

103 年特種考試地方政府公務人員考試試題

等 別：三等考試

類 科：教育行政

科 目：教育測驗與統計

高明 老師 解題

一、某能力測驗適用對象為小學四到六年級學生，測驗有兩個平行版本，甲式及乙式。測驗發展者抽取 100 名四到六年級的學生作為信度研究的對象。這 100 名學生同時完成甲、乙式測驗，然後求取兩式測驗成績的相關，其值為 0.8。今有汪老師以此測驗測量小學四年級學生，為確保測驗結果的信度，抽取 100 名四年級學生，同時完成甲、乙兩式的測驗。

(一)測驗發展者估計的是何種信度？並解釋信度 0.8 的意義。

(二)請問汪老師所得的信度是否會高於原測驗信度？為什麼？

(三)測驗發展者覺得信度不夠理想，打算將信度提升到 0.9，如果以增加平行試題方式，測驗長度要如何調整？

【擬答】：

(一)1. 複本信度：當同一種測驗有二種以上的複本測驗 (Alternate-form Test) (測驗性質類似在內容、試題格式、長度、難度、鑑別度均一致) 時，可將此二種複本分別實施於同一群受試者，再根據此二份測驗分數計算其相關係數，即得複本信度，又稱等值係數 (Coefficient of Equivalence)。

2. 解釋信度 0.8 的意義：複本信度愈高，即表示二份測驗測量同一特質和屬性的程度愈高。複本信度的測量誤差來源，主要是來自試題抽樣產生的誤差，亦即試題取樣誤差為 20%，而複本信度真實分數占實得分數變異為 80%。

(二)1. 不會。在其他條件相等下，團體變異性愈大，其信度愈高。從相關係數來看，當團體內個別差異愈大，測驗分數分散範圍愈大，其相關係數也愈大，其測驗信度亦較高。

2. 四到六年級的學生變異程度較大，而小學四年級學生變異程度較小，因此汪老師測得信度會較小。

(三)1. 測驗愈長則其試題的取樣愈適當，測驗分數受猜測之影響亦較小，其信度也愈高。這是因為大量試題測驗比少量試題更能對欲測量特質作適當取樣，且新增加試題應與原試題同質性，才不致於產生試題異質性誤差。(葛樹人，民 91)

2. 應增加題目長度亦即增加同質性試題題目。

二、隨著教育評量的改革趨勢，建構式反應 (constructed-response) 題型/實作評量已普遍出現於各種考試，但部分學者、教師、家長、學生質疑此種題型的評分過於主觀，可能造成偏誤，因而對此種形式的試題持保留態度或反對立場。請就你對這類評量的瞭解，及它在班級評量的應用情形，至少舉出三種措施使此種題型或評量的偏誤最小化。

【擬答】：

(一)詳細說明細節行為的項目，以及教師所期望學生達成的表現標準，即評分標準與評量重點須事先決定。美國「全國專業教學標準協會」(NBPTS) 提出當受試者充分地瞭解計分標準、知道評分者如何評定其作業時，有助於增加測驗評分一致性。

(二)提供適當的表現情境：評量標準界定後，教師便需準備可供觀察表現成果的施測情境，情境選擇視進行評量表現或成果特質而論。其選擇情境原則有：①教室裡自然發生表現的頻率；②決定的重要性；③到底需要多少訊息才能決定。

(三)主試者評分過程產生月暈效應、評分前後寬鬆不一、邏輯誤差、偏見等因素進而影響評分一致性。透過嚴謹適當主試者訓練，可提高評分者間一致性、評分前後標準需一致、預擬一份評分要點、使用適當的評分方式、匿名彌封。

三、某房屋仲介公司的網站，定期公告每個經紀人售出房價的集中量數，一般來說，每個房屋經紀人所賣出的房屋中，會有少數特別昂貴，不過每個人所賣出的房屋數相差無幾。網站上也

公職王歷屆試題 (103 地方政府特考)

呈現該地區兩個地段 A 與 B 的房價標準差，這個月 A 地段售出的房子有 20 戶，標準差為 12 萬元/坪，B 地段售出的房子有 30 戶，標準差為 8 萬元/坪。

- (一)若房屋經紀人的佣金為賣出總房價的 5%，請問那個集中量數與房屋經紀人的收入最相關？為什麼？
- (二)王先生想利用該網站粗估一下該地區的房屋售價散布情形，請問將兩個地段資料合併後，該地區的房價標準差是多少？
- (三)若該仲介公司改變佣金制度，將佣金從 5% 減少為 4%，請逐一說明房屋經紀人的收入之變異量數分別有何影響？(1)全距 (2)四分差 (3)變異數。

【擬答】：

(一)中位數。因為題目中會有少數特別昂貴為極端值，平均數容易受極端值影響，叫不適合探討收入與平均數相關，而中位數不受極端值影響，且測量量尺為次序變項功能比眾數強。

$$(二) S_p^2 = \frac{(n_1 - 1) \times \hat{S}_1^2 + (n_2 - 1) \times \hat{S}_2^2}{n_1 + n_2 - 2} = \frac{(30 - 1) \times 12^2 + (20 - 1) \times 8^2}{30 + 20 - 2} = 112.33$$



高明小叮嚀.....

