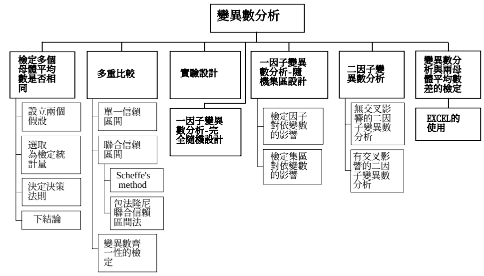
14 變異數分析

學習目的

1. 討論變異數分析的一般觀念，變異數分析的方法與步驟。
2. 學習檢定多個母體平均數是否相同。
3. 學習如何做多重比較。
4. 了解實驗設計的意義。
5. 學習一因子變異數分析-完全隨機設計方法。
6. 了解一因子變異數分析-隨機集區設計方法。
7. 學習二因子變異數分析的方法
8. 利用Excel統計軟體來做變異數分析

本章結構



變異數分析的意義

- **變異數分析的意義**

  檢定三個或三個以上的母體平均數是否相等的統計方法，或

  檢定因子對依變數是否有影響的統計方法。

檢定多個母體平均數是否相同

變異數分析的步驟

(1) 設立兩個假設

(2) 選取$\dfrac{MSF}{MSE}$為檢定統計量

(3) 決定決策法則

(4) 下結論

變異數分析的假設

- **變異數分析的假設**

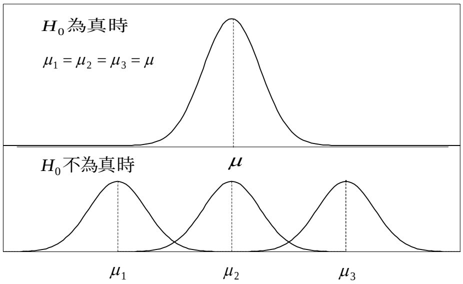
  ①假設因子對依變數的影響效果是固定的。亦即$\mu_i - \mu$為一常數而不是隨機變數。

  ②每個小母體均為常態分配，表為：$Y_i \sim N(\mu_i, \sigma_i^2), i = 1, \cdots, k$。

  ③變異數齊一性(Homogeneity)，即$\sigma_i^2 = \sigma^2$，每個小母體的變異數均相等。

  ④抽樣方法為獨立簡單隨機抽樣，即自k個小母體分別抽取獨立之隨機樣本。

圖14.1 虛無假設與對立假設



檢定多個母體平均數是否相同

- **總差異**

$$Y_{ij}-\mu=(\mu_i-\mu)+(Y_{ij}-\mu_i)$$

總差異=因子引起的差異+隨機差異

檢定多個母體平均數是否相同

- **總變異可分解為因子變異與隨機變異**

總變異 ＝因子變異 ＋隨機變異

$$SST = SSF + SSE$$

$$SST = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(Y_{ij}-\overline{\overline{Y}})^2$$

$$SSF = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(\overline{Y}_i-\overline{\overline{Y}})^2$$

$$SSE = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(Y_{ij}-\overline{Y}_i)^2$$

檢定多個母體平均數是否相同

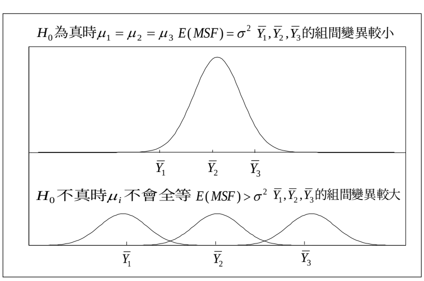
- **因子變異數(平均變異)**

$$MSF=\frac{SSF}{k-1}$$

- **隨機變異數(平均變異)**

$$MSE=\frac{SSE}{\sum n_i-k}$$

圖14.4 H_0 為真與不真的組間變異



檢定多個母體平均數是否相同

- **當H_0為真時$\frac{MSF}{MSE}$的抽樣分配**

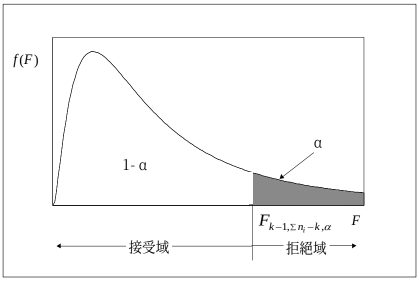
$$\frac{MSF}{MSE}\sim F_{k-1,\sum n_i-k}$$

