Chapter 11

變異數分析

11.1 變異數分析之觀念

在實務應用中，常常需要探討相依變數與一個或多個自變數間之關係，亦即要討論某些因素是否對資料產生變動之研究。例如不同品牌之機器對產品品質是否有影響、不同品牌的飼料是否影響豬仔的重量，或顏色不同對產品銷售量是否有影響，此類問題皆屬於變異數分析之範圍。

前章已討論過單一及兩個常態母體平均數之檢定，本章將繼續探討關於三個或三個以上常態母體平均數是否相等之檢定方法，稱之爲變異數分析（Analysis of Variance；ANOVA）方法。

一、變異數分析之術語

㈠因子或因素：

因子（factor）爲研究分析中之獨立變數，亦稱之爲研究之分類標準。例如探討不同品牌飼料對豬仔重量是否有影響之問題中，品牌即爲此研究主題之因子。

㈡因子水準：

因子水準（factor level）是指所研究的因子之狀態及特殊形式，例如飼料有 A 、 B 、 C 三種，則此品牌因子即表示有三個水準。又如四種顏色紅、白、藍、綠對銷售量是否有影響，則顏色爲因子，且此因子有四個水準。

㈢**一因子與多因子分類：**

在變異數分析中，所探討之對象若祇包含一個獨立變數時，稱之爲一因子分類，但在研究時若同時針對二個或二個以上獨立變數作分類，則稱之爲多因子分類。

㈣**處理：**

在變異數分析中，處理（treatments）的意義隨研究時所探討因子之多寡而異。在一因子分析中，處理數即指因子之水準數，例如在研究三種品牌之飼料時，每種不同品牌即爲一個處理。在多因子分析中，所謂處理數是表示不同因子其因子水準的組合數，例如在探討飼料與地區不同對豬仔重量是否有影響之問題中，飼料因子有三個水準，而地區因子亦有三種（北、中、南三區），則其處理數共有$3\times3=9$種。

㈤**屬質因子及屬量因子：**

屬質因子（qualitative factor）是指因子的因子水準不可以用計量方式表達者，亦即屬性類別之型態，例如A、B、C三種品牌飼料，又如甲、乙、丙、丁四種不同之機器。而屬量因子（quantitative factor）即表示所研究之因子其因子水準可以用計量之方式表示者，如溫度、價格等。

二、變異數分析之基本假設

在作變異數分析時，爲了要發展優良的檢定統計方法，必須針對統計資料作下列的一些基本假設，即

㈠每一個處理所對應之母體分配，皆要假設服從常態母體。

㈡每一個常態母體之變異數皆相同，亦即變異數有同質性。

㈢來自每個常態母體之樣本資料皆具有隨機獨立性。

Remark

在變異數分析之三個基本假設中，對於常態假設而言，一般來說偏離常態母體是「穩健的」（robust），換句話說只要母體資料的分配不極端背離常態分配，則對檢定的檢定力影響不大，特別是在觀察資料很多的情況下。

11.2 一因子變異數分析

在變異數分析（ANOVA）中，若研究主題所探討對象，只包含一個獨立變數（因子）稱之爲一因子變異數分析（one factor ANOVA），又稱之爲單向變異數分析（one way ANOVA）。它所研究的問題是對於此因子的不同因子水準是否對研究對象有顯著差異。

一、模式之建立及其意義

㈠ANOVA模式：

若因子有 k 個水準，則可由此 k 個獨立的常態母體中各抽出一組隨機樣本（見表11-1），即

表11-1　一因子ANOVA之樣本資料

	母體						
	1	2	…	i	…	k	
	X_{11}	X_{21}		X_{i1}		X_{k1}	
	$\vdots$	$\vdots$	…	$\vdots$	…	$\vdots$	
	X_{1j}	X_{2j}		X_{ij}		X_{kj}	
	$\vdots$	$\vdots$	…	$\vdots$	…	$\vdots$	
	X_{1n_1}	X_{2n_2}	…	X_{in_i}	…	X_{kn_k}	
樣本和	$T_{1\bullet}$	$T_{2\bullet}$	…	$T_{i\bullet}$	…	$T_{k\bullet}$	$T_{\bullet\bullet}$
平均數	$\bar{X}_{1\bullet}$	$\bar{X}_{2\bullet}$	…	$\bar{X}_{i\bullet}$	…	$\bar{X}_{k\bullet}$	$\bar{X}_{\bullet\bullet}$

由表11-1中資料知，k 個因子水準皆對應於常態母體，故知每組隨機樣本之 X_{ij} 皆服從常態分配，即

$$X_{ij} \overset{\text{iid}}{\sim} N(\mu_i, \sigma^2) \quad ，i = 1,2,\ldots,k \text{ ，} j = 1,2,\ldots,n_i$$

因每一個觀察值 X_{ij} 與平均數 μ_i 之間可能有差異，故可將上式寫爲

$$X_{ij} = \mu_i + \varepsilon_{ij}$$

其中 $\varepsilon_{ij} \overset{\text{iid}}{\sim} N(0, \sigma^2)$ ，$i = 1,2,\ldots,k$ ，$j = 1,2,\ldots,n_i$ ，此即爲一因子變異數分析之模式。