

101 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、(一)何謂標準分數？其公式為何？

(二)將原始分數標準化之目的為何？

(三)舉例說明將偏態分布的分數加以常態化之目的為何？

【擬答】：

(一)標準分數

1. 表示某一測驗分數在團體平均數以上或以下多少個標準差；將原始分數加以轉換成可以互相比較之分數。當二變數欲比較大小，同時考量到不同單位、不同平均水準與不同變異（離散）程度時，就可採用標準分數。係以平均數為參照點，以標準差為單位的相對地位量數。

2. 轉換方式有直線轉換，稱直線標準分數；非直線轉換，稱為常態化標準分數。

$$\text{定義：} Z = \frac{X - \bar{X}}{S} = \frac{X - \mu}{\sigma}$$

(二)原始分數標準化之目的

原始分數本身沒有意義，將原始分數轉換為衍生分數（Derived Score）參照測驗常模表，以瞭解個人的測驗分數在團體中相對地位。主要目的有二：

1. 表示受試者測驗分數在常模團體中的相對位置，以表個別間比較（Interindividual Comparison）（大都指一般同儕間比較）。

2. 提供相同量數比較，以直接比較個人在不同測驗的分數，顯示個人內比較（Intraindividual Comparison）（使用側面圖解釋）。

(三)將偏態分布的分數加以常態化之目的

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。常態化標準分數適用時機為二個分數分配形狀很明顯不同，為了使不同分配型態分數可以直接比較，就必須將原始分數轉換為符合常態化分配的標準分數；且依據大而有代表性樣本，有理由相信非常態的原因是測驗缺陷所造成，或影響測量行為的其他因素導致；常態分配曲線擁有許多數量特徵，對於分數解釋與數學運算有很大助益。

二、(一)何謂變異數？其公式為何？

(二)從教育理論觀點而言，我們期望在班級團體的學科成績與大型升學考試科目成績上之兩種變異數的比較：

1. 何者變異數越大越好，其理由為何？

2. 何者變異數越小越好，其理由為何？

【擬答】：

(一)1. 變異數

變異數（Variance）係一群數值與其算術平均數之差異平方和的平均數，稱為變異數，又稱變方或均方變異數，開根號即為標準差。二者為用途最廣的變異量數。當群體中標準差愈小，即表群體中大部分數值集中於平均數附近，則平均數代表性強；相反地，若標準差大則表大部分數值比較分散，平均數代表性較弱。

2. 公式

$$(1) \text{未分組資料：} \sigma_x^2 = \frac{\sum(X - \bar{X})^2}{N}$$

$$(2) \text{分組資料: } \sigma_x^2 = \frac{\sum f(X - \bar{X})^2}{N}$$

- (二) 1. 大型升學考試科目成績變異數越大越好，因為大型升學考試科目必須競逐高低區辨程度，其學習範圍較廣，每一範圍試題較少，強調試題的鑑別力區辨個別差異。
2. 班級團體的學科成績變異數越小越好，因為班級團體學科成績大都關心學生精熟情形個體表現，其學習範圍較窄，每一範圍試題較多，強調試題在學生學習工作表現。

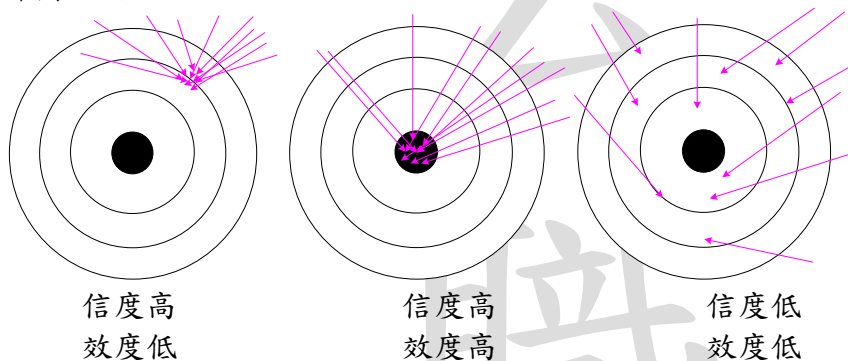
三、請以射箭或打靶為例(紅心表示所欲測量之心理特質)，作圖說明測驗之信度與效度的關係。

【擬答】：

(一) 信度是效度必要條件而非充分條件：

1. 信度是效度必要條件而非充分條件，效度是測量的首要考量條件，而信度則是效度的輔助品。二者可以推論為效度高，信度必高；效度低，信度不一定低；信度高，效度不一定高；信度低，效度一定低。
2. 效度需要測驗分數的正確性與一致性，但信度只需一致性即可。

(二) 作圖說明



四、(一) 何謂內部一致性信度？

(二) 內部一致性信度可支持何種效度？其理由為何？

【擬答】：

(一) 內部一致性信度 (Internal Consistency Reliability Coefficient)

