

99 年公務人員特種考試原住民族考試試題

等 別：四等考試

類 科：教育行政

科 目：教育測驗與統計概要

一、某教育機構編製了一份適用於國小三至六年級的標準化成就測驗（原始分數之平均數為60、標準差為10；常態化量尺分數之平均數為100、標準差為20），下表為400位學生的部分原始數據，以及美美的成績。

原始分數	次數	累積次數
...		
71	10	352
70	12	342
69	8	330
...		

美美的成績：

原始分數	70
1. PR	?
2. 常態化量尺分數	?

- (一)美美是一位四年級的學生。請根據上表數據算出美美的PR值、常態化量尺分數，並向美美的家長分別解釋其意義。
- (二)何謂量尺分數 (scaled score)，其優點為何？常態與非常態轉換的量尺分數有何不同？你認為在何種情況下適合（或不適合）使用常態轉換的量尺分數？請說明理由。
- (三)假如該份測驗為診斷測驗，請說明以上表提供的各項指標解釋受試者的表現是否恰當？若不恰當，你的建議為何？

【擬答】：

(一) $Z = \frac{X - \bar{X}}{S} = \frac{70 - 60}{10} = 1$ ，PR=84。表示一個人的分數在某一團體中所占的地位，亦即團體中按某一屬性依序排列且分成一百個等級情況下，某一個人分數能勝過多少個百分等級，以 PR 表示。例如：美美得分 70 分在團體中 100 人勝過 84 人，則表示 PR=84， $P_{84} = 70$ （原始分數 70 分便相當於第 84 個百分位數）。百分等級為次序變項。

2.

原始分數	次 數	累積次數	cf 中點	cp 中點
71	10	352		
70	12	342	336	.84
69	8	330		

∴ PR = 84，常態化 $Z = 1$

常態化量尺分數 = $20Z + 100 = 20 \times 1 + 100 = 120$ 。

- (二) 1. 量尺分數：國際上大型標準測驗通用的方式，因為具有和群體比較的意義，好像一把用來測「量」個人表現的「尺」，所以稱為「量尺分數」，量尺分數的等化 (Equating) 為將題庫中的每道試題都透過「試題反應理論」(Item Response Theory, 簡稱 IRT) 的估計程序，將試題難度連結在同一把尺上，所以每道題目的難度值是已知的。著名的 TOEFL、GRE、國中基測等電腦化適性測驗也都是採用這種作法。

2. 量尺分數優點

- (1) 考生之間可在單一學科能力進行比較，考生自己各學科亦可進行比較，以及考生們在不同測驗日期學科能力表現比較。
 - (2) 分數解釋有一定標準，以增進測驗分數的解釋。
 - (3) 不同能力考生所得分數精確程度會類似，不至於有高低能力考生測量誤差比能力中等考生測量誤差來得大。
3. (1) 常態轉換適用不同分配型態的分數若須直接比較，就必須將原始分數轉換成符合常態分配的標準分數，此種分數稱常態化標準分數；亦即不論原來分配型態為何，經常態化轉換均成為常態分配（郭生玉，民 94）。
- (2) 非常態轉換為將原始分數轉換成 Z 分數時，分數會有負數和小數之值出現，為此經由 $Z = az + b$ 的方式將 Z 值都變成正值，以利解釋；其中直線轉換，並未改變原始分數之分配情形。原量表與新量表分配型態相同均為常態分配，將原量表轉換成標準常態分數的方法稱之。
- (三) 1. 診斷測驗：教學過程中或教學後，評量學生學習困難之處，以作為補救教學依據。採加分項分析或選題分析學生的反應，以確定學生學習困難地方。依測驗時間區分診斷性評量為教學中進行評量。只適用少數學習成就低落或身心障礙學生，日後進行補救教學。
2. 診斷測驗難度較低（P 值較大—分配形狀應為負偏形狀），只做為診斷學生錯誤所在。因此標準化成就測驗與常態化量尺分數均為標準分數轉換方式，屬於團體性常模用來表示相對地位量數，表示某一特質在群體中相對所佔的地位之量數，並不適合探討診斷測驗；應使用診斷學習困難的策略包括剖析內容領域的優點與弱點、鑑定先備知識的缺失、鑑定未精熟的學習目標、鑑定學生學習錯誤、確認學生的知識結構、確認解決文字問題的能力。

