

《教育測驗與統計概要》

| | |
|--------------------|--|
| <p>試題評析</p> | <p>今年考題命題主題稍嫌集中，相關係數與共變數就有兩題，信度也有兩題。不過難度普遍不高，其中又以第三題實作評量與客觀式測驗題的6個面向的回答需要費點心思，考生答題時需要先行構思最佳的表達方式。第五題的類別資料卡方統計法，因為計算最多，沒有準備或是公式不熟的考生可能會付出代價而成為分數高低之關鍵。綜觀今年考題之特點，可看出本考科的命題趨勢逐年傾向於深度觀念的質性試題。一般程度考生應可以有70~80分的成績，程度較佳同學可能在85分左右。</p> <p>第一題：本題命題核心在於兩個變項之間的相關程度探討的兩個統計量-相關係數與共變數計算上的重要性質。難度不高，數字計算簡單，由於沒有太多理論論述的空間，答題與高分關鍵在於能否寫出計算公式輔助說明。</p> <p>第二題：本題涉及測驗品質指標之一的信度探討與應用。信度研究設計上的變更考的是影響信度大小的因素檢視；信度與測量標準誤的關係所衍生的真分數信賴區間估計與解釋則需要計算公式的熟練。</p> <p>第三題：本題命題涉及實作評量的主軸議題，並就不同面相比較與傳統客觀式測驗的差異。其中以「學生反應的控制」與「課程內容取樣」較難意會與表達。有準備的同學比較有機會相對拿高分。以表格呈現答案較不易限於文字陳述的泥沼。</p> <p>第四題：本題涉及共變數與相關係數的運算性質，與第一題不同的是必須算出共變數的數值，加上，對於相關係數的解讀與變異數轉換的計算，雖然看似簡單，卻需要考生非常細心思考題意，謹慎作答，否則容易大意失荊州！</p> <p>第五題：本題命題環繞在卡方統計法的幾個重要觀念與計算式。包括：列聯表的期望值、卡方值(卡方統計量)以及phi相關係數。題目難度低，一般考生應該能夠輕易完成作答！</p> |
| <p>考點命中</p> | <p>第一題：《高上教育測驗與統計講義》第一回，傅立葉編撰，Ch.7，頁39-41。《高上104地方特考重點題神》最熱門之議題重點整理(一)迴歸分析。</p> <p>第二題：《高上教育測驗與統計講義》第三回，傅立葉編撰，Ch.15，頁11-12、14-15。</p> <p>第三題：《高上教育測驗與統計講義》第三回，傅立葉編撰，Ch.19，頁49-52。《高上104地方特考重點題神》最新之議題重點整理3. 各種重要測驗彙編，頁4-10~4-11。</p> <p>第四題：《高上教育測驗與統計講義》第一回，傅立葉編撰，Ch.7，頁40-41；Ch.4，頁13。《高上104地方特考重點題神》最熱門之議題重點整理(一)迴歸分析，p.4-13。</p> <p>第五題：《高上教育測驗與統計講義》第二回，傅立葉編撰，Ch.12，頁36-37。《高上104地方特考重點題神》最新之議題重點整理(二)推論統計，頁4-5。</p> |

- 一、(一) 假設兩變項 X 與 Y 之相關係數 $r_{xy} = 0.3$ ，當兩變項都轉換成標準 Z 分數後，兩變項之相關係數變成多少？(5分)
- (二) 當 $r_{xy} = -1.0$ 且 $z_x = -0.8$ 時， z_y 為多少？(5分)
- (三) 當 X 變項及 Y 變項的各個觀察值都乘上4倍後，其共變數的值又會變成原來共變數的多少倍？(5分)

答：

- (一) 當兩變項的相關係數為0.3，因為相關係數本身是一個標準化的統計量數，在兩變項皆轉為標準 Z 分數後，兩變項的相關係數仍維持原始變項0.3之相關強度。
- (二) 由於 $z_y = \beta z_x$ ，其中 $\beta = r_{xy}$
代入題目給定的數據資訊可得
 $z_y = -1.0(-0.8) = 0.8$ **【版權所有，重製必究！】**
- (三) 當原始兩變項的各個觀測值都乘上4倍，以共變數的計算公式：

$$C_{xy} = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{n-1}$$

影響所及為分子的計算，將兩變項乘上4倍的倍數提出相乘，則共變數的值會變成原來共變數的16倍。

- 二、某數學成就測驗適用於小學四年級到六年級，測驗發展者報告間隔6個月的再測信度為0.64，研究樣本來自某所小學的四年級到六年級的學生100名。該測驗的平均數為250，標準差為50。
- (一)請從分數變異的觀點解釋信度值0.64的意義。真分數與觀察分數的相關為何？(6分)
- (二)若由你來設計信度研究，你會變更那些研究設計(思考目前的信度研究有何缺失)?(10分)
- (三)請計算該測驗的測量標準誤，若小華在該測驗的得分為200分，則其分數的68%信賴區間為何?解釋此一信賴區間的意義。(9分)

