

《教育測驗與統計》

試題評析

今年的考題目延續以往傳統，兩題申論、兩題計算。

第三題計算題的區間估計與假設檢定所使用的公式在高上《教育測驗與統計》講義的第10章P.42-43有詳細表列，類似題目在課堂上也一再演練，班內學員應可以拿到全部分數。第四題在講義的P.51頁也曾出現，惟其第二小題考出較偏的自由度公式，同學須經仔細思考後方能完整作答；至於第一題的電腦化適性測驗，是課堂上一再強調的重點，同學應不難發揮；至於第二題兩種測驗的差異，則可以透過平日上課的舉例(如：智力與性向測驗的陸軍普通分類測驗等)大致回答其內容。整體而言，此次考題屬於簡單中帶有些許困難，估計班上考生可以拿到75分至80分，一般的考生應可獲得65至70分。

一、試述電腦化適性測驗 (computerized adaptive testing) 的優缺點。(二十五分)

答：

電腦化適性測驗(computerized adaptive testing, CAT)所進行方式，其呈現給受試者的試題順序，是依據個人在前一個試題的表現好壞來作決定，因此對每一位考生而言，題目會不盡然完全相同。

其優點有：

1. 接續的題目應該是對受試者能力估計精確性最有貢獻的最大訊息量的試題。既能縮減測驗長度，又不犧牲測量的精確性。
2. 有因材施教的效果。
3. 閱卷可捨人工方式而由電腦執行。
4. 應用試題反應理論，雖然不同受試者考不同的試題，但由於精確估計的結果，可以比較其分數之差異。
5. 增強施測的標準化過程。

其缺點有：

1. 必須先建立一套完整的題庫，題庫的品質好壞可能影響CAT的效果。
2. 須搭配電腦系統實施，在電腦系統不穩定的情況下，易導致測驗中斷而影響受試者情緒。

二、何謂「能力」(ability)？何謂「成就」(achievement)？而測量兩者所編製出的測驗之間有何差異？(二十五分)

答：

1. 通常我們習慣以智力來稱呼能力，因而智力測驗在測量受試者的學習能力，分析受試者心理活動的歷程，所以取材以能獲得推理、判斷、記憶、與解決問題等能力為主；成就測驗則以測量受試者由學習所獲得的知識或技能的成就為主。

2. 智力測驗經常採用年級常模；而成就測驗則採年齡常模。

3. 智力測驗測量的是受試者一般的心智能力；而成就測驗則是測量受試者就某學科學習所養成的特殊能力。

另外，在目的、內容、與重點等有所不同。有時不易劃分、分辨。

三、假設225位大學男教授的育嬰性向測驗分數平均為110，標準差為5，全國的男性在此測驗上的平均數為100，標準差為15，倘若 $\alpha = 01$ ：(二十五分)

(一)以區間估計法求大學男教授的育嬰性向測驗真正分數。請列出公式、計算式、答案及結論。

(二)請問我們可否說大學男教授的平均育嬰性向比全國男性平均為高？請列出假設、公式、計算式、答案及結論。

請選用合適者

參考數值 $Z_{1-01} = 2.23$ $Z_{1-05} = 2.58$

參考公式 $Z = \frac{(\bar{X}) - \mu}{\sigma_{\bar{X}}}$ $Z = \frac{(X - M)}{SD}$ $T = 10Z + 50$

答：

1. 使用公式： $\bar{X} \pm Z_{\alpha/2} \frac{S}{\sqrt{n}}$

$$\text{算式與答案： } 110 \pm 2.58 \frac{15}{\sqrt{225}} = 110 \pm 2.58 = (107.42, 112.58)$$

結論：以此225位大學男教授的育嬰性向測驗平均分數，我們有99%的信心估計全國大學男教授的真正平均分數介於(107.42, 112.58)的區間中。

2. H_0 : 此225位大學男教授的育嬰性向測驗平均分數不高於全國男性平均分數

H_1 : 此225位大學男教授的育嬰性向測驗平均分數高於全國男性平均分數

$$\text{使用公式： } Z = \frac{\bar{X} - m}{s_{\bar{X}}}$$

$$\text{算式與答案： } Z = \frac{110 - 100}{15 / \sqrt{225}} = \frac{10}{1} = 10$$

結論：由於 $10 > 2.33$ ，因此拒絕虛無假設，在0.01的顯著水準下有充分證據證明該225位大學男教授的育嬰性向測驗平均分數確實高於全國平均分數。

四、假設甲乙兩校六年級學生之體重均符合常態分配，今自甲校隨機抽取10位學生，乙校隨機抽取12位學生，得樣本平均數與標準差分別為： $\bar{X}_1 = 34$ ， $\hat{s}_1 = 5$ ， $\bar{X}_2 = 31$ ， $\hat{s}_2 = 2$ ，試問

(一)兩者體重之變異數差異是否達到 $\alpha = .05$ 顯著水準？(九分)

(二)兩者體重之平均數差異檢定 t 值為多少？(八分) 自由度為多少？(八分) [$F_{0.05}(9, 11) = 2.90$]

答：

$$1. H_0 : s_1^2 = s_2^2$$

$$H_1 : s_1^2 \neq s_2^2$$

$$F = \frac{5^2}{2^2} = 6.25, 6.25 > 3.59, \text{足證兩母體變異數不等。}$$

2. 由於前一小題得知兩母體變異數不等，因此在檢定兩平均數的差異時，使用t-test

$$t = \frac{34 - 31}{\sqrt{\frac{5^2}{10} + \frac{2^2}{12}}} = 1.782$$

$$df = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{\left[\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1} \right]} = \frac{(5^2/10 + 2^2/12)^2}{\frac{(5^2/10)^2}{10 - 1} + \frac{(2^2/12)^2}{12 - 1}} = 11.394 \cong 11$$