本題並無給定 A、B 地段的平均數並無法使用 $S^2 = \frac{\sum_{i=1}^k n_i [S_i^2 + (\bar{x}_i - \bar{\bar{x}})^2]}{\sum_{i=1}^k n_i}$ ，且題目中「每個人所賣出的房屋數相差無幾」（具有強韌性）必須使用合併變異數，同學亦可將此公式寫於考卷中。

- (三) 1. 全距會變小。因為 $R = \text{Max} - \text{Min}$ 其中 Max、Min 原本乘以 .05 改為 .04 則其 R 會變小。例如賣出最高單價 1000 萬最小單價 100 萬，則原本 $R = 1000 \times .05 - 100 \times .05 = 45$ 萬，新制 $R = 1000 \times .04 - 100 \times .04 = 36$ 萬。
2. 四分差會變小。因為 $QD = \frac{Q_3 - Q_1}{2}$ ，其中 Q_1 與 Q_3 都會變小其差距亦會變小。
3. 變異數變小。 $S_x^2 = \frac{\sum (x - \bar{x})^2}{n}$ 必須考量每一位佣金收入，原本 $S_{.05x}^2 = .0025 S_x^2$ 改為 $S_{.04x}^2 = .0016 S_x^2$ ，其變異數會變小。

四、王老師目前任教小學三年級，他想瞭解班上同學的體重會不會有過重的情形。在全班所有 30 名小朋友量過體重後，王老師想找個數據來描述小朋友體重資料的集中情形而來請教你。請問你如何向王老師說明平均數、眾數、中位數三種集中量數指標的意義？

【擬答】：

(一)平均數

1. 意義

平均數 (Arithmetic Mean) 乃將所有觀察值的數值資料的總和除以總觀察值的個數，由於其計算簡單方便且易於瞭解，是集中量數中最常使用的量數，常以 M ， μ ， \bar{X} (唸 x bar) 表示。

$$\text{公式} : \bar{X} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n}$$

2. 適用時機：用於等距變項及比率變項，推論統計時可進行推論、用於比較兩個團體之成就水準、作其他較複雜統計分析之用、作為測驗常模基準、最為穩定集中量數。

(二)中數 (Median)

1. 意義：又稱二分位數或中位數 (Mid-score)，係指順序數列或次數分配中心項之數值，一般皆以 M_d 表示。將資料排序 (由小到大)，最中間位即表中位數，代表此一團體在此變項的集中情形。中位數等於第二「四分位數」、第五「十分位數」、第五十個「百分位數」。適用情形為資料中有極端值、次序變項、偏態時使用。

$$\text{公式} : M_d = \frac{N+1}{2} \text{ 項或 } M_d = L + \left(\frac{N}{2} - F_L \right) \cdot \frac{h}{f}$$

公職王歷屆試題 (103 地方政府特考)

2. 適用時機：適用於有極端值資料與次數分配為偏態資料、觀察值與中位數絕對差和最小、可做無母數統計推論、簡單明瞭容易計算。

(三)眾數 (Mode)

1. 意義：眾數亦為集中量數之一，指連續數列資料中出現次數最多的數值或最多數人所得到的分數。一般都以 M_o 表示。眾數並非唯一，也可不存在，通常使用於名義變項。

公式： $\bar{X} \quad M_o = 3(\bar{X} - M_d) \quad M_o = \bar{X} - 3(\bar{X} - M_d)$

2. 適用時機：適用於有極端值、適用於偏態資料、適用質的資料。眾數是唯一在質的資料中，可求得集中量數、分組次數表有不確定之組距仍能求得眾數、近似眾數計算，甚為簡單，易於瞭解。

五、研究調查不同學院的大學生之打工經驗滿意度是否不同，在文、理、工學院各抽取 100 名具打工經驗的學生，其中對打工經驗表示滿意的分別各有 70、75、78 名，其餘表示不滿意。

(一)請利用統計檢定不同學院的學生其打工經驗滿意度是否不同？請說明檢定須查何表？如何在表中找到臨界值 A？並說明如何利用臨界值 A 做決策。

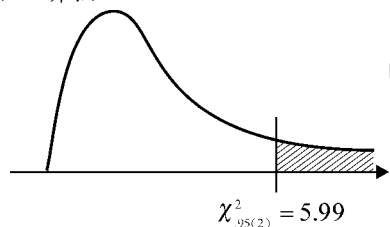
(二)這個假設檢定的基本假定(assumptions)為何？

【擬答】：

(一)1. $H_0 : P_1 = P_2 = P_3$

$H_1 : P_i$ 不全等

2. 臨界點



3. 蒐集到資料：

	文	理	工		期望次數		
滿意	70	75	78	223	74.33	74.33	74.33
不滿意	30	25	22	77	25.67	25.67	25.67
	100	100	100	300			

計算：

$$\chi^2 = \sum \sum \frac{(f_o - f_e)^2}{f_e} = \frac{(70 - 74.33)^2}{74.33} + \frac{(75 - 74.33)^2}{74.33} + \frac{(78 - 74.33)^2}{74.33} + \frac{(30 - 25.67)^2}{25.67} + \frac{(25 - 25.67)^2}{25.67} + \frac{(22 - 25.67)^2}{25.67} = 1.712$$

4. 結論：

$\chi^2 = 1.712 < \chi^2_{0.95(2)} = 5.99$ ，落入接受區，亦即無充分證據支持不同學院的大學生之打工經驗滿意度百分比有顯著差異。

(二)基本假定

- χ^2 考驗只能處理間斷變項，列聯表中只能處理人數或次數，無法處理受試者分數。
- χ^2 考驗表中每一次數必須獨立，每個觀察值只能歸類於一個類別或水準，亦即相同受試者不可能在聯列表出現二次。
- 所有觀察值必須全部列舉，不只要彼此獨立，亦不可省略任何一個觀察資料。
- χ^2 考驗中期望次數少於 5 (或 10)，就必須進行耶氏校正，用來避免 χ^2 值被高估產生錯誤估計。
- 當 χ^2 考驗達顯著差異時，想進一步瞭解到底那二組有差異，必須進行事後比較。(謝廣全，民 82)