

《教育測驗與統計》

一、何謂標準化測驗？（5分）舉兩個具體例子，對照說明標準化與非標準化測驗的不同適用場合及其主要理由。（15分）

試題評析	本題主考測驗學的標準化測驗，並以其與非標準化測驗的差異要求考生舉例說明其適用場合與理由。難度不高，本班生與一般有準備的考生應該都能掌握好分數。
考點命中	1.《高上教育測驗與統計講義》第三回，傅立葉編撰，第十四章，頁6。 2.《高上高普考重點題神》教育測驗與統計，重點三，頁3-4。

答：

- (一)標準化測驗的定義是：在標準化的情境下，測量個人目標能力或心理特質的工具或歷程。其使用原則含下列四點：
- 1.標準化測驗之實施必須顧及施測步驟、計分及結果解釋的一致性；
 - 2.在施測步驟上，舉凡指導語、時間限制、物理環境和受試者的身心特質都應力求標準化；
 - 3.測驗的計分要力求客觀與正確，結果解釋則需同時考量個別受試者的身心特性，測驗結果才能真實的反應個案的能力；
 - 4.必須挑選能適當評量個別受試者能力的測驗，因此須先詳讀測驗手冊之內容（包括適用對象、測驗目的、常模與信效度、測驗題型等）。
- (二)有關標準化與非標準化測驗之間的相異在於，標準化測驗乃依指導手冊之編製程序編成的測驗，使用常模表決定分數之等級且建立常模、信度、效度等資料，常又稱為正式測驗。大多數的智力與性向測驗皆屬此類測驗。非標準化測驗則為非正式的依個人教學需要所編製的測驗，實施較具彈性，通常也不建立常模，或無建立常模的必要，又稱非正式測驗，大多數成就測驗屬之。基於此，升學測驗多屬標準化測驗，因為統一的題本與施測程序規範，並使用相對位置量數，如百分等級進行施測結果的解釋與分派安置作業；而課堂上的隨堂測驗則屬非標準化測驗，因為命題內容、題數、難度、施測環境與物理條件的控制都不特殊規定，施測後也不一定審慎評分，純為掌握學生學習進度與調整教學進度或難度之用。

二、王校長在新學期進行閱讀教學實驗，他選用國語文能力測驗甲、乙兩式測驗做為前、後測工具，依據測驗手冊資訊，兩式測驗平均數都是50，標準差都是10，兩測驗的相關為0.84。試回答下列問題：（一）以複本信度估計這兩測驗的測量標準誤為何？（5分）（二）王校長在實驗結束後發現實驗效果達顯著水準，但對實驗組學生而言，這兩式測驗的相關只有0.5，你認為在實驗有效情況下，實驗組學生前、後測成績相關低於常模的資料，這個現象是否合理？為什麼？（10分）（三）王校長希望頒發最佳進步獎給前、後測進步最多的10位小朋友，對於後測減前測這項進步量，其誤差變異數會不會小於前測的測驗誤差變異數？為什麼？（10分）

試題評析	本題命題涉及測量標準誤、前後測差值誤差變異數以及單一母體的誤差變異數為命題主軸。難度適中，唯一的挑戰是第三小題，重點在於考生對題目意義的閱讀理解能力。
考點命中	1.《高上教育測驗與統計講義》第三回，傅立葉編撰，第十五章，頁15。 2.《高上高普考重點題神》教育測驗與統計，重點十六，頁15。

答：

(一)依據題意，以複本信度估算測驗的測量標準誤可以計算得：

$$SE_{meas} = 10\sqrt{1-0.84} = 4$$

- (二)實驗結束後，實驗效果達顯著，但是兩測驗之間0.05的相關卻低於常模的0.84，在實驗有效的情況下，此一現象是合理的，可能原因是王校長所進行閱讀教學實驗的學校學生，因屬於偏鄉地區或其他學習能力與意願等特殊條件，使得該校學生之國語文能力遠低於常模的平均水準，而導致前後測分數的相關較低。
- (三)而欲頒獎給進步最多的前10名學生，以後測減去前測成績作為進步量，其誤差變異數可以依以下計算公式

