

《教育測驗與統計概要》

- 一、測驗編製及實施時，信度是重要的考量，請針對穩定係數、延宕複本信度、內部一致性(α)、以及評分者信度等信度估計方法來分析各種方法可能的誤差來源。此外，請說明在比較不同評量工具的信度係數時，需考量的因素有那些？(25分)

試題評析	本題主考測驗學的信度指標，是不可能被考生掠過忽視的重點命題範圍。往年也考過類似題目，答分高下在於第二小題的作答內容與陳述。
考點命中	《高上教育測驗與統計講義》第三回，傅立葉編撰，第十五章，頁14。

答：

- (一)針對題目所提及的穩定係數、延宕複本係數、內部一致性(α)與評分者信度的誤差來源可以整理說明如下：

信度種類	誤差來源
穩定係數	時間取樣
延宕複本係數	時間與內容取樣
內部一致性(α)	內容取樣與內容異質性
評分者信度	評分者誤差

- 1.穩定係數：同一份測驗受在兩次施測之間隔時間長短期間的身心條件與因素變化之影響。
- 2.延宕複本係數：兩個複本在不同時間施測，所可能受複本內容與時間因素的影響。
- 3.內部一致性(α)：受到組成題項與題項之間同質性高低的影響。
- 4.評分信度：受不同評分者主觀判斷差異的影響。

- (二)在比較不同評量工具的信度係數時，須考量的因素有以下：

- 1.不同測驗目的的測驗，所需要的信度資料不同：
就性向測驗而言，最重要的信度資料是穩定性數，因為性向測驗的結果主要是用來做未來長時間的預測。就成就測驗而言，需要的信度資料是複本信度，因為要依據測驗結果推論個人精熟的重要知識與技能的程度。
- 2.不同類型的測驗，所需要的信度資料不同：
常模參照測驗與校標參照測驗使用的信度指標就不同，後者常用的Kappa係數與百分比一致性與常模參照不同。
- 3.信度係數的取得所基於常模樣本的差異性：
不少目的相同，但是基於受試者常模樣本就性別、年齡或其他特質的差異，在比較時不得不注意。

- 二、和平國小吳老師編製了一份課堂數學評量，題型包含客觀及紙筆實作試題兩類。

(一)請從內容(content)、歷程(process)、後果(consequences)三個面向，說明吳老師應如何建立該份測驗結果的效度。(15分)

(二)「構念代表性不足」(construct-underrepresentation)及「無關構念」(construct-irrelevant)是影響評量結果的兩大威脅，請解釋其意義及影響，並具體說明那些因素會造成此現象。(10分)

試題評析	本題涉及效度的核心概念-構念效度以及內容效度與效標效度等。創新的以內容、歷程、後果要求考生回答自編數學評量的效度探討。至於，構念代表性不足和與無關構念的定義與影響則為效度不足與效標汙染為同義詞，難不倒實力好的考生。卻是今年在同類科分出高下的最重要題目。
考點命中	《高上教育測驗與統計講義》第三回，傅立葉編撰，第十六章，頁24-25。

答：

【版權所有，重製必究！】

(一)由於吳老師所編製的課堂數學評量屬於教師自編測驗，雖然屬於非正式測驗，似乎不須探討與建立效度證據，但是，基於試題分為客觀與紙筆實作兩類題型，該評量的效度若依內容、歷程與後果三個面向加以探討與建立其效度，則可考慮的作法如下：

1.內容：

就客觀題型而言，可以利用雙向細目表就教學目標與教材內容兩個構面進行試題出處與組合的適切性探討，以便取得題數與所占百分比的基本資訊並加以合理調整；而在紙筆實作類型的題目上，評量內容要與教學目標一致，數學的教學目標主要涵蓋的內容包括：科學方法的運用、科學精神的培養及科學知識的獲得、以及提昇其處理問題的能力，因此評量的內容項目應包括這些。

