

《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：（每小題5分，共25分）

- (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？

試題評析	本題命題以敘述統計之統計量數為主要核心，並著重於標準分數之T分數與偏態係數之計算與解釋。雖然小題很多，但是皆僅止於計算性質的說明，難度偏易。
考點命中	《高上教育測驗與統計講義》第一回，傅立葉編撰，第三章至第六章。

答：

(一) 利用甲、乙兩班數學期末考成績的統計量數，可以分別計算林生與李生得分的T分數如下：

$$\text{甲班的林生 } z = \frac{85 - 80}{5} = 1, T = 50 + 10(1) = 60$$

$$\text{乙班的李生 } z = \frac{73 - 65}{4} = 2, T = 50 + 10(2) = 70$$

李生的T分數較高。

兩班的平均數差異過大，使用T分數而不用原始分數進行比較的前提是，當兩班數學期末考成績都近似常態分配時，在不同平均數與標準差的參數情境下，標準分數才能真正比較林生與李生兩人的數學能力程度。

- (二) 當張生的成績因登錄錯誤而少了5分，修正後因全班數學總分增加，乙班的平均數將變大；而因誤植的61分與正確的66分都比全班中位數的60分高，修正後的乙班數學期末考分數中位數則不受影響。
- (三) 若甲班老師將每位同學分數加5分，則該班平均數會提高至85分。中位數也因每人分數皆等量增加5分而向右位移至89分。
- (四) 若甲班林生得分的百分等級為63，加5分之後因為是每位同學都等量加分，屬於相對位置量數的百分等級因同學之間的成績高低關係維持不變，林生的分數百分等級也不會變動。
- (五) 依題目提供的數據，若要探討兩班分數的偏態，可利用平均數與中位數的差值是幾倍的標準差之概念加以判斷，計算如下：

$$\text{甲班：} \frac{80 - 84}{5} = -0.8$$

$$\text{乙班：} \frac{65 - 60}{4} = 1.25$$

由兩班分數的偏態係數可知，甲班呈現負偏，而乙班呈現正偏。

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？（25分）

試題評析	本題針對測驗結果解釋應注意事項進行命題。雖然測驗種類繁多，測驗結果解釋重點與方向不一，但是以智力測驗為例則容易發揮，測驗解釋涉及的倫理議題如能配合統計名詞與方法，要拿高分並不困難。
-------------	--

考點命中

《高上教育測驗與統計地特考前重點整理》，傅立葉編撰，最新之議題重點整理，(三)測驗學2.(3)施測後。

答：

當我們解釋一個重要的測驗結果時，以智力測驗為例，應注意的重要事項有以下幾點：

- (一)測驗結果要經過計分、參照常模、換算、畫側面圖、和解釋等幾項重要工作。
- (二)解釋測驗結果時應注意一般性說明、專業說明、及特殊說明。在解釋的過程中，一般性說明可依個別受試者表現與分數相對於群體受試者的優劣進行初步判斷；專業及特殊說明則可再次著重於該測驗的目的，並依特定測驗工具的形式甚至請受試者回想答題時的情境，狀況與心態，然後再針對其所得的測驗結果加以分析說明，並解釋各分項所代表與顯示的意義。
- (三)輔導者應鼓勵個別受試者參與討論，給予積極正面的協助。如果測驗結果涉及個人隱私，測驗結果需加以保密。
- (四)告知受試者不要互相比較，給自己貼標籤，使其可以為自己保留適度的彈性。
- (五)應將測驗結果與其他資料(如家庭社會經濟地位、晤談紀錄、教師評語、自然情境中的觀察、及學生本身健康情形等)配合使用，避免以單一測驗結果作為決定或解釋的依據。
- (六)如果有必要，更可運用測驗結果於個別諮商中，以蒐集資料作為證實，或是作為轉介至其他單位安置前的準備。

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。(10分)
- (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?(5分)
- (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。(10分)

