

# 《教育測驗與統計》

## 試題評析

今年高考教育行政的「教育測驗與統計」一科，罕見的出了五題，其中，前三題考的是統計學，共占55分；而後兩題考的是測驗學，共占45分。

第一題：出自本班教材的第七章(積差相關)，簡單的說明兩小題正/負相關之意義，並輔以圖示，10分要全拿，易如反掌！

第二題：出自本班教材第十三章(變異數分析)，在教材內容與考古題習作上，甚至是總複習班皆再三強調，尤其是填表的20分絕對沒有問題，交互作用的檢定結論也難不倒本班考生。

第三題：考的則是本班教材第五章(相對位置量數)與第六章(常態分配)，PR值的求算與常態直線轉換的圖示，也是平日再三練習的重點。

第四題：考的是效度的檢測方法，只要考生在一開始讀題時，就能抓到題目要求作答的方向與主軸，25分至少應可拿下20分。本題出自本班教材第十六章(效度)，難度不高。

第五題：是本次試題的黑馬，個人認為以第四次的比西量表修訂版作為考題似乎太過偏頗，相信沒有幾位考生可以順利作答。不過若能將本班教材第十四章之g與non-g(上課補充過)提出，仍可獲得基本分。

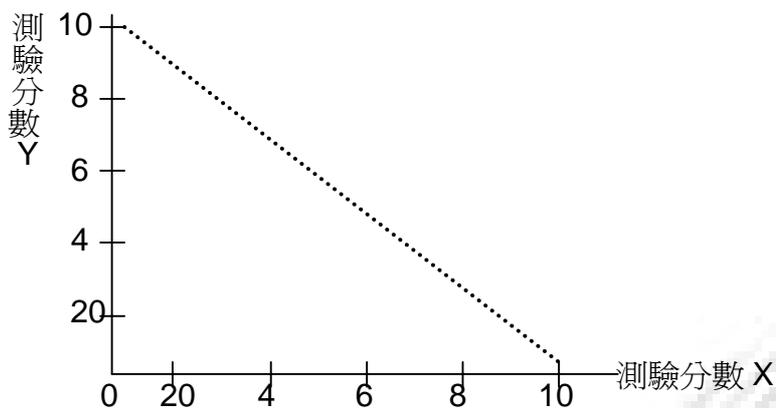
前三題的55分，本班絕大多數考生至少可囊括50分；綜合以上評論，此科本班考生應可有70分的基本成績，對於最後一題曾有涉獵者，還可有更高的成績表現。

一、以測驗分數X與測驗分數Y為座標之橫軸與縱軸，畫出二個散佈圖並分別說明X與Y的積差相關為-1，.85的意義。(10分)

**答：**

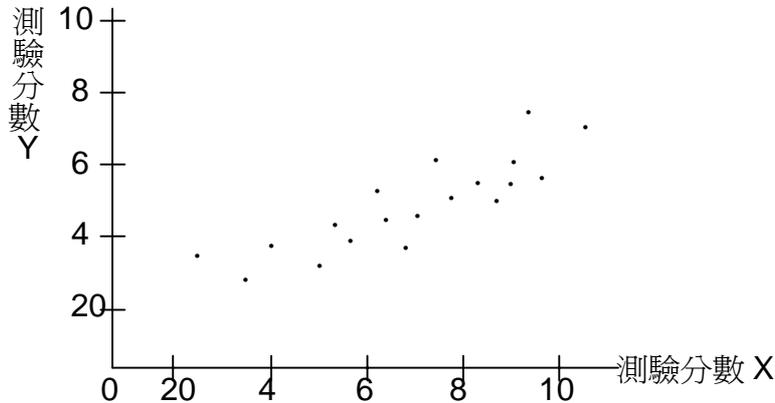
(一)X與Y的積差相關為-1的散佈圖如下：

其表示X、Y兩測驗分數呈完全或絕對負相關，而所有的每一個(x,y)對應點皆位於同一條直線上。



(二)X與Y的積差相關為0.85的散佈圖如下：

其表示X、Y兩測驗分數呈高度正相關，而所有的每一個(x,y)對應點呈現出由左下到右上的趨勢散佈，但不會位於同一條直線上。



二、某研究探討地區(A，分為北、中與南三個地區)與討論式、演講式、自修式三種教學法(B)對高中學生學習興趣的影響，每一細格各有10名不同的學生，共有90名學生參與此研究。請填寫以下變異數分析摘要表(每格2分，共20分)，並說明此研究中交互作用效果顯著代表的意義。(5分)(請將下表繪製於試卷上作答，於本試題作答者，不予計分。)