2.歷程：

客觀測驗題目應經過預試(tryout或pre-test)以進行效標關聯效度的計算與探討；而紙筆實作評量部分則需要準備操作的器材，並清楚界定主題、所需時間與活動過程，以確保測驗結果的效度。

3.後果：

這指的是測驗結果的評分、解讀、分析與應用。就客觀測驗部分，應該進行測驗結果基於常模的相對表現好壞的對照與解釋；而在紙筆實作評量中，由於操作計畫的書寫、登錄資料的表格設計、研判的結論等等都是用文字來描述，評閱時可以由作答的文字敘述情形判斷學生作答的心路歷程，因此，可獲得許多資訊。但是，對於這種「開放性回答」的評分需要事先設定評分標準或登錄答案的方式，以建立同一評分標準。

(二)測驗的效度探討，雖然有幾個效度指標可供進行深入探討，但是避免構念代表性不足和與無關構念的變異是構念效度的重要資訊，當然是了解評量效度的主要努力方向，也是影響評量結果的兩大威脅。其中，在實作評量中，若無法評量學生在真實情境下應用知識與技能的能力，則屬評量工具本身構念代表性不足的問題，需進行題目或問題的重新調整處理。無關構念則指問題混雜了非目標特質構念的題項內容，當然影響評量結果的效度。其實，構念代表性不足也就是效度不足，題目無法準確測量到所要理解的數學能力與資訊；而無關構念正是效標污染的概念，評量放入不合格、不好的問題，也就無法準確測量目標能力。至於，造成構念代表性不足現象的因素為，對於測量目標能力定義不清，導致試題組合結構比例不當，沒有適當地選擇某些測驗內容，或採用某些心理活動程序，或引導出測量該構念所需要的反應；而造成無關構念的因素為受到構念以外因素的系統性影響。與構念無關因素可能包括：對於測驗內容上的情緒反應、對於數學學習材料的熟悉程度、或是回答問題時所需要的寫作技巧等。

三、研究員欲了解某一課程補救教學介入程度(X, 1至4, 數字愈大表愈強)對學習成績(Y)的影響。16名學生隨機分成相同人數的四組，並接受不同程度的補救教學介入，課程結束後，16名學生學習成績平均數 \bar{Y} 為68.75，而 $\sum_{i=1}^{16} (Y_i - \bar{Y})^2 = 4,453$ 。試回答下列問題：

(一)請利用簡單線性迴歸分析及單因子變異數分析方法，陳述研究問題與對應的研究假設(H_0 & H_1)。(10分)

(二)已知X與Y的相關係數為0.831，而單因子變異數分析表中，組間的平方和為3,237，計算兩分析方法中X與Y的關聯強度，並說明差異的原因。(10分)

試題評析	本題考的是因果關係推論可以使用的兩種最基本統計分析方法。研究問題與研究假設相當容易，兩種方法的關聯強度也是相當好的命題。答題與得分關鍵在於關聯強度的計算與兩者數值間差異的說明。
考點命中	1.《高上教育測驗與統計講義》第一回，傅立葉編撰，第八章，頁52。 2.《高上教育測驗與統計講義》第二回，傅立葉編撰，第十三章，頁48。

答：

(一)欲研究某一課程之補救教學介入程度(X)對學習成績(Y)的影響，其屬於因果關係的推論。因果關係推論的兩種統計分析方法為簡單線性迴歸分析與單因子變異數分析，兩者的差別在於自變項(X)的操作為連續變項(1-4的涉入程度)或是間斷變項(實驗處理組別)。以下分別回答兩種可採用統計方法下的研究問題與研究假設：

【版權所有，重製必究！】

1.簡單線性迴歸分析：

研究問題：補救教學介入程度(X)對學習成績(Y)的高低是否存在顯著影響？

研究假設： H_0 ：補救教學介入程度對於學習成績不存在顯著影響(回歸模式配適不佳)

H_1 ：補救教學介入程度對於學習成績存在顯著影響(回歸模式配適良好)