得出：

$$\sqrt{\frac{s_1^2 + s_2^2 - 2rs_1s_2}{n}} = \sqrt{\frac{10^2 + 10^2 - 2(0.5)10^2}{n}} = \sqrt{\frac{10^2}{n}}$$

而前測成績的誤差變異數可以計算如：

$$\sqrt{\frac{s_1^2}{n}} = \sqrt{\frac{10^2}{n}}$$

如此比較看來，後測減去前測成績之進步量的誤差變異數與前測成績的誤差變異數相同。

三、下表為52位國一學生在一項實作測驗上的分組次數統計。原始分數之 $\sum X$ 為130， $\sum X^2$ 為385.18。請根據表中資料回答問題：

分數	次數	累積次數
3.6-4.0	7	52
3.1-3.5	9	45
2.6-3.0	10	36
2.1-2.5	9	26
1.6-2.0	6	17
1.1-1.5	5	11
0.6-1.0	3	6
0.1-0.5	3	3

- (一)將成績以累積次數多邊圖表示，並說明如何應用此圖？(6分)
- (二)以原始分數之資料求平均數與變異數；以分組資料求中位數、眾數及四分差(請採真正上、下限解題)。根據所得資料，何種量數最能反映數據之集中與離散的情形？(12分)
- (三)研究者在事後更正一位學生的成績，由3.6變更為4.0。何種集中與離散量數會改變？何者不會改變？(4分)
- (四)為確保分數的精確性，另請一位評分者評量學生表現，並取兩者的平均分數為學生的最終分數。第二位評分者給分的平均數為2.0，變異數為1.2，兩位評分者給分的相關為0.80。請問學生最終分數的變異數是多少？(3分)

試題評析	本題考的是分組資料下幾個重要集中與離散量數的計算。分組資料的統計量數估算通常因為有不同方式的選擇，以及相對於原始資料常有誤差，多年來不在高普考中出現考題。今年的命題出乎意料之外，算是冷門大翻身的題目。預估少有同學在此看似容易的題目取得高分。
考點命中	1.《高上教育測驗與統計講義》第一回，傅立葉編撰，第三章，頁13。 2.《高上教育測驗與統計講義》第一回，傅立葉編撰，第四章，頁18。

答：

- (一)(圖形，略)累積次數多邊形圖是將這52位國一學生在該實作測驗分數的分組次數，從低而高依序累加人數而得累積次數後，利用橫軸為分數、縱軸為累積次數繪製而成的統計圖表。將圖中因累積次數所呈現越墊越高的條狀物之頂端中點連結而成的圖形，可以提供分數組距與累積人數的資訊，在大型考試結果受測者分數的統計與排序特別有用。
- (二)從原始分數所求的平均數與變異數，分別計算如下：

$$\mu = \frac{130}{52} = 2.5$$

$$\sigma^2 = \frac{385.18 - (130^2 / 52)}{52} = 1.157$$

利用分組資料求中位數、眾數與四分差，分別計算如下：

$$Md = l + \left(\frac{\frac{N}{2} - F}{f_{Md}} \right) h = 2.55 + \left(\frac{\frac{52}{2} - 26}{10} \right) (0.5) = 2.55$$

$$Mo = l + \left(\frac{f_a}{f_b + f_a} \right) h = 2.55 + \left(\frac{9}{9+9} \right) (0.5) = 2.80$$