重測信度與複本信度測驗，均需對受試者實施二次測驗，不僅增加編製的負擔，同時受試者合作意願降低、動機減少、疲勞與厭煩增加，均會影響施測的品質；內部一致性係數是由一個題本一次測量結果估計信度，亦即同一測驗之測量結果內各試題間的相關係數。因此學者嘗試一次施測產生正確估計信度結果，會較受歡迎；此方法包括：①折半信度；②庫李方法；③ α 係數方法；④霍意特變異數分析法。

(二) 1. 內部一致性分析法：類似效標關聯效度分析方式，以測驗總分為「內在效標」，逐一分析個別試題與總分間的相關係數。一般人格測驗皆用此方法。

2. 理由

- (1) 相關分析法：新編測驗和公認有效相似測驗之相關，常用來做為新編測驗，也能測出相關特質的依據。計算每一試題的二元化（及格或不及格）與總分的二系列相關或點二系列相關，凡相關達顯著水準，予以保留否則淘汰；若分測驗與總分相關太低，即表二者並不一致，應該予以刪除。
- (2) 團體對照法：將學生測驗總分，分高低二組，比較二組學生答對試題百分比。經統計考驗，若高分組答對試題百分比顯著高於低分組答對試題百分比，即表示試題有較高的內部一致性，反之，試題無效應刪除。

公職王歷屆試題 (101 地方政府特考)

五、某人要在甲乙丙丁四種智力測驗中選一種較穩定者來使用，他檢查四個測驗的 Mean，SD 與信度如下表。

	Mean	信度	SD
甲	100	.91	15
乙	100	.84	10
丙	100	.75	10
丁	100	.64	5

請問：

(一)那一個測驗在解釋個人分數時變動最小？理由為何？

(二)某生在上述變動最小的測驗上之 IQ 如為 110，試問在 95%信賴水準下，其 IQ 之範圍為何？

【擬答】：

(一)丁測驗。因為丁測驗的測量標準誤最小，在解釋個人分數時變動最小。

$$\text{甲測驗 } SE_{meas} = S_x \sqrt{1-r_{XX}} = 20\sqrt{1-.91} = 6$$

$$\text{乙測驗 } SE_{meas} = S_x \sqrt{1-r_{XX}} = 10\sqrt{1-.84} = 4$$

$$\text{丙測驗 } SE_{meas} = S_x \sqrt{1-r_{XX}} = 10\sqrt{1-.75} = 5$$

$$\text{丁測驗 } SE_{meas} = S_x \sqrt{1-r_{XX}} = 5\sqrt{1-.64} = 3$$

(二) $X - Z_{.975} \cdot SE_{meas} \leq t \leq X + Z_{.975} \cdot SE_{meas}$

$$110 - 1.96 \cdot 3 \leq t \leq 110 + 1.96 \cdot 3$$

$$104.12 \leq t \leq 115.88$$

表 IQ 測驗真實得分介於 104.12 與 115.88 間，可以信賴程度為 95%。惟犯錯機率 5%。

六、隨機自小幼生中抽 60 人之樣本，要他們在三種色紙選一種最喜歡的，結果人數為紅(30)、黃(20)、藍(10)，請利用所附的資料及空表格，考驗「小幼生對顏色有偏好」的假設。

(一)寫出統計假設： H_0 與 H_1 。

(二)訂 $\alpha = .05$ ，本題要採單側或雙側檢定的理由為何？

(三)計算之 $\chi^2 = ?$

(四) $df = ?$

(五) χ^2 的臨界值為 5.99，請問計算之 χ^2 有無顯著？

(六)統計裁決結果為何？

	人數	期望值			
紅	30				
黃	20				
藍	10				

(此表格僅為方便應考人計算之用，不列入評分項目，亦不須謄寫於試卷上。)

【擬答】：

(一) $H_0: P_1 = P_2 = P_3$

$H_1: P_1$ 不全等

(二) $\alpha = .05$ ，本題要採單側檢定。當實際次數與理論次數的差距愈大，亦即二者人數相差愈大，則其 χ^2 值也將會愈大，此時若 χ^2 值落入拒絕區，研究結果即拒絕虛無假設，即實際次數與理論次數存有顯著差異。 χ^2 考驗通常只用雙側考驗，我們常視 χ^2 、ANOVA 為單尾右側考驗。

$$(三) \chi^2 = \sum \frac{(f_o - f_e)^2}{f_e} = \frac{(30-20)^2}{20} + \frac{(20-20)^2}{20} + \frac{(10-20)^2}{20} = 10$$

公職王歷屆試題 (101 地方政府特考)

	紅	黃	藍
觀察人數	30	20	10
理論次數	20	20	20

(四) $df = K - 1 = 3 - 1 = 2$

(五) $\chi^2 = 10 > \chi_{.95(2)}^2 = 5.99$ ，落入拒絕區，拒絕虛無假設有無顯著。

(六) $\chi^2 = 10 > \chi_{.95(2)}^2 = 5.99$ ，落入拒絕區，拒絕虛無假設，亦即有充分證據支持小幼生對顏色有偏好人數有顯著差異，惟此一裁決犯型 I 錯誤 5% 的機率。

公
職
王