答：

- (一)題目中該數學成就測驗的6個月重測信度0.64表示，該所小學之四到六年級100位學生的樣本資訊告訴我們該測驗的真分數變異數佔觀察(實得)分數變異數的64%。也表示真分數與觀察分數之間的相關為0.64。
- (二)本測驗的信度研究，在研究設計上可以進行的合理變更，主要從影響信度的幾個因素進行思考：
- 1.測驗長度：
從平均分數250而言，除非每個題目的配分偏高，否則該測驗的題數應該不少。理論上，測驗長度愈長，信度愈高。然而，過多的題數因為作答時間的限制可能造成猜答反而影響信度。雖然題目未提及題數，卻值得考量評估。
 - 2.測驗難度
以小學4到6年級的數學成就測驗而言，難度的掌控會是影響信度的要素之一。應在施測前進行難度檢視與調整。難易適中的試題組合才有較高的信度。
 - 3.團體變異量
100位樣本學生的組合應進行適當的分層隨機抽樣，團體變異愈大，信度愈高。
 - 4.重測法的替代方式
因為重測法的進行常有受試者流失的缺點，而且，時間間隔也需要特別注意，6個月幾乎已經跨越一個學期的教學期間，可能因為教學進度、教學成果、甚至學習心態的轉變而影響信度。因此，可以將重測時間間隔縮短為一或兩個月，如果該測驗品質夠好，應較能呈現較高的信度。
- (三)該測驗的測量標準誤可以計算如下：

$$\begin{aligned} SE_{meas} &= S_x \sqrt{1 - r_{XX}} \\ &= 50 \sqrt{1 - 0.64} \\ &= 30 \end{aligned}$$

接著，若小華在該測驗的得分為200分，則其真分數的68%信賴區間為

$$\begin{aligned} X \pm Z_{\alpha/2} (SE_{meas}) \\ &= 200 \pm 1(30) \\ &= (170, 230) \end{aligned}$$

對同一個體以同一評量工具重複評量，所有評量結果會落在真分數附近形成一個分配。這個觀察分數分配的期望值為真分數，其標準差則為測量標準誤。而以觀察分數加減包含測量標準誤在內的測量誤差所形成的信賴區間，得以說明真分數落在該信賴區間的機率。依題意的計算結果，此區間的意義為：研究者有68%的信心說明小華在該數學成就測驗的真分數會落在(170, 230)分的區間。

【版權所有，重製必究！】

- 三、試就下列6個面向：欲測量之學習結果、試題的準備、課程內容的取樣、學生反應的控制、評

分、信度，分別說明實作評量與客觀式測驗的差異。(20分)

答：

以觀察和專業判斷來評量學生學習成就的評量方式都可以稱為實作評量，其型式非常的多元化，例如建構反應題、書面報告、作文、演說、操作、實驗、資料蒐集、作品展示等，實作評量具有的特徵如：要求學生執行或製作一些需要高層思考或問題解決技能的事或物；評量的作業(tasks)是具有意義性、挑戰性且與教學活動相結合；評量的作業能與真實生活產生關聯；歷程和作品通常是評量的重點；評量的重要層面與給分標準，也就是表現的規準(criteria)和標準(standards)要事先確定。

實作評量是介於評量認知能力和在真實情境之應用能力兩者之間，在模擬各種不同真實程度的測驗情境之下，提供教師一種有系統的評量學生實作表現的方法。例如，自然課程比較重視學生在實驗室的操作技能、數學課程比較重視學生的演算及實際問題解決技能、語言課程比較重視學生的口語表達技能等，這些實作表現技能都是典型的客觀式和論文式測驗所無法測量出的，在這種情境下，唯有實作評量才可以提供直接測量上述學習成果的有用訊息。

