

《教育測驗與統計》

- 一、某位教育學者自編一份數學成就測驗，試圖用來預測學生在「國中會考」的數學科成績。但該學者在發展測驗的過程中，於進行效標關連效度分析時發現，他的自編測驗缺乏信度。請問：你會建議他該如何改善自編測驗缺乏信度的問題？（25分）

試題評析	本題屬於發揮應用題，以教師自編測驗為主軸，考生必須找到試題分析時信度改善的可能方法加以系統陳述才是得分的關鍵。
考點命中	1.《高點·高上教育測驗與統計講義》第三回，傅立葉編撰，Ch.16，頁28；Ch.17，頁36。 2.《高點·高上地方特考「教育測驗與統計」重點題神》，傅立葉編撰，頁2。

答：

某位教育學者自編一份數學成就測驗，目的在於預測學生在國中會考的數學科成績，由於測驗發展過程中，在效標關聯效度分析時發現其測驗缺乏信度，建議其改善信度問題的做法如下：

(一)針對以下問題仔細檢視與分析：

- 1.該測驗目標與教學目標的一致性為何？
- 2.欲評量的知識、技能、情意目標為何？
- 3.界定的教學目標是否具體明確？
- 4.測驗編製時所擬定的雙向細目表完整性為何？
- 5.測驗題目能否測得所欲評量的教學目標？
- 8.測驗題目類型能否測得所欲評量的教學目標？

(二)由於是在效標關聯效度分析時發現信度不足問題，基於其目的在預測未來的國中會考表現，建議應從效標的選擇上努力找到該自編測驗成績之適當合理的比較基準與對象。

如此回饋至測驗編製的前面步驟，應能有效透過後續的試題修編、刪除或替換，並進行相關的信效度分析而獲致較佳測驗的成果！

- 二、請根據下列試題分析結果的局部摘要數據，回答下列問題並說明其原因：（每小題5分，共25分）

學生注意係數表

學生	得分	答對率	注意指標	判定類別
A	54	67.50	0.69*	B'
B	61	76.25	0.24	A
C	44	55.00	0.27	B
D	73	91.25	0.55*	A'

試題注意係數表

試題	答對人數	答對率	刪題後信度	注意指標	判定類別
1	16	34.78	0.76	0.90**	B'
2	40	86.96	0.74	0.81**	A'
3	24	52.17	0.73	0.43	A
4	14	30.43	0.72	0.37	B

平均答對人數：30.25，受試者總人數（N）：46.00

平均答對率（P）：0.76，試題總題數（K）：80.00

內部一致性係數：0.75，平均點二系列相關係數：0.45，差異係數：0.65

(一)那一位學生是屬於「粗心大意型」的學生？

(二)那一題是屬於需要大幅修改甚至是刪除的試題？

- (三)此次試題分析結果，顯示本測驗的作答難易度為何？
 (四)此次測驗結果，是否反映出學生作答能力的真實情況？
 (五)此次測驗結果，是否反映出試題編製的品質良好？

試題評析	本題以各種指標數據的提供進行試題分析。考生除了必須清楚信效度與難度、鑑別度指標外，仍須利用S-P表的注意係數回答部分子題。本題為今年整份試卷一分高下的關鍵題目。
考點命中	1.《高點·高上教育測驗與統計講義》第三回，傅立葉編撰，Ch.15，頁13；Ch.18，頁38~39。 2.《高點·高上地方特考「教育測驗與統計」重點題神》，傅立葉編撰，頁5；最新之議題重點整理，(二)推論統計，3.相關係數，頁15~16；最熱門之議題重點整理，(二)推論統計，4.難度與鑑別度分析。

答：

注意係數(caution index)是S-P表針對「個別」學生與試題所使用的一類係數，又可分為學生注意係數與試題注意係數兩種。通常用來作為判斷學生或試題在反應組型中是否有異常現象的指標，教師可利用這些指標瞭解學生或試題的狀況與問題所在。重要的判斷依據如下：

- 1.當學生或試題注意係數介於0~0.5之間，表示該試題或學生反應組型的不尋常情況並不嚴重，屬正常程度。
- 2.當學生或試題注意係數介於0.5~0.75之間，表示不尋常已是嚴重狀況，教師應該加以注意。
- 3.當學生或試題注意係數大於0.75時，表示不尋常情況已是非常嚴重，教師應更加特別注意。

依題意茲分別回答各子題如下：

(一)基於91.25%答對率，0.55的注意指標，以及A'的判定類別，學生D屬於粗心大意型。

(二)需要大幅修改甚至刪除的試題是第一題，理由是：

- 1.答對率34.78%偏低，表示難度太高；
- 2.刪除後剩餘三題的信度為最高的0.76；
- 3.注意指標0.90顯示非常嚴重的不尋常情況；
- 4.被判定為B'的試題類型。