- **決策法則**

①若 $F>F_{k-1,\sum n_i-k,\alpha}$， 則拒絕$H_0$。

②若 $F\le F_{k-1,\sum n_i-k,\alpha}$， 則接受$H_0$。

圖14.5 *F* 檢定的接受域與拒絕域



多重比較

- 單一信賴區間

$$(\bar{Y}_i-\bar{Y}_j)-t_{v,\frac{\alpha}{2}}\sqrt{MSE}\sqrt{\frac{1}{n_i}+\frac{1}{n_j}}\le\mu_i-\mu_j\le(\bar{Y}_i-\bar{Y}_j)+t_{v,\frac{\alpha}{2}}\sqrt{MSE}\sqrt{\frac{1}{n_i}+\frac{1}{n_j}}$$

- Scheffe's method的聯合信賴區間

$$\bar{Y}_i-\bar{Y}_j\pm\sqrt{(k-1)F_{k-1,\sum n_i-k,\alpha}}\sqrt{MSE(\frac{1}{n_i}+\frac{1}{n_j})}$$

$$i,j=1,\cdots,k,\ i\ne j$$

多重比較

- **聯合信賴區間**

若有 k 個小母體，則可求 $m=C_2^k$ 個母體均數差 $1-\alpha$ 的聯合信賴區間如下：

$$\bar{Y}_i-\bar{Y}_j\pm t_{\sum n_i-k,\alpha/2m}\sqrt{MSE}\sqrt{\frac{1}{n_i}+\frac{1}{n_j}}$$