變異數分析摘要表

變異來源	SS	自由度(df)	均方(MS)	F值
A(地區)	( )	( )	3	1
B(教學法)	( )	2	( )	4
AB	( )	( )	( )	5*
Error	( )	( )	( )	
Total	333	89		

\* $p < .05$

**答：**

(一)ANOVA摘要表

變異來源	SS	自由度(df)	均方(MS)	F值
A(地區)	( 6 )	( 2 )	3	1
B(教學法)	( 24 )	2	( 12 )	4
AB	( 60 )	( 4 )	( 15 )	5*
Error	( 243 )	( 81 )	( 3 )	
Total	333	89		

\* $p < .05$

(二)針對地區(A)與教學法(B)兩個因子的交互作用(interaction effects)所進行的顯著性檢定，其虛無與對立假設為：

**H0:** 地區與教學法並不交互作用以影響高中生的學習興趣

**H1:** 地區與教學法確實交互作用以影響高中生的學習興趣

因其F值=5，且達0.05的顯著水準。其代表的意義為：由該90名參與研究的學生提供充分的統計證據，足以結論地區與教學法之間存在一影響高中學生學習興趣的交互作用。

- 三、如果某校學生的身高平均數為165公分，標準差為10公分，體重平均數為55公斤，標準差為5公斤。小張身高與體重是175公分與65公斤。
- (一)請以Z分數的概念來說明小張的身材在該校學生中是不是很正常？是身高還是體重比較不正常？(10分)
- (二)如果該校學生的身高與體重在分佈狀況呈常態，那麼小張的身高與體重的PR值為多少？請以常態分配的圖形來協助回答。(10分)

**答：**

(一)

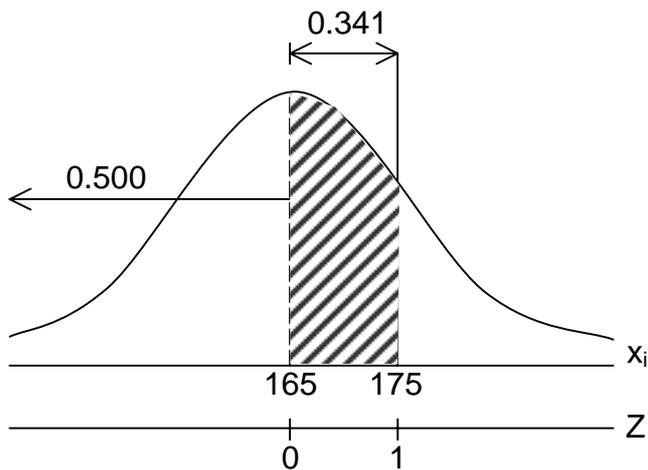
$$Z_{\text{身高}} = \frac{175 - 165}{10} = 1$$

$$Z_{\text{體重}} = \frac{65 - 55}{5} = 2$$

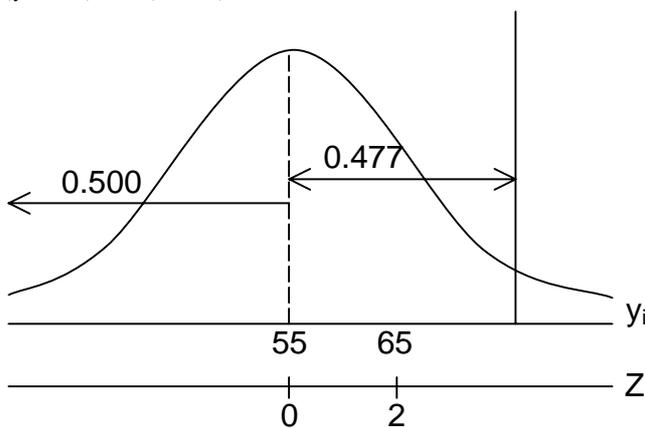
將小張的身高與體重進行上述的常態分配Z分數轉換後，可知小張的身材無論是身高或體重皆較該校學生的平均值來得高，其中又以其體重與該校平均體重的差異達兩倍的標準差，相對於其身高與平均身高的差異僅一倍標準差，表示小張的體重相對較不正常！