實作評量與客觀性測驗在題目所指的6個面向的差異比較可以用下表整理得知：

| | 實作評量 | 客觀式測驗 |
|---------|--|---|
| 欲測量學習結果 | 以教師上課的內容為主，沒有固定的教材範圍 | 侷限於教材範圍 |
| 試題的準備 | 測量結合知識與複雜動作協調的技能領域。可以用口語表達方式進行，如演講、戲劇、角色扮演等；或以紙筆完成建構式作業，如編擬計畫、建造模型、繪製圖表等；或操作工具或設備(電腦、樂器、打字機)等。 | 由「題庫」中選取出已標示難易度，也顯示具有評量那項能力的功能，挑選數個題目並作適當的整合，即可成為一份合適的試卷。 |
| 課程內容的取樣 | 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。 | 單一目標範疇(認知能力) |
| 學生反應的控制 | 以應用綜合層次的、建構式的反應為主。可參考資料或共同完成活動來設計；計畫、建構和傳送原始的反應 | 以知識理解層次的、選出式的反應為主。不可參考資料，且要獨立測試；閱讀、評鑑、選擇 |
| 評分 | 由師生在評量前共同設計的評分規準 | 教師在評量前已有固定的標準答案 |
| 信度 | 穩定性及一致性程度。對行為表現的多次觀察及以一群有共識者以相同準則觀察同一的行為表現、作品、卷宗是提高實作評量信度的二個有效方法。 | 用傳統信度指標的計算方式，容易計算，但須要進行題目的調整與取舍等前置作業，方能達成對信度的期許。 |

四、某公司想了解該公司員工每天工作時數、薪水等級(分為1~10共10等級)與工作滿意度之關係，乃在全公司隨機抽樣1000人為樣本進行研究，請各人填交一天的工作時數及工作滿意度。

結果發現1000人每天工作時數的平均數為11，標準差為3，眾數為9。工作滿意度以0~100測量的平均數為60，標準差為15，與每天工作時數的相關為-0.30，與薪水等級的相關為-0.55。

- (一)工作滿意度與每天工作時數的共變數為多少？(4分)
 (二)如果該公司研究的結論是：工作時數越長的人工作滿意度越低，薪水越少的員工工作滿意度亦較低。工作滿意度100代表對工作非常滿意還是非常不滿意？薪水等級10是薪水最多或最少？請說明你的理由。(8分)
 (三)如果該公司員工每週均只上班5天，員工每週工作時數的變異數約為多少？每週工作時數與工作滿意度之相關為多少？(8分)

答：

(一)利用題目提供的統計量資訊，工作滿意度與每天工作時數兩個變數的樣本共變數計算公式為

$$\begin{aligned} C_{xy} &= r_{xy} \cdot s_x \cdot s_y \\ &= (-0.3)(15)(3) \\ &= -13.5 \end{aligned}$$

(二)由於工作滿意度與工作時數的相關-0.3表示兩者為負相關，因此，當工作滿意度以0-100測量，工作時數愈長則其數字愈大，則工作滿意度的測量值應該愈小，工作滿意度為100(最大值)的任一員工，表示其滿意度最高，等同於對工作非常滿意。另外，依題意，工作滿意度與薪水等級的相關也是負相關的-0.55，而且該公司結論薪水愈少，工作滿意度也較低，綜合這兩個資訊可知：薪水等級10表示薪水最低，對應較小的工作滿意度數值。

(三)由於每天工作時數的標準差為3，依題意，每週只上班5天，則員工每週上班時數的平均數成為55，變異數可以計算得：

$$5^2(3^2) = 225$$

又，每週工作時數與工作滿意度之相關由於不因由每天轉換為每週之倍數而影響，因此兩者的相關維持-0.30。

五、某房屋仲介商調查100名民眾(男女各半)對於購屋的意願，發現有60%的人想要購屋，40%沒有意願，有意願的人當中，男生高達80%，女生僅有20%。請寫出性別與購屋意願的4個細格期望值(可使用敘述或2×2列聯表表示)、卡方值與phi係數。(20分)

答：

(一)依題意，該房屋仲介對於100名民眾的調查結果，可以使用2×2列聯表呈現如下：

| | 想購屋 | 不想購屋 | 總計 |
|----|-----|------|-----|
| 男 | 48 | 2 | 50 |
| 女 | 12 | 38 | 50 |
| 總計 | 60 | 40 | 100 |

四個細格期望值可以使用公式

$$E = \frac{\text{行和} \times \text{列和}}{\text{總樣本}}$$

進一步計算得：

| | 想購屋 | 不想購屋 |
|---|-----|------|
| 男 | 30 | 20 |
| 女 | 30 | 20 |

(二)而卡方值為

【版權所有，重製必究！】

$$\chi^2 = \sum \frac{(O-E)^2}{E} = \frac{(48-30)^2}{30} + \frac{(12-30)^2}{30} + \frac{(2-20)^2}{20} + \frac{(38-20)^2}{20} = 54$$

(三)最後，phi係數可以計算得

$$\phi = \sqrt{\frac{\chi^2}{n}} = \sqrt{\frac{54}{100}} = 0.7348 \cong 0.74$$

高點
·
高上

【版權所有，重製必究！】