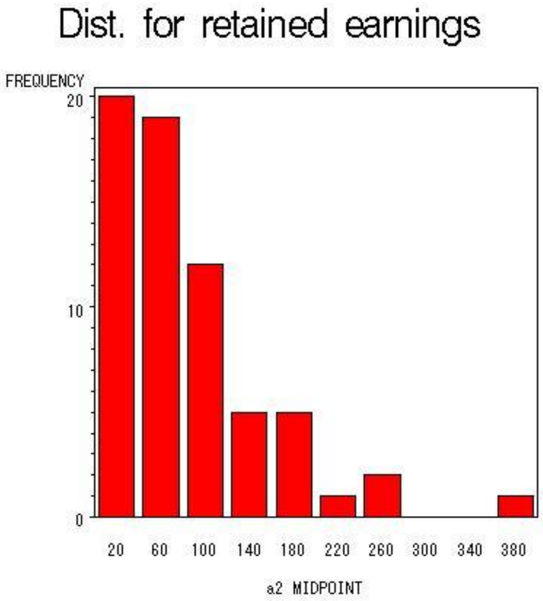
- 開方根轉換，如，蟲數，產蛋數。
- 對數轉換
- 角度轉換
- 倒數轉換
- 指數轉換

$$\sqrt{\mathrm{Y}},\ \log(\mathrm{Y}),\ 1/\mathrm{Y},\ \alpha Y^{\beta}$$

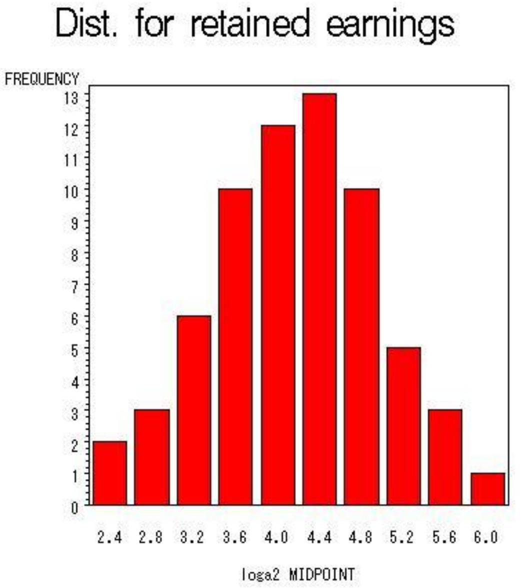
變數轉換的影響

- 線性轉換 通常不會改變原資料的分佈型態
- 次方或對數的轉換會改變分佈的形態。

原資料分佈



對數轉換後分佈



無母數 1-way ANOVA分析法

使用ANOVA F test時必須假設母體滿足同質性及常態性，但資料不一定符合假設，此時ANOVA分析的結果是受到質疑的，當面對此問題時，可採用Wilcoxon signed-rank test(二個樣本時)或Kurskal-Wallis檢定(分組類別有三個以上時)等無母數統計方法。

Kruskal-Wallis 檢定法

- 適用於同時對三個以上母體進行比較。
- 資料不符合常態性或同質性。
- 原理：以資料的序位代替資料大小值，來檢定中位數是否相等。

方法：

1. 將 k 組資料混合排序，列出序位， R_{ij}(遇到數值相同時，以平均序位取代)。
2. 計算：$R_{i.}$ = 第 i 組資料序位之和

3、 $S^2 = \frac{1}{N-1}[\Sigma\Sigma R_{ij}^2 - \frac{N(N+1)^2}{4}]$，

檢定量：$H = \frac{1}{S^2}[\Sigma \frac{R_{i.}^2}{n_i} - \frac{N(N+1)^2}{4}]$

此亦為序位的組內變異與組間變異的比值

4、 H-值 遵循 卡方分佈
5、H-值愈大，p-值愈小，顯示各組間之差異愈顯著。

例：檢定四品種豆子雞蛋白質含量的差異

- 在無母數法中，H_0 ： 4組中位數相等
- 例：

原資料

	A	B	C	D
	8.5	10.6	6.4	5.4
	7.9	9.6	7.3	6.4
	7.5	10.6	11.5	4.0
平均	7.97	10.27	8.4	5.27

A組均值比C組小，未反應實情

資料之序位

	A	B	C	D
	8	10.5	3.5	2
	7	9	5	3.5
	6	10.5	12	1
和	21	31	20.5	6.5

A組序位和比C組大，有反應實情

下例資料由ANOVA F-test 檢定出差異不顯著，因為資料非同質，使用K-W test 時，則檢定出差異顯著

	A	B	C
	8.5	10.3	6.4
	7.9	11.0	7.3
	7.5	10.6	13.0
	8.2	9.6	7.0
	7.8	10.6	7.2
平均	7.98	10.42	8.18

ANOVA F-test:

p-value = 0.0623

三組均值差異不顯著

Bartlett test ：

p-value= 0.0006

變異數差異顯著

K-W test:

p-value = 0.0385

三組差異顯著

多組資料比較的統計分析步驟：

1、作圖：各組均值的誤差長條圖。

2、變異數分析：檢定各組的差異是否顯著。

3、前提條件的檢測，包括同質性及常態性。

若無法滿足前提條件，應使用無母數之變異數檢定。

4、差異顯著時 → 作 multiple comparison (如對對比較)，或，

→ 以迴歸分析尋找關係。

5、針對特殊組合的比較，執行對比量檢定。

6、結論。

或稱為
Post hoc

1、輸入資料：定義及輸入: 組別佔一行，資料佔一行

2、變異數分析：分析 → 比較平均數法 → 單因子變異數分析

指定 (1) 依變數，固定因子

Post Hoc 檢定　選一方法　比對

選項　v描述性統計量 v同質性檢定 v平均數圖

3、常態性檢定：先輸入殘差

分析 → 描述性統計 → 預檢資料

統計圖　v 常態機率圖附檢定

4、作圖：統計圖 → 誤差長條圖

SAS guide

針對一維的完全隨機設計，以1-way ANOVA分析資料，其餘的實驗設計資料，大都使用 Linear model 分析。

1、匯入資料：檔案 → 匯入資料: 組別一行，測量值一行

2、分析 → ANOVA → 單因子 ANOVA

工作角色：組別為自變數，測量資料為應變數

檢定：選擇檢定等變異性的方法

平均值：選擇多重比較的方法，或欲列印的統計值

標繪圖：選擇平均數比較圖（可指定寬度）

結果：儲存結果

標題：

SAS guide

3、輸入殘差： 在excel資料檔中算出 residual

4、檢測常態性：描述 → 分佈分析

工作角色：殘差為分析變數

分佈：常態（可改變線條的顏色）

標繪圖：選擇 常態機率圖（可改變背景的顏色）

表格：只勾選 常態性檢定

標題：

5、無母數分析：分析 → ANOVA → 無母數單因子 ANOVA

工作角色：組別為自變數，測量資料為應變數

分析：勾選 Wilcoxon

精準p值：