- (二)當該校學生的身高與體重皆呈常態分配，則依經驗法則，小張的身高(令其為X)與體重(令其為Y)之PR值分別為：

$$P(x \leq 175) = P(Z \leq 1) = 0.5 + 0.3413 = 0.8413 = 84.13\% \cong 84$$



$$P(y \leq 65) = P(Z \leq 2) = 0.5 + 0.4772 = 0.9772 = 97.72\% \cong 98$$



四、李芸編了一份測量同理心的量表，請問她可以如何檢驗這份量表所測的特質的確是同理心？請列舉方法並加說明。(25分)

**答：**

李芸要驗證其所編製的同理心量表是否確實在測同理心，也就是李芸在探討其編製測驗的效度(validity)，即該量表能否正確測量到每一位受測者同理心特質的高低。一般要檢驗量表或測驗的效度，依美國心理學會(APA)、AERA、NCME等教育團體所組成之聯席委員會所出版的標準，其建議應報告之效度有三，分別是內容效度、效標(關聯)效度、構念效度。

(一)內容效度：

指測驗內容的代表性或取樣的適切性，因取決於編製者的經驗與邏輯判斷，又可稱為邏輯效度。通常內容效度的證據屬於質化，李芸可以請該領域的幾位權威專家，針對其量表中的試題內容與組合提供意見，若多數專家共同認定其量表的品質，則可初步取得內容效度的證據；或是李芸可以準備雙向細目表，由測驗目標與測驗所依據教材內容兩個面向，清楚呈現其量表中題目之出處與分佈，以確保該測驗的內容效度。

(二)效標(關聯)效度：

李芸可以進行受試者團體在其量表所得之測驗分數，與另一適當的外在效標間之關聯性計算，由其獲得的相關係數即為效標效度。依測驗分數與效標分數是否在同一時間內發生或取得，又分為同時效度與預測效度兩種。效標效度可以呈現該同理心量表在一特定情境下，正確預測某一個體行為的有效程度。

(三)構念效度：

指該同理心量表能夠測量到理論上的構念或特質的程度，通常在測量前所未有的、或是以前測的不是很好的事物時特別注重。李芸為求得該同理心量表的構念效度，她可以採取的方法有：

1.內部一致性分析；

2.多項特質-多項方法分析(multi-trait multi-method matrix)

計算出收斂效度 (convergent validity)與區別效度 (discriminant validity)；

3.因素分析法

探索性因素分析 (exploratory factor analysis, EFA)在進行量表中題目與所測構念之正確歸屬與組合的分析，以確保題目測對其對應的構念以及題目的內部一致性；

4.實驗研究

5.團體差異分析

6.相關研究

五、1986年美國第四次修訂的比西量表 (Stanford-Binet Scale) 和以前的版本最大不同之一是明確地將Cattell的流體智力 (fluid intelligence) 與結晶智力 (crystallized intelligence) 的理論納入其測驗編製所依賴的架構。請說明流體智力與結晶智力的含意與其間的關係。(20分)

**答：**

(一)Spearman(1927)之智力二因論，將智力分為一般 (g) 與特殊 (s或non-g) 兩個要素。其中，g因素為個人的智力總能量，本質上以思考上的推理與演繹能力為主，為任何心智因素所必須的，是決定個人在智力測驗上表現的主要因素；而s因素與個別的特殊能力有關，例如空間關係或運動協調等，僅存在於個別的測驗中。

(二)Cattell(1963)則認為g因素智力又可分為流體智力和結晶智力。

(三)流體智力是對於非語文性之訊息間的關係，能夠加以洞察和推理的思考能力，主要來自遺傳，與大腦的功能有關，因而此種智力與個人的學習能力和速度有高度的正相關，但不受環境經驗或文化背景之影響；

(四)結晶智力則是經由學習和經驗而逐漸累積得來的智力，為個人智識和專門性技能之總體。也是流體智力與文化知識混合的結果，是後天經驗的產物，包括：常識、語文瞭解、數學能力與應付社會情況的能力等。

(五)一般智力測驗所測量者以結晶智力為多。