

## 105 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

- 一、某位教育學者自編一份數學成就測驗，試圖用來預測學生在「國中會考」的數學科成績。但該學者在發展測驗的過程中，於進行效標關連效度分析時發現，他的自編測驗缺乏信度。請問：你會建議他該如何改善自編測驗缺乏信度的問題？(25 分)

### 【擬答】

效標關連效度的外生效標本身必須具備信度(信賴性)、穩定性，否則測驗結果會受情境因素影響產生誤差。改善自編測驗缺乏信度的問題方法分述如下：

#### (一)選取較大變異性的受試者樣本

1. 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. 由信度定義公式實得分數變異量愈大，表示團體間個別差異也愈大，若誤差分數變異量不變，則測驗結果信度就愈高。

#### (二)選擇難度適中的測驗

1. 試題太難或太易均會使大多數學生答錯與答對，均會產生二種極端情形(如正偏分配與負偏分配)，使其測驗分數的變異程度較小，信度會較低。而測驗題目難易適中，可使測量特質的範圍較大，提高信度。
2. 試題難度分布的範圍越廣，求出來的信度會越高，例如 P(難度值)在.25~.90 之間(題目難度的變異情形大，總分的變異情形也大)會比 P 在.35~.75 之間的信度高。

(三)採用客觀式的測驗題型：測驗的評分如受評分者個人主觀意見與判斷的影響，則不同評分者評分的結果含有很大評分差異，如此，測驗的信度係數則較低。因此客觀測驗如是非題、選擇題、配合題應有較高信度係數。

(四)測驗間隔時間適中，不宜過長或過短：複本信度二次施測間隔時間愈短，信度愈高，反之，間隔時間愈長，其他變項介入可能性增加，信度降低。

#### (五)當信度低時可以用二種方式補救

1. 增加試題或刪除不良試題：刪除題目與總分負相關的題目、刪去題目與總分相關最低的題目。
2. 校正相關係數的萎縮：當我們使用測驗工具均存在誤差，大部分工具信度並非完全信度，此種情況下必須使用校正相關係數的萎縮，否則得到研究結果會有低估相關係數可能存在(余民寧，民 95)。Nunnally (1978) 提出校正相關係數(Correction for Attenuation)公式，以具體的數學模式，說明信度與效度的關係，當某一個測驗(X)與效標(y)求得實際的效度係數為  $r_{XY}$  時，若二者的信度分別為  $r_{XX}$ 、 $r_{YY}$ ，完美效度係數以  $r_{XY}$  最大值表示，為二個完美的測驗所求得的預期相關：

$$\text{公式：} \hat{r}_{XY} = \frac{r_{XY}}{\sqrt{r_{XX}r_{YY}}}$$

二、請根據下列試題分析結果的局部摘要數據，回答下列問題並說明其原因：(每小題 5 分，共 25 分)

學生注意係數表

學生	得分	答對率	注意指標	判定類別
A	54	67.50	0.69*	B'
B	61	76.25	0.24	A
C	44	55.00	0.27	B
D	73	91.25	0.55*	A'

試題注意係數表

試題	答對人數	答對率	刪題後信度	注意指標	判定類別
1	16	34.78	0.76	0.90**	B'
2	40	86.96	0.74	0.81**	A'
3	44	52.17	0.73	0.43	A
4	14	30.43	0.72	0.37	B

平均答對人數：30.25，受試者總人數(N)：46.00

平均答對率(P)：0.76，試題總題數(K)：80.00

內部一致性係數：0.75，平均點二系列相關係數：0.45，差異係數：0.65

- (一)那一位學生是屬於「粗心大意型」的學生？
- (二)那一題是屬於需要大幅修改甚至是刪除的試題？
- (三)此次試題分析結果，顯示本測驗的作答難易度為何？
- (四)此次測驗結果，是否反映出學生作答能力的真實情況？
- (五)此次測驗結果，是否反映出試題編製的品質良好？