52(1/4) = 13, 52(3/4) = 39, 因此, Q1與Q3分別落在1.6~2.0以及3.1~3.5兩組中
接著以等差法分別計算Q1與Q3, 得

$$\frac{Q1-1.6}{2.0-1.6} = \frac{13-11}{17-11}, Q1 = 1.73$$

$$\frac{Q3-3.1}{3.5-3.1} = \frac{39-36}{45-36}, Q3 = 3.23$$

$$QD = 3.23 - 1.73 = 1.5$$

根據計算結果, 因為實作測驗得分呈現負偏分配, 並不存在最能反映數據之集中與離散情形的量數。

(三)從3.6分更正其中一名學生的分數為最高的4.0分, 會造成平均數的改變(稍微提高)以及變異數的改變(稍微擴大)外, 其餘如中位數、眾數的集中與四分差的變異量數皆不會改變。

(四)因為取兩位評分者的平均分數為最終分數, 欲求最終分數的變異數, 可以計算如下:

$$\sigma_y^2 = \frac{\sum y_i^2 - 2^2(52)}{52} = 1.2, \sum y_i^2 = 270.4$$

$$\sigma^2 = \frac{385.18 + 270.4 - 52 \left(\frac{2.5+2}{2} \right)^2}{52} = 7.545$$

四、某教育研究者想了解不同師資培育背景教師的教學專業知能有無差異, 隨機挑選了師專、師院、師資班三種師培背景的小學教師共150名接受教育專業知能的測驗, 測驗結果如下:

	師專	師院	師資班
N	50	50	50
\bar{X}	46.6	42.5	43.6
s^2	9.3	11.3	8.6

針對上述測驗結果, 研究者懷疑三種師培背景教師的教育專業知能是否相同。

(一)針對研究者所關心的問題來看, 你會建議採用何種統計方法回答上述問題?(5分)

(二)在進行上述統計方法之前, 須符合那項統計假設(assumption)?(5分)

(三)請寫出(一)統計方法之虛無假設與對立假設。(8分)

(四)針對上述資料, 研究者同時關心各組資料的變異數是否相等, 研究者決定使用Bartlett與Cochran二法進行檢定, 試問二法的檢驗分別為何?(8分)結果有無不同($\alpha = 0.05$)?(4分)

試題評析	本題的命題圍繞於單因子變異數分析的重點假設條件與推論統計基礎內容。前三小題的難度不高, 一般考生在答題上應無太大障礙; 但是第四小題對於變異數同質性檢定的檢測方法, 卻提到寇克蘭檢定, 該檢定常屬於非母數統計法, 似乎在此有故弄玄虛或命題錯誤之嫌。當場會造成考生很大的壓力, 不過, 大家在相等的測驗條件下, 不應受其影響。
考點命中	《高上教育測驗與統計講義》第二回, 傅立葉編撰, 第十三章, 頁47-55。

答:

(一)依題目針對研究者所關心的問題, 由於「師資培育背景」屬於類別的間斷型變項, 而「教育專業知能」的

測驗結果屬於連續型變項，三種不同師資培育背景的教育專業知能是否不同的因果推論，建議採用單因子變異數分析。

- (二)在進行單因子變異數分析之前，須符合的統計假設不只一項，其中，最重要的是變異數同質性的假設。也就是各組變異數應具有齊一性的假設，以確保變異數分析結果的說服力。
- (三)依題意，在單因子變異數分析的統計方法下，虛無假設與對立假設分別為：

H_0 ：不同師資培育背景的教育專業知能測驗平均成績無顯著差異

H_1 ：不同師資培育背景的教育專業知能測驗平均成績不全等

- (四)對於各組資料的變異數是否相等的問題，研究者採用Bartlett與Cochran兩種檢定方法，兩種方法的檢定差異為，當資料分佈是常態時，Bartlett法是很好的方法，但是它對分佈型態很敏感；也就是說，它不是很Robust，當常態的假設不成立，它的精確度就值得商榷。另外，精確的拒絕域已有專表可查，而卡方乃是近似解。Cochran檢定在此應是指當違反變異數同質性的假設時，必須利用Cochran & Cox所發展的T考驗法。應無比較的意義與需要！

高
點
·
高
上

【版權所有，重製必究！】