二、假設你自行發展一份資源班學生篩選用的「數學能力測驗」，為支持該工具的適用性，請擬定一個以因素分析法（factor analysis）及多項特質-多項方法（multitrait-multimethod method）進行建構效度（construct validity）資料的蒐集計畫。計畫中請列明：

(一) 何謂建構效度？

(二) 假設將「數學能力架構」界定為數與量、統計、代數。如何以因素分析法（factor analysis）驗證此數學能力的建構效度？

(三) 如何以多項特質-多項方法（multitrait-multimethod method）進行數學能力建構效度資料的蒐集？

【擬答】：

(一) 建構效度

指測驗能測量理論的概念或特質的程度，換言之指測驗分數能夠依據某種心理學的理论概念加以解釋的程度，凡用心理學的概念來分析測驗分數的意義皆為建構效度。（郭生玉，民 95）根據社會學與心理學所建構出來描述某種特質的理论，並對測驗分數所做的分析與解釋，如動機、性向、批判思考、社會性、內外向、情緒智商 EQ、利他行為、愛情……。（余民寧，民 95）

(二) 因素分析

1. 驗證性因素分析(Confirmatory Factor Analysis)：邱皓政（民 95）指出研究者對於依變項的內容與性質，在測量之前即有明確說明、文獻或理論依據，可以事先決定相對應的依變項的數目。進行因素分析時，即在考驗這事先提出因素結構適切性，甚至檢驗理論架構，即為驗證性因素分析。

2. 計算各變項間積差相關，組成相關矩陣、估計共同性(Communality)。

3. 抽取共同因素(Common Factor)得到未轉軸因素負荷量(Unrotated Factor Loading)。抽取方法有「主軸法」和「形心法」。

4. 進行因素轉軸：使每個變數只歸屬於一個或少數幾個因素，以得到轉軸因素負荷量。因素轉軸方法有「正交轉軸法」與「斜交轉軸法」。

公職王歷屆試題 (99 原住民特考)

- (1) 正交轉軸法是「最大變異法」(Varimax)。目的在使轉軸後每一共同因素本身的因素負荷量大小差距達到最大，以利共同因素的區辨與命名。
- (2) 通常採用學者凱塞轉軸標準，只留取「特徵值大於 1」的特徵向量。特徵值為將每個變項其所有變異因素組型之平方加總而得。
- 5. 進行因素命名(數與量、統計、代數)，是一種依據學理的猜測工作。(林清山，民 87)

(三) 學者 Campbell 與 Fiske 在 1959 年提出一份良好的構念效度，須符合下列二條件：

1. 聚斂效度(輻合性效度)(Convergent Validity)
一份測驗分數必須和相同構念或潛在特質其他測驗的分數間存在高相關，即不同方法測量相同特質所得的相關。使用三種方法(數學測驗一、數學測驗二、數學測驗三)測量數學能力架構三項特質(數與量、統計、代數)。
2. 區別效度(辨別效度)(Discriminate Validity)
一份測驗分數必須和不同構念或潛在特質其他測驗的分數間存在低相關，即相同方法測量不同特質與不同方法測量不同特質所得的相關應低相關。
3. 如果構念效度令人滿意條件，相同方法測量相同特質與不同方法測量相同特質所得的相關，應該比相同方法測量不同特質與不同方法測量不同特質所得的相關高。
4. 多特質—多方法矩陣表

特質	方法 1			方法 2			方法 3		
	數與量	統計	代數	數與量	統計	代數	數與量	統計	代數
方法 1	數與量 統計 代數								
方法 2	數與量 統計 代數								
方法 3	數與量 統計 代數								

三、某研究探討國小四年級的數學成就表現，蒐集了 30 名學童的數學測驗分數，以及他們的年齡。其中一名學生未填寫年齡，研究者登錄資料時，輸入 99 的數據，但分析時並未界定為遺失值，因此，統計軟體分析結果得到年齡的平均數 12.6 歲，眾數 9.25 歲，中位數 9.4 歲。請問：

- (一) 若研究者將「99」重新宣告為遺失值，則此組資料的眾數、中位數和平均數會變成多少？標準差和全距的變動那一個較大？還有那些離散(變異)量數比前兩者較不受此一資料更正的影響？請說明理由。
- (二) 研究者想將年齡改以月份為單位，即將年齡的數據乘上 12，那麼平均數和標準差會如何改變？請以簡單圖形分別畫出以年及以月份為單位的次數分配圖形，並標示出眾數、平均數的位置與橫軸的單位(請以修改後的數據為準)。

