

## 98 年特種考試地方政府公務人員考試試題

等 別：四等考試

類 科：教育行政

科 目：教育測驗與統計概要

### 一、簡答：

(一)全班共 50 人，排名第 3 名學生的百分等級為多少？

(二)甲生在 WISC 的全量表智商為 85，那麼與他同年齡的學生中有多少百分比的學生智商比他高？

(三)甲班共 25 人，施測魏克斯勒兒童智力量表，全量表 IQ 的平均是 110，要作 95%信心水準的區間估計，問其值應該落於多少與多少之間？

(四)某生在某測驗的 T 分數( $M=50$ ,  $SD=10$ )為 45，若要換算為 AGCT 分數( $M=100$ ,  $SD=20$ )，問其 AGCT 分數為多少？

【擬答】：

$$(一) PR = 100 - \frac{100 \times 3 - 50}{50} = 95$$

$$(二) WISC = 15Z + 100 = 85, Z = -1, PR = 16$$

與他同年齡的學生中有 84% 的學生智商比他高。

$$(三) \bar{X} \pm Z_{.975} \times \frac{\sigma}{\sqrt{n}} \Rightarrow 110 \pm 1.96 \times \frac{15}{\sqrt{25}} \Rightarrow 104.12 \leq \text{真實分數} \leq 115.88$$

進行區間估計施測魏克斯勒兒童智力量表介於 104.12 至 115.88 間信心水準為 95%，惟犯型一錯誤 5%。

$$(四) T = 10Z + 50 = 45, Z = -.5, AGCT = 20Z + 100 = 20 \times (-.5) + 100 = 90。$$

### 二、解釋下列名詞：

(一)自比性測驗格式(Ipsative test format)

(二)等距量尺(Interval scale)

(三)期望表(Expectancy table)

(四)多重切截分數(Multiple cut-off scores)

【擬答】：

(一)自比性測驗格式(Ipsative test format) 用來測量個人在多種興趣或人格特質上相對強弱，此類題型提供二個或二個以上具有相等吸引力或無吸引力題目，選出一個較能代表受試者特質敘述，但無法做不同受試者間比較，如愛德華斯個人興趣量表 EPPS、KOIS 量表全部為強迫選擇題技術（顯示自己本身各種興趣相對優勢）。

(二)屬於四大變項分類方式，不僅能標示事物與事物間不同特質或類別而指定的數字，亦能指出大小次序（等級順序）、多寡、優劣、高低，且能以數值方式表示二事物差別之大小變數；等距變數具有一基本特性「相等單位」，但無絕對零點、無倍數關係。平均數、標準差都有意義，尚有溫度、明暗度、音量、智商、成績分數。

(三)期望表(Expectancy table) 一種表示效標關聯效度方法。預期表為一種雙向圖表預測分數排在左邊（成績高低由上而下順序排列），效標（分數）排在表上端（成績高低由左而右排序），建立各層次測驗分數來預測而得到效標分數的機率。

(四)使用測驗來甄選人員時，測驗分數決斷點或切截分數，用來探討標準參照測驗錄取或拒絕二分計分。當切截分數訂得較高時，錄取率則會相對減低，而提高命中率與增加效度；而若當切截分數訂得較低時，錄取率則會相對增加，而降低命中率與增加效度。

### 三、何謂常模(Norms)，它有那些分類方式？選用常模時應注意那些原則？

【擬答】：

## 公職王歷屆試題 (98 地方政府特考)

### (一)常模(Norms)

常模(Norm)解釋測驗分數的依據，指特定參照團體或標準化樣本(Standardization sample)在測驗上所得到的平均數水準典型表現，因此有時常模會等於平均數。根據對標準化樣本施測所得的結果，經統計分析，整理出一個系統的分數分配表，按照高低排列，所得的平均數，即是常模。

### (二)分類方式

- 1.發展性常模：將受試者的測驗分數，在不同發展階段中比較其他個體獲得發展水準，此一分數稱之發展性常模。又分為年齡常模(Age norm)與年級常模(Grade Norm)。
- 2.團體內常模(Within-group Norm)：將個人測驗分數在團體中與其他成員比較相對位置，以確定分數高低程度。團體內常模分為百分等級、標準分數，標準分數又分為直線轉換(T 分數、WISC、BIS、GRE、SAT、AGCT)與非直線轉換標準分數(T 量表、標準九、STEN 分數、C 量表、常態曲線當量)。

### (三)選用常模時應注意那些原則

建立常模表時，不可能對整個母群體施測，必須只抽取部份樣本構成常模樣本(Norm Sample)或稱標準化樣本。測驗編製者所提供常模是否恰當，可由三方面價值判斷。

#### 1. 新近性

測驗常模隨著時間，會愈來愈不適合用來解釋測驗分數。因為隨著時代演進，如教學方法更好、課程重點改變，常模即不再合適，甚至連測驗分數本身也不準確。因此測驗常模應該採用最新的，但不能用測驗手冊的出版日期判斷。例如九十年代人們價值取向和六十年代必有不同，隨著社會變遷時代演進，課程內容不斷更新、傳統測驗與常模不再適用。

#### 2. 代表性

代表性指選取常模樣本的人口學特徵須與母群體特徵相吻合。測驗分數經常有測量誤差與抽樣誤差。其中抽樣誤差是由於抽樣不當所致，建立常模樣本其代表性比樣本大小重要，因為草率抽樣來的二萬名樣本，比正確取樣過程 200 個樣本更不具代表性。隨機取樣所得樣本最具代表性。

#### 3. 適切性

使用常模解釋分數，必須了解常模樣本性質是否和測驗對象相似，測驗對象背景條件如學業性向、教育經驗、文化背景、年齡是否和常模團體相近。總結任何測驗分數的正確解釋，均須顧及常模團體性質與常模類型。

四、學科診斷測驗相較於一般學科成就測驗，有何特色？教師在解釋學科診斷測驗的結果時，應注意那些事項？

【擬答】：

### (一)學科診斷測驗特色

#### 1. 意義

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。採用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. 診斷測驗與一般學科成就測驗相異處

	診斷測驗	一般學科成就測驗
相同處	(1)皆成就測驗一環 (2)皆可了解學生學習困難	
相異處	(1)較多分測驗，更多題目測量每項能力。 (2)測驗難度較低，只做為診斷學生錯誤所在。 (3)測驗題目無難度、有層次之分。	(1)大都是一份測驗，較少有分測驗，題目較少。 (2)測驗難度較高，做為區分學生學習成就。 (3)測驗題目有難度、無層次之分。

## 公職王歷屆試題（98 地方政府特考）

### （二）使用診斷測驗注意事項（郭生玉，民 95）

1. 診斷測驗目的在發現學習缺失，但無益於指出學習熟練程度，適用於平均水準以下學生。
2. 診斷測驗可以推論學生錯誤類型，但無法指出錯誤原因所在，因為原因常是多元複雜。
3. 診斷測驗僅提供學生學習困難部份資料，尚須考量學生其他因素（如年齡、智力、文化不公平、情緒…等）。
4. 診斷測驗結果信度較低，大部份診斷錯誤題目都較少。

公  
職  
王