

3. 通常我們用什麼樣的統計指標評估一估計值 (estimator) 的「有效性」(efficiency)? 請列出式子並舉例說明之。【10分】

命中講義 P.89

心理與教育統計

#### 2. 有效性(efficiency)

- (1) 在其他條件不變(如  $n$  固定), 母群參數的所有不偏估計量中, 具有最小變異數者, 稱為「最有效的估計量」。
- (2) 若  $E(\hat{\theta}_i) = \theta, i=1, 2, 3, \dots, k$ , 則  $\min \text{Var}(\hat{\theta}_i)$  就是「最有效估計量」(most efficient estimator)。
- (3) 例如,  $\bar{X}$  和  $Md$  皆為母群平均數  $\mu$  的不偏估計量, 但因為 R.E(相對有效性) =  $\frac{\text{Var}(Md)}{\text{Var}(\bar{X})} = 157\%$ , 所以  $\text{Var}(Md) > \text{Var}(\bar{X})$ , 亦即  $\bar{X}$  比  $Md$  的變異數更小, 所以  $\bar{X}$  比  $Md$  的變異數更小, 所以  $\bar{X}$  比  $Md$  更合乎有效性。

#### 3. 一致性(consistency)

所謂「一致性」是指樣本容量無限增大時, 估計值應越來越接近他的母群參數。即  $n \rightarrow \infty$  時,  $\bar{X} \rightarrow \mu, \hat{S}^2 \rightarrow \sigma^2, \hat{S}_{YX}^2 \rightarrow \sigma_{YX}^2$ 。

#### 4. 充份性(sufficiency)

指一個容量為  $n$  的樣本統計量, 是否充分地代表  $n$  個數據, 去反映母群的性質。例如  $\bar{X}$  較  $Mo, Md$  能充分地代表所有數據, 去反映母群性質; 同樣  $S^2$  比  $AD$ (平均差)和  $Q$ (四分差)更具充分性。

#### 5. 抗拒性(Resistance):

指的是樣本估計量是否能夠不受極端值影響, 例如: 中位數比平均數就不受極端值影響。

#### 8-2 區間估計

一、根據點估計值、抽樣分配及機率原理去推估母群參數在某一區間或範圍的統計方法, 即稱為「區間估計」。

二、原理:

$$\begin{aligned} \text{母數} &= \text{點估計值} \pm \text{抽樣誤差} \\ &= \text{點估計值} \pm (\text{臨界值} \times \text{點估計值的標準誤}) \end{aligned}$$

4. 以下為  $t$  test 的部份臨界值 (critical values) 表格。請據此算出  $F$  test 在分子自由度為 1 及分母自由度為 4，位於 .05 顯著水準 (significance level) 的臨界值 (四捨五入至小數點後第三位)。【10 分】



df	$t_{.40}$	$t_{.70}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$
1	.325	.727	1.376	3.078	6.314	12.706	63.657
2	.289	.617	1.061	1.886	2.920	4.303	9.925
3	.277	.584	.978	1.638	2.353	3.182	5.841
4	.271	.569	.941	1.533	2.132	2.776	4.604
5	.267	.559	.920	1.476	2.015	2.571	4.032

## 命中講義 P.79

心理與教育統計

- 每一種自由度就有一種  $t$  分配曲線，如圖 9-3。
- 曲線下的總面積為 1，故為機率分配，其機率值可查  $t$  分配表。
- $t^2(df) = F_{(1,df)}$ ，即  $t_{1-\frac{\alpha}{2}(df)} = F_{1-\alpha(1,df)}$

證明：

$$t_{(df)}^2 = \frac{Z^2}{\chi^2/df} = \frac{\chi^2/1}{\chi^2/df} \quad (df = 1 \text{ 時}, \chi^2 = Z^2) = F_{(1,df)}$$

### h. 卡方( $\chi^2$ )分配

自常態母群中，每次隨機抽取  $n$  個變量  $X_1, X_2, X_3, \dots, X_n$ ，並將其轉換為  $Z_i$ ，再加以平方，即  $Z_i^2 = (\frac{X_i - \mu}{\sigma})^2$ ，然後將這  $n$  個  $Z_i^2$  加起來，求其總和  $\sum Z_i^2$ ，則命名  $\sum Z_i^2$  為卡方( $\chi^2$ )，讀做 chi square)。如此進行抽樣及轉換無窮多次，即可得到無限多個  $\chi^2$  所形成的分配，稱為自由度  $df$  的  $\chi^2$  分配，以  $\chi_{df}^2$  表示。

#### ※ $\chi^2$ 分配的特性

- 因為  $\chi^2$  為  $Z$  的平方和，所以  $0 < \chi^2 < \infty$ ，即  $\chi^2$  均為正值。
- 每一種  $df$ ，就有一條卡方分配曲線。
- $\chi^2$  分配為正偏態分配，但當  $df$  增加， $\chi^2$  分配逐漸接近常態分配，可轉換成  $Z$  值，再查常態分配表。
- 自由度為  $df$  的卡方分配，其平均數為  $df$ ，標準差為  $\sqrt{2df}$
- 曲線下的面積為 1，故為機率分配，其機率值可查  $\chi^2$  分配表。

5. 考慮  $n$  對觀察值  $(x_i, y_i)$  ( $i = 1, 2, \dots, n$ ) 的簡單回歸 (simple regression)。我們可將最小平方法 (method of least squares) 用於 regression of  $Y$  on  $X$  以及 regression of  $X$  on  $Y$ ，而分別得到其回歸線的斜率估計值。請問如何藉此二斜率估計值來推估  $X$  與  $Y$  的相關係數 (correlation coefficient)？請詳述理由。【15分】

## 命中講義 P.183

心理與教育統計

### 十、相關比(曲線相關)( $\eta/\text{eta}$ ):

當  $X$ 、 $Y$  變項之間的相關，不是直線相關時，則積差相關便不適用。例如：焦慮和成績通常呈二次函數關係，則必需使用相關比。

$$\eta = \sqrt{\frac{SS_b}{SS_t}}$$

### 12-4 淨相關、部分相關及複相關

#### 一、淨相關(partial correlation)(偏相關)：

兩個變項在它們與一個或多個變項的共同解釋力被排除之外，所剩下的純相關程度以  $r_{12.3}$  表示。

例如：國語成績和數學成績的相關極高，但是國語、數學都會受到智力的影響，因此將智力自果與、數學成績中分別排除之後所剩下的純相關即是「淨相關」。

$$r_{12.3} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{1 - r_{13}^2}\sqrt{1 - r_{23}^2}}$$

#### 二、部分相關(part correlation)(半淨相關)：

指「第一變項」和「排除第三變項共同解釋力後的第二變項」之間的相關程度。

以  $r_{1(2.3)}$  表示。

例如：國語成績和排除致力因素後的數學成績間的相關。

$$r_{1(2.3)} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{1 - r_{23}^2}}$$

#### 三、複相關(multiple correlation)(多元相關)：

指某依變項和數個自變項間的共變關係。

例如：數學成績( $Y$ )可能和智商( $X_1$ )練習次數，作答習慣( $X_2$ )都有相關。

$$R = \sqrt{\beta_1 r_{Y1} + \beta_2 r_{Y2}}$$