其中 $t_{\sum n_i-k,\alpha/2m}$ 是指在自由度為 $\sum n_i-k$ 的 t 分配右尾機率為 $\alpha/2m$ 的臨界值。

圖14.7 完全隨機實驗

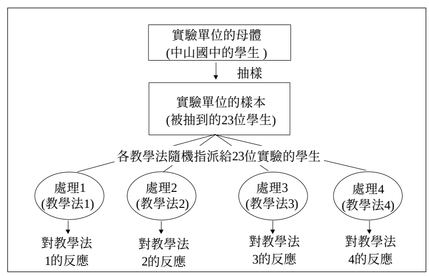


圖14.8 一因子完全隨機實驗設計

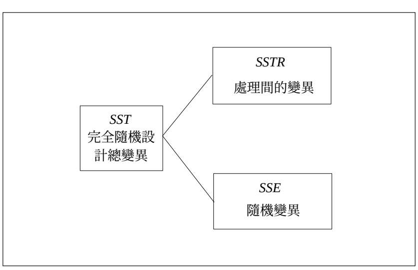


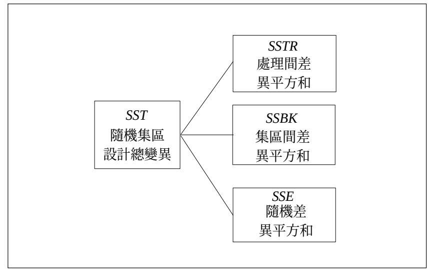
圖14.9 隨機集區設計

B_1	A_3	A_1	A_2
B_2	A_3	A_2	A_1
B_3	A_1	A_3	A_2
B_4	A_2	A_3	A_1

表14.12 隨機集區實驗的統計資料

集區	因子(處理)				$\bar{Y}_{i\cdot}$
	A_1	A_2	$\cdots$	A_k	
B_1	Y_{11}	Y_{12}	$\cdots$	Y_{1k}	$\bar{Y}_{1\cdot}$
B_2	Y_{21}	Y_{22}	$\cdots$	Y_{2k}	$\bar{Y}_{2\cdot}$
$\vdots$	$\vdots$			$\vdots$	$\vdots$
B_b	Y_{b1}		$\cdots$	Y_{bk}	$\bar{Y}_{b\cdot}$
$\bar{Y}_{\cdot j}$	$\bar{Y}_{\cdot 1}$	$\bar{Y}_{\cdot 2}$	$\cdots$	$\bar{Y}_{\cdot k}$	$\bar{\bar{Y}}$

圖14.10 一因子隨機集區實驗設計



一因子變異數分析—隨機集區設計

- **總變異**

$$SST = \sum_{i=1}^{b}\sum_{j=1}^{k}(Y_{ij} - \bar{\bar{Y}})^2$$

- **因子變異(處理間的變異)**

$$SSF = \sum_{i=1}^{b}\sum_{j=1}^{k}(\bar{Y}_{\cdot j} - \bar{\bar{Y}})^2 = b\sum_{j=1}^{k}(\bar{Y}_{\cdot j} - \bar{\bar{Y}})^2$$

- **集區間變異**

$$SSBK = \sum_{i=1}^{b}\sum_{j=1}^{k}(\bar{Y}_{i\cdot} - \bar{\bar{Y}})^2 = k\sum_{i=1}^{b}(\bar{Y}_{i\cdot} - \bar{\bar{Y}})^2$$

- **隨機變異**

$$SSE = \sum_{i=1}^{b}\sum_{j=1}^{k}(Y_{ij} - \bar{Y}_{i\cdot} - \bar{Y}_{\cdot j} + \bar{\bar{Y}})^2 = SST - SSF - SSBK$$

一因子變異數分析—隨機集區設計

- **因子平均變異(變異數)**

$$MSF = \frac{SSF}{k-1}$$

式中：$k-1$ 為 SSF 的自由度。

- **集區間平均變異(變異數)**

$$MSBK = \frac{SSBK}{b-1}$$

式中：$b-1$ 為 $SSBK$ 的自由度。

- **隨機平均變異(變異數)**

$$MSE = \frac{SSE}{(k-1)(b-1)}$$

式中：$(k-1)(b-1)$ 為 SSE 的自由度(或表為 $kb-k-b+1$)。

一因子變異數分析—隨機集區設計

檢定因子對依變數的影響

- F 檢定統計量

$$\frac{MSF}{MSE} \sim F_{k-1,kc-k-c+1}$$

- 決策法則

①若 $F > F_{k-1,kc-k-c+1,\alpha}$，則拒絕 H_0。

②若 $F \le F_{k-1,kc-k-c+1,\alpha}$，則接受 H_0。

一因子變異數分析—隨機集區設計

檢定集區對依變數的影響

- F 檢定統計量

$$\frac{MSBK}{MSE} \sim F_{c-1,kc-k-c+1}$$

- 決策法則

①若 $F > F_{c-1,kc-k-c+1,\alpha}$，則拒絕 H_0。

②若 $F \le F_{c-1,kc-k-c+1,\alpha}$，則接受 H_0。

表14.14　四種教學法的學習效果

集區	處理				集區和	集區平均$\overline{Y}_{i.}$
	教學法1	教學法2	教學法3	教學法4		
學生1	76	75	73	80	304	76
學生2	78	71	76	87	312	78
學生3	73	80	69	81	303	75.75
學生4	79	78	65	80	302	75.5
學生5	80	75	83	82	320	80
學生6	67	77	78	81	303	75.75
學生7	72	86	79	79	316	79
處理和	525	542	523	570	2,160	
處理平均數$\overline{Y}_{.j}$	75	77.43	74.71	81.43		