2.單因子變異數分析：

研究問題：不同補救教學介入程度(X)組別的學習成績(Y)是否存在顯著差異？

研究假設： H_0 ：不同補救教學介入程度組別的平均學習成績無顯著差異

H_1 ：不同補救教學介入程度組別的平均學習成績存在顯著差異

(二)兩種統計分析法下兩變項之間的關聯強度如下：

1.簡單線性迴歸分析：

依據X與Y的相關係數0.831，可以計算決定係數如下：

$$R^2 = (r)^2 = (0.831)^2 = 0.690561 \cong 69.06\%$$

2.單因子變異數分析：

利用題目提供的組間平方和與總平方和的數據資料與變異數分析摘要表的計算關係，可計算關聯強度如下：

$$\hat{\omega}^2 = \frac{SS_b - (k-1)MS_w}{SS_t + MS_w} = \frac{3237 - (4-1)(101.33)}{4453 + 101.33} = \frac{2933.01}{4554.33} = 0.644 = 64.4\%$$

3.兩個關聯強度存在差異的原因在於單因子變異數分析透過簡單隨機分派的方式，將存在於16名受試者之間的誤差透過實驗處理加以分解，但是在組數為相對稍多的4組而實驗對象為稍嫌太少的16位之條件下，使得64.4%的關聯強度不如簡單線性迴歸分析的69.06%。

四、下圖是61位學生在某門課上成績的莖葉圖，請根據圖回答下列問題（回答下列問題時，請將男女生的得分一併考慮）：（30分）

男生	莖	女生
	2	1
6	3	
	4	
	5	
2	6	03
8876655320	7	0145566899
8766543221	8	0112334566677889
42	9	0023356

(一)成績的次數分配圖會呈現何種型態？眾數和中位數各是多少？全距為何？

(二)已知整組資料的標準差為12.45，請問變異係數 (coefficient of variation) 是多少？其使用時機為何？

試題評析	本題的命題屬於敘述統計的次數分配、統計圖表與統計量數範圍。難度不高，惟若答題時太快下筆，而未能注意題幹中特別說明應以男女生得分合併考慮，則將犯下大錯。題目雖偏容易，卻是來自本班平日授課講義的幾個大章，屬於綜合性考題。
考點命中	1.《高上教育測驗與統計講義》第一回，傅立葉編撰，第二章，頁6-7。 2.《高上教育測驗與統計講義》第一回，傅立葉編撰，第三章，頁13。 3.《高上教育測驗與統計講義》第一回，傅立葉編撰，第四章，頁18。 4.《高上高普考重點題神》教育測驗與統計，重點十一、枝葉圖。 5.《高上教育測驗與統計總複習講義》，傅立葉編撰，肆、重點摘要，四.解釋名詞之13.(變異係數)。

【版權所有，重製必究！】

答：

- (一)依題目所提供的男女生在某課程成績的莖葉圖，因為分數群聚於高分處，所以次數分配圖將呈現左偏(或稱負偏)的分配型態。至於，成績的中位數與眾數因為題幹要求男女生合併考慮，因此，需要將男女生的成績重新依高低排序後進行回答。排序後，61名學生成績的眾數是出現最多的5次的86分，而中位數是高低排序位於正中間之第31位學生的成績82分。最後，全距是最高分與最低分之間的差值，所以96分減21分得75分。
- (二)1.要計算整組分數資料的變異係數(CV)需要平均數與標準差兩個統計量，在已知標準差為12.45的條件下，先行計算平均成績得79.754，因此，變異係數計算如下：
 $CV = 12.45/79.754 = 0.1561 = 15.61\%$ 。
- 2.變異係數為一種變異量數，其使用時機有：
- (1)比較說明性質不同的兩個變數資料的變異大小時，例如：成年男性體重與男童體重的變異；
 - (2)比較說明單位不同的兩個變數資料的離散情形時，例如：身高(公分)與體重(公斤)的離散；
 - (3)比較說明單位相同，平均數差異過大的兩個變數資料的變異大小時。

高
點
·
高
上

【版權所有，重製必究！】