(三)本次測驗的作答難易度應以0.76的平均答對率(P)為根據，有高達76%的學生答對表示難易度偏低。

(四)由於題目難度偏易，儘管部分試題仍待改善，大致反映學生作答能力的真實狀況。

(五)依測驗結果，試題編製的品質是否良好可以由相關資訊取得：

- 1.內部一致性係數= 0.75，此為高信度的證據。
- 2.平均點二系列相關係數= 0.45，此為普通鑑別度的證據。
- 3.平均答對率= 0.76，此為低難度的證據。
- 4.差異係數= 0.65，大於0.6的數值表示試題的異質性存在。

因此結論：試題編製品質還稱不上良好，仍有改善空間！

三、某教育統計學者擬進行以「成就動機」(X變項)預測「學業成績」(Y變項)的迴歸分析研究，但因為熬夜工作，頭昏眼花，錯把初步獲得的標準化迴歸公式寫成下列的式子：

$$\text{成就動機} = 0.5 \times \text{學業成績}$$

請回答下列問題：

- (一)他原本打算獲得的標準化迴歸公式應該表示為何？(5分)
- (二)成就動機可以預測到學業成績多少百分比的變異量？(5分)
- (三)事後他仔細一想，若改成以「學業成績」預測「成就動機」的結果才是合理的話，則該預測公式的決定係數為何？(5分)
- (四)經過這兩次的測試分析，他終於發現以「成就動機」預測「學業成績」的標準化迴歸係數，和以「學業成績」預測「成就動機」的標準化迴歸係數，彼此間有何不同？請說明之。(10分)

試題評析	本題涉及標準化迴歸分析的係數與決定係數之簡單關係與計算。知道自變項與依變項互為因果下的兩個迴歸模式其標準化迴歸係數的乘積正是決定係數的特性，將有助於回答各子題。
------	--

考點命中	1.《高點·高上教育測驗與統計講義》第一回，傅立葉編撰，Ch.8，頁54。 2.《高點·高上地方特考「教育測驗與統計」重點題神》，傅立葉編撰，頁13，最熱門之議題重點整理。
-------------	---

答：

對於該教育學者以成就動機預測學業成績的迴歸分析研究，依題意茲分別回答如下：

(一)他原本打算獲得的標準化迴歸公式應為：

$$\text{學業成績} = 0.5 \text{成就動機}$$

(二)利用前述(一)標準化迴歸公式中的係數0.5，取平方即得決定係數為0.25，因此，成就動機可以預測學業成績25%的變異量。

(三)當事後該學者覺得以學業成績預測成就動機才是合理，基於自變項與依變項互為因果的兩個標準化迴歸係數的乘積為決定係數，則該預測公式的決定係數仍為0.25(25%)。

(四)經過這兩次的測試分析，由於題目提供資訊不足，只得基於自變項與依變項互為因果的兩個標準化迴歸係數的乘積為決定係數的特性，斷定兩個迴歸模式的標準化係數皆為0.5而並無不同。

四、某教育學者根據某城市小學六年級國語科基本能力檢測結果，進行性別差異的檢定分析，獲得如下表的結果：

性別	人數	平均數	標準差	t值	p
女生	8398	528.38	77.35	27.55	.000**
男生	7731	516.23	78.98		

請回答下列的問題：

(一)那一種性別的個別差異較大？(3分)

(二)那一種性別的檢測成績較佳？(3分)兩種性別差多少分？(3分)

(三)該教育學者係使用何種統計檢定方法？(6分)

(四)請根據上表數據，解釋該城市小學六年級國語科基本能力檢測結果，是否具有性別差異存在？請說明之。(10分)

試題評析	本題命題主旨為兩個獨立母體的平均國語科基本能力差異之t檢定。命題沒有讓考生有計算的表現機會，反而考的是解讀檢定報表的能力，本班考生理應能夠輕鬆應付！
考點命中	1.《高點·高上教育測驗與統計講義》第二回，傅立葉編撰，Ch.11，頁23~25。 2.《高點·高上地方特考「教育測驗與統計」重點題神》，傅立葉編撰，頁14~15，最熱門之議題重點整理，(二)推論統計，2.各種樣本統計量的抽樣分配。

答：

該教育學者針對某城市小六國語科基本能力檢測結果，進行性別差異的檢定分析，依題意茲分別回答如下：

(一)因為從檢定分析結果所得之表格中得知，男生測驗分數的標準差78.98大於女生的77.35，男生的個別差異較大。

(二)女生的測驗平均數528.38高於男生的516.23，因此男生的檢測成績較佳。兩種性別的差異分數為：

$$528.38 - 516.23 = 12.15(\text{分})$$

(三)該教育學者使用的統計檢定方法為兩個獨立母體的平均數差異t檢定。

(四)該城市小六國語科基本能力檢測結果，是否具有性別差異存在的檢定，其假設為：

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

對於t檢定的決策法則與結論可以利用表格中的p-值檢定加以取代。由於.000的p-值表示在任何顯著水準下皆得以拒絕虛無假設，因此，利用該城市小六男女生國語科基本能力檢測結果的樣本資訊，足以提供充分證據結論小六國語科的基本能力確實具有男女性別的差異存在，檢定達顯著。

【版權所有，重製必究！】