

101 年特種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、某教授想瞭解班上 100 位學生是否精熟，他請了 A、B 兩位評分員協助評定學生成就，評定結果如下：

	A 評分員	
B 評分員	精熟	不精熟
精熟	40	20
不精熟	10	30

(一)有那些方法可以瞭解此測驗的信度？

(二)請使用這些方法，計算此測驗的信度。

【擬答】：

(一)標準參照測驗的信度估計「決定」的重要性，採用下列二種方法：

1. 百分比一致性 (Percent Agreement, P_A)：又稱符合度，分析前後二次分類結果是否為一致的一種簡單統計方法，並用百分比之和來表示。當分類決定愈一致，即表示採用分類標準 (效標) 愈適當。適用類別變數需要作歸類評定。

2. 柯恆 K 係數：學者 Cohen 在 1960 年提出，目的在消除 P_A 中團體成份的機會影響，以估計由測量程序造成決定的一致性。亦即實際評定一致性百分比與評分者在理論上評定一致性最大可能次數百分比之比率。

$$(二) 1. P_A = \frac{A+D}{N} = \frac{40+30}{100} = \frac{70}{100} = 0.7$$

∴表示分類決定的一致性是 70%，即學生成就測驗的信度 0.7

$$2. P_C = \frac{(a+b)(a+c)}{N^2} + \frac{(c+d)(b+d)}{N^2} = \frac{50 \times 60}{100 \times 100} + \frac{50 \times 40}{100 \times 100} = 0.5$$

$$\therefore K = \frac{P_A - P_C}{1 - P_C} = \frac{0.7 - 0.5}{1 - 0.5} = 0.4$$

表示在團體中，依團體組成性質的影響，可預期有 40% 一致性分數。將此一影響去除， $K = 0.4$ ，表除團體組織或預期影響外，測驗可貢獻 40% 一致性分類。

二、(一)如果想要計算非對即錯試題的難度，可使用那些指數？請說明這些指數的意義。

(二)如果想要計算非對即錯試題的鑑別度，可使用那些指數？請說明這些指數的意義。

【擬答】：

(一)難度分析：

1. 答對百分比分析：計算全體受試者答對每個試題的人數占全體總人數的百分比值，這個百分比值稱為「難度指標 (數)」 (Item Difficulty Index)；通過百分比多代表這個題目簡單，通過百分比少代表這個題目較難。

$$\text{公式：} P = \frac{R}{N} \text{ 或 } P = \frac{P_H + P_L}{2}$$

2. 等距量尺分析：

美國教育測驗服務社 (Educational Testing Service, 簡稱 ETS) 創用，以 13 為平均數，4 為標準差，以 Δ (Delta) 代表試題難度指標。下限為 1，上限為 25。

$$\text{公式：} \Delta = 13 + 4Z$$

(二)鑑別力分析：

鑑別力分析的目的，主要在瞭解試題是否具備區別學生能力高低的作用。某個試題鑑別度愈高，表示能明確分辨學生答對與答錯功能很強，亦即高分組會傾向答對，低分組會傾向答錯，試題具備此種辨別作用，即為「試題鑑別度」。一個優良的測驗試題，具有較高鑑

公職王歷屆試題 (101 地方政府特考)

別度，表示其區辨力愈高。其方法分為三類：

1. 內部一致性 (Internal Consistency)：探討個別試題得分與整份測驗總分間是否合乎一致性，目的在使測驗的內部一致性變得最大。

教師自編測驗較難尋找合適的外在效標，此種分析可以改進內容效度與建構效度，但無法增加效標關聯效度，又稱諧度分析。

$$\text{公式：} D = P_H - P_L$$

2. 外在效度分析：

(1) 分析學生在試題上的反應與在效標上表現之間的關係，目的在試題反應與測驗的外在效度變得最大，用來增強效標關聯效度。

(2) 根據學生在某題試題反應答對或答錯與測驗總分相關係數，此一相關係數即用來表示試題鑑別度指標；點二系列相關係數值愈高，代表試題的作用與測驗總分的作用間愈一致，其鑑別功能愈強。即一個「二分變