試題評析

本題命題為推論統計之信賴區間估計，相依樣本T檢定的資料條件與計算結果推論能力的應用。考生雖不需計算，但應能嫻熟信賴區間估計結果的解釋與檢定顯著性的判讀。第三小題則為真正考驗考生統計實力的部分，應該也是今年四大考題中最難拿分的一個小題。

考點命中

《高上教育測驗與統計》第二回，傅立葉編撰，第十一章，25頁。

答：

(一)依題意，該輔導老師想進行性別知識輔導方案之相依樣本T檢定，收集資料應具有以下特性：

- 1.符合兩群體特性具有或近似常態分配的條件。
- 2.前後測變異數同質性條件的檢定。
- 3.成對資料的數量夠大的特性。

(二)若輔導老師得到前後測性別知識進步分數的95%信賴區間為(3.2, 4.8)，其表示在100次的抽樣中，約有95次性別知識進步分數的差異在該區間中，由於上限與下限都是正值，因此，在該輔導方案下，兩性知識確實有顯著進步。

(三)針對老師收集資料的缺點，主要是在老師教學與行政的掌握下，學生資料常陷於「便利抽樣」而有母體代表性不足，進而影響推論統計結果推論上的限制。在樣本數維持32的情況下，資料收集方式建議注意以下細節，以便提高推論統計結果對於因果關係的推論能力：

- 1.擴大樣本抽樣範圍，也就是打破同班的抽樣對象，擴大該輔導方案於該校其他年級班級的推行後，再行抽樣，應能提高母體代表性。
- 2.對於學生在性別知識的前測分數進行常態分配條件的檢測，確認其不違反常態分配的假設。
- 3.對於前後測分數不違反變異數同質性條件的檢測。
- 4.利用學生的性別或其他特質，進行適當的配對，以控制可能影響因果關係的干擾變項。

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？（10分）

(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。（15分）

試題評析	本題針對測驗結果解釋應注意事項進行命題。雖然測驗種類繁多，測驗結果解釋重點與方向不一，但是以智力測驗為例則容易發揮，測驗解釋涉及的倫理議題如能配合統計名詞與方法，要拿高分並不困難。
考點命中	《高上教育測驗與統計地特考前重點整理》，傅立葉編撰，最新之議題重點整理，(三)測驗學2.(3)施測後。

答：

(一)林老師為建立個人所編製之數學成就測驗的信度證據，分別使用了四種信度取得方式，其一為重測信度，其二為內部一致性信度的Alpha係數，其三為評分者信度，最後則為效標關聯效度。其中，較不恰當的信度建立方式有三，分別說明如下：

1.對錯計分的選擇題計算Alpha係數

由於Alpha係數只適合對於多重計分的題目或態度量表進行回答、得分的複相關係數計算，對於對錯計分選擇題應改採庫李信度為宜。

2.學生月考成績與上學期數學總成績的相關係數

這是屬於自編測驗與外在效標之間的效度概念，無法凸顯本學期教學成效與學生學習成就外，也未能提供信度證據。

3.針對多元計分應用題的相關

因為多元計分應用題只是該測驗的一種題型，部分題目的信度並無法提供整份測驗的信度證據。

(二)

1.重測信度：

使用同一測驗，針對同一群受試者，在不同時間內前後測兩次，求兩次得分間的相關係數(可依資料類型選擇計算積差相關係數或等級相關係數)，又稱穩定係數或再測信度。影響信度高低的因素為兩次測驗間隔時間長短。適當地掌握兩次測驗之間約6~8週的間隔應能提高測驗的信度。

2.Alpha信度：

針對多元計分的題目答題結果進行複相關係數的計算，通常測驗所欲測量能力如有不同面向或是存在分量表或分測驗，則需先分開進行計算與解讀，以取得內部一致性的證據。影響其信度高低的因素為題數，題數夠多而係數也高，則能提高測驗的信度。

3.評分者信度：

涉及評分者主觀判斷的測驗題型，如能取得不同評分者之間的評分一致性，則可提高測驗信度。命題時準備標準答案，以及提供評分參考標準應能提高測驗信度。

【版權所有，重製必究！】