**【擬答】**

- (一)D 生；學生答對百分比 91.25%，學生注意係數.55，表示粗心大意不細心造成錯誤。
- (二)試題 1；判定類別 B'區為拙劣型試題，試題非常不當，可能含相當多異質成分，或資料登錄錯誤，或試題題意模糊不清，必須加以修改或刪除。
- (三)平均答對率=.76，顯示 P 值較大試題題目偏易。
- (四)學生 A 注意係數.69 與學生 D 注意係數.55，即表示該學生的反應組型為不尋常的情況已是很嚴重，應予注意；學生 B 注意係數.24 與學生 C 注意係數.27，即表示該學生的反應組型發生不尋常的情況並不嚴重，或還在容許的作答誤差範圍之內。
- (五)一般常模參照測驗大都測量綜合性能力， $D^* = .5$  左右的標準值，當  $D^* > .6$  時表示有相當多異質因素在內，此份試題的差異係數=.65，可能是學生學習動機或意願低落、學生對某些觀念有所誤，或者可能是命題有問題所造成的，應對學生反應組型予以注意，對試題檢討或修改。

三、某教育統計學者擬進行以「成就動機」(X 變項)預測「學業成績」(Y 變項)的迴歸分析研究，但因為熬夜工作，頭昏眼花，錯把初步獲得的標準化迴歸公式寫成下列的式子：

$$\text{成就動機} = 0.5 \times \text{學業成績}$$

請回答下列問題：

- (一)他原本打算獲得的標準化迴歸公式應該表示為何？(5 分)
- (二)成就動機可以預測到學業成績多少百分比的變異量？(5 分)
- (三)事後他仔細一想，若改成以「學業成績」預測「成就動機」的結果才是合理的話，則該預

## 公職王歷屆試題 (105 地方特考)

測公式的決定係數為何？(5分)

(四)經過這兩次的測試分析，他終於發現以「成就動機」預測「學業成績」的標準化迴歸係數，和以「學業成績」預測「成就動機」的標準化迴歸係數，彼此間有何不同？請說明之。(10分)

### 【擬答】

(一)標準分數迴歸方程式  $\hat{Z}_Y = \beta Z_X = r Z_X = .5 Z_X$

(二)成就動機預測學業成績其  $r^2 = .25$ ，即表「在學業成績中，可被成就動機解釋到變異量部分為 25%，即成就動機預測學業成績預測正確部分高達 25%」

(三)決定係數  $r^2 = .25$ 。

(四)1. 應該相同。

2. 簡單迴歸時自變項與依變項的積差相關不變，無論「成就動機」預測「學業成績」的標準化迴歸係數，和以「學業成績」預測「成就動機」的標準化迴歸係數，兩者標準化迴歸係數應該相同。

3. 多元迴歸自變項與依變項對調則標準化迴歸係數會改變。

四、某教育學者根據某城市小學六年級國語科基本能力檢測結果，進行性別差異的檢定分析，獲得如下表的結果：

性別	人數	平均數	標準差	t 值	p
女生	8398	528.38	77.35	27.55	.000**
男生	7731	516.23	78.98		

請回答下列的問題：

(一)那一種性別的個別差異較大？(3 分)

(二)那一種性別的檢測成績較佳？(3 分)兩種性別差多少分？(3 分)

(三)該教育學者係使用何種統計檢定方法？(6 分)

(四)請根據上表數據，解釋該城市小學六年級國語科基本能力檢測結果，是否具有性別差異存在？請說明之。(10 分)

### 【擬答】

(一)男生的個別差異較大，因為男生標準差 78.98 比女生標準差 77.35 大。

(二)女生的檢測成績較佳，因為女生平均數 528.38 大於男生平均數 516.23。兩種性別差 = 528.38 - 516.23 = 12.15。

(三)獨立樣本 t 考驗。

(四)  $t = 27.55$ ， $P$  值 = .000 <  $\alpha = .05$ ，落入拒絕區，拒絕虛無假設，亦即有充分證據支持男女生國語科基本能力檢測平均數達顯著差異，惟此裁決有犯型 I 錯誤 5% 機率。