【擬答】：

(一) 1. 眾數 = 9.25；中位數 = 9.4；平均數 = 9.62。

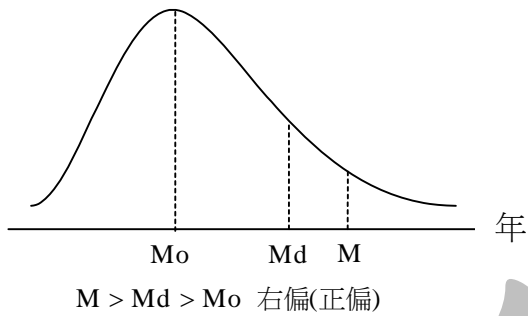
$$\bar{X} = \frac{30 \times 12.6 - 99}{29} = 9.62 \text{ (遺失值則將其數值刪除不計算)}$$

2. 全距變動較大(全距為最大值減去最小值，易受極端值影響，故十分不穩定)。

公職王歷屆試題 (99 原住民特考)

3. 四分位差較不受極端值 99 的影響 ($QD = Q = \frac{Q_3 - Q_1}{2}$ 只重視中間 50% 部份, 故反應不靈敏不受極端值影響)。

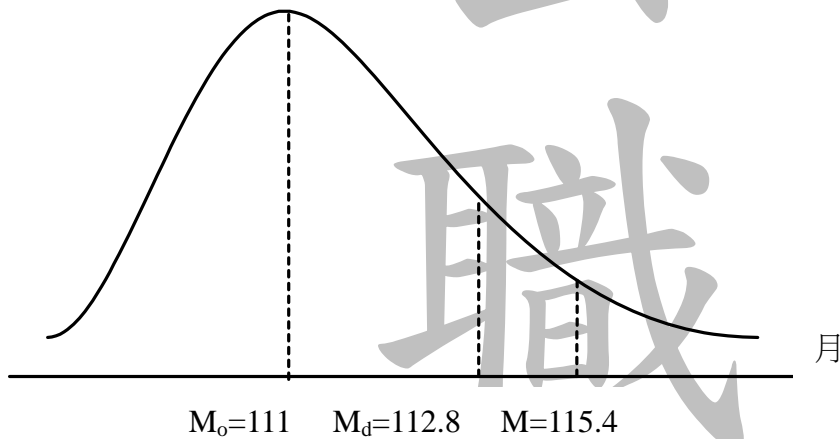
(二) 1. 以「年」為單位次數分配圖



2. 以「月」為單位次數分配圖

$$Y = 12X \Rightarrow \bar{Y} = 12\bar{X} = 12 \times 9.62 = 115.44 ; \text{新 } M_o = 12 \times 9.25 = 111 ;$$

$$\text{新 } M_d = 12 \times 9.4 = 112.8$$



四、某位大學教授在學期初和學期末針對大一學生對於專業課程應用與發展的了解情形進行調查, 結果如下表, 請問學生學期末時是否對於專業課程與發展更為了解?

		學期末	
		不了解	了解
學期初	了解	6	15
	不了解	11	18

(一) 請列出本題之虛無假設及對立假設。

(二) 試求出本題之考驗統計數。

(三) 請問學期末大一學生是否對於專業課程應用與發展更為了解?

$$(\alpha = .05, \chi^2_{.95(1)} = 3.841)$$

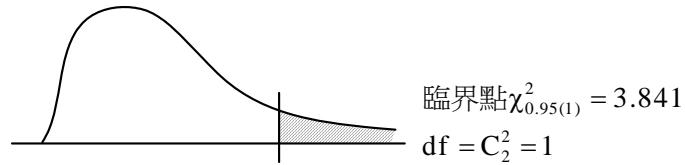
【擬答】：

用麥氏考驗進行

(一) $H_0 : A = D$

$H_1 : A \neq D$

(二) 1.



2. 計算

$$\chi^2 = \frac{(A-D)^2}{A+D} = \frac{(18-6)^2}{18+6} = 6$$

(三) 結論

$\chi^2 = 6 > \chi^2_{0.95(1)} = 3.841$ ，落入拒絕區，表拒絕虛無假設，支持對立假設，亦即有充分證據支持大一學生在學期初及學期末對於「專業課程應用與發展」有顯著改變，惟犯型 I 錯誤的機率 0.05。

公
職
王