圖14.11　二因子的變異數分析

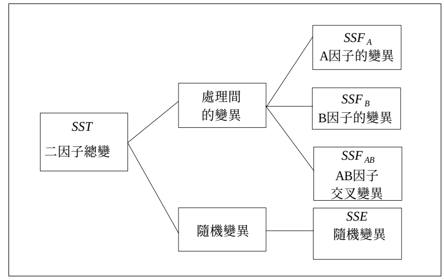


圖14.13　無交叉影響

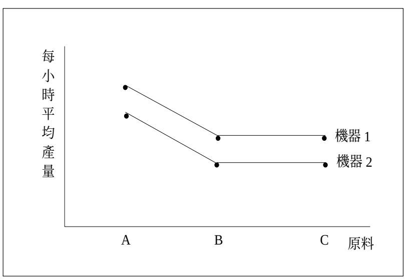
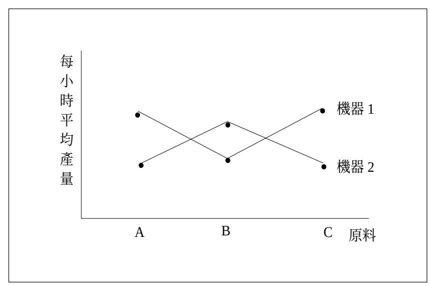


圖14.14　有交叉影響



無交叉影響的二因子變異數分析

檢定A因子有無影響

- ***F*檢定統計量(A 因子)**

$$F_A = \frac{MSF_A}{MSE} \sim F_{(r-1),(r-1)(c-1)}$$

- **決策法則**

在選定的顯著水準α下，決策法則為：

①若$F_A > F_{(r-1),(r-1)(c-1),\alpha}$，則拒絕$H_0$。

②若$F_A \le F_{(r-1),(r-1)(c-1),\alpha}$，則接受$H_0$。

無交叉影響的二因子變異數分析

檢定B因子有無影響

- ***F*檢定統計量(B因子)**

$$F_B = \frac{MSF_B}{MSE} \sim F_{(c-1),(r-1)(c-1)}$$

- **決策法則**

在選定的顯著水準α下，決策法則為：

①若$F_B > F_{(c-1),(r-1)(c-1),\alpha}$，則拒絕$H_0$。

②若$F_B \le F_{(c-1),(r-1)(c-1),\alpha}$，則接受$H_0$。

有交叉影響的二因子變異數分析

檢定A因子有無影響

- **檢定統計量**

$$F_A = \frac{MSF_A}{MSE} \sim F_{(r-1),(n_T-rc)}$$

- **決策法則**

在選定的顯著水準α下，決策法則為：

①若 $F_A > F_{(r-1),(n_T-rc),\alpha}$，則拒絕 H_0。

②若 $F_A \le F_{(r-1),(n_T-rc),\alpha}$，則接受 H_0。

有交叉影響的二因子變異數分析

檢定B因子有無影響

- **檢定統計量**

$$F_B = \frac{MSF_B}{MSE} \sim F_{(c-1),(n_T-rc)}$$

- **決策法則**

在選定的顯著水準α下，決策法則為：

①若 $F_B > F_{(c-1),(n_T-rc),\alpha}$，則拒絕 H_0。

②若 $F_B \le F_{(c-1),(n_T-rc),\alpha}$，則接受 H_0。

有交叉影響的二因子變異數分析

檢定AB因子是否有交叉影響

- **檢定統計量**

$$F_I = \frac{MSI}{MSE} \sim F_{(r-1)(c-1),(n_T-rc)}$$

- **決策法則**

在選定的顯著水準α下，決策法則為：

①若 $F_I > F_{(r-1)(c-1),(n_T-rc),\alpha}$，則拒絕 H_0。

②若 $F_I \le F_{(r-1)(c-1),(n_T-rc),\alpha}$，則接受 H_0。

圖14.17　兩種車輛耗油平均數差的檢定