

103 年特種考試地方政府公務人員考試試題

等 別：四等考試

類 科：教育行政

科 目：教育測驗與統計概要

一、學生能力程度類似的甲、乙兩班，因為授課教師不同，所用試卷亦不同，學期末的數學期末考試成績之描述資料如下：甲班平均數為 80 分，中位數 84 分，標準差 5 分；乙班的平均數 65 分，中位數 60 分，標準差 4 分。請根據這些數據回答下列問題：

(一)甲班林生得 85 分，乙班李生得 73 分，如果將林、李兩生的成績化為 T 分數，則誰的 T 分數較高？兩班平均數差異如此大，不以原始分數比較而以 T 分數來比較，主要是建立在何種前提之下？

(二)若乙班張生實際得分為 66 分，後來老師發現誤將張生的成績登錄為 61 分，試問修正過後乙班的平均數與中位數受何影響？

(三)若甲班老師將每位同學的分數加 5 分，則平均數與中位數有何種變化？

(四)若甲班林生(得分 85 分)原本的百分等級為 63，則加 5 分之後，其百分等級有何變化？

(五)由所提供的數據來判斷，甲班和乙班分數之偏態有何差異？

【擬答】：

$$(一)甲班林生 Z_1 = \frac{85-80}{5} = 1 \Rightarrow T = 10Z + 50 = 10 + 50 = 60$$

$$乙班李生 Z_2 = \frac{73-65}{4} = 2 \Rightarrow T = 10Z + 50 = 10 \times 2 + 50 = 70$$

(1)∴ 乙班李生 T 分數較高

(2)甲、乙二班試題難易度不同，所以原始分數不能比較大小，應轉化成標準分數才可比較大小

(二)平均數會變大，中位數不變

(三)平均數及中位數會各加 5 分，即新的平均數為 $80+5=85$ ，中位數 $84+5=89$

(四)百分等級不變，即 PR=63

(五)甲班： $\mu < Me \Rightarrow$ 圖形為左偏

乙班： $\mu > Me \Rightarrow$ 圖形為右偏

二、當我們解釋一個重要的測驗結果(例如：智力測驗得分)時，應注意那些事項？

【擬答】：

(郭生玉著：教育測驗與評量 P162-P166)

(1)解釋測驗者應瞭解測驗的性質與功能

(2)測驗分數應為學生保密

(3)解釋分數應參考其他有關資料

(4)解釋分數應避免只給數字

(5)對低分者的解釋應謹慎小心

(6)解釋分數時應設法瞭解學生的感受

(7)解釋分數只做建議勿做決定

(8)採用一般可信範圍解釋

三、某輔導老師研究其針對國小高年級學生所設計的性別知識輔導方案是否能有效提高學生的性別知識。輔導老師以前測及後測的方式，收集了 32 名學生輔導前及輔導後的性別知識表現。

(一)若輔導老師以相依樣本 T 檢定來檢定其問題。說明其所收集的資料需符合那些特性，以確保使用 T 檢定的適當性。

(二)若輔導老師得到性別知識前後測進步分數的 95%信賴區間介於 3.2 分與 4.8 分之間，根據該區間，說明 32 名學生的性別知識表現是否有顯著的進步？

公職王歷屆試題 (103 地方政府特考)

(三)輔導老師資料收集的一個缺點為「學生性別知識表現有顯著進步未必表示該輔導方式有效」。若樣本數依然為 32, 建議一個資料收集方式, 以改善該缺點。

【擬答】：

(一)

- (1)常態性:母體為常態分配
- (2)變異數同質性:二個母體變異數須相同
- (3)隨機獨立性:樣本取得須隨機且獨立

(二)其信賴區間 $3.2 \leq \mu_2 - \mu_1 \leq 4.8$ 為正, 即 $\mu_2 > \mu_1$, 即在信賴度 95%之條件下 性別知識表現有顯著的進步

(三)資料收集方法應隨機才可以收集到有代表性的樣本, 這樣可以降低型 I 錯誤, 以改善其缺點。

四、長興國中林老師編製了一份數學科的月考測驗(包含 40 題對錯計分的選擇題及 4 題多元計分的應用題)。下列為林老師建立這份測驗信度的方式: (1)他隨機抽取 100 位學生, 在月考後 1 週重新施測了一次, 得到兩次測驗分數的相關為.50; (2)他分別求得選擇題及應用題的 α 係數, 前者為.80、後者為.40; (3)他隨機抽取 50 份試卷, 請兩位評分者批改多元計分的應用題, 並求得兩位評分者給分的相關為.60; (4)他求得學生月考分數與上學期數學總成績的相關為.90。

(一)請說明林老師建立課堂數學成就測驗信度的方式中有那些是不恰當的?理由為何?

(二)針對林老師所蒐集的適當信度係數, 說明這些信度的意義及影響信度的因素, 並提出提高信度的方法。

【擬答】：

(一)

- (1)重測信度為 0.5 太低, 可能間隔時間太長
- (2)後者 α 係數 0.4 太低, 代表 4 題應用題的同質性偏低
- (3)評分者信度為 0.6, 即二位評分者在應用題的評分一致性為 0.6
- (4)月考分數與上學期總成績效標關聯效度為 0.9, 而非信度

(二)(郭玉生著: 教育測驗與評量 P63-P69)

(1)測驗的長度:

題數愈多, 信度愈高, 所以增加題數可提高信度

(2)團體的變異量: 團體的變異性愈大, 信度愈高。所以增加樣本變異程度可提高信度。

(3)測驗的難度: 難度愈適中, 信度愈高。所以增加難度適中的題目可提高信度。

(4)測驗的客觀性: 評分愈客觀信度愈高

(5)信度的估計方法: 採用適當的信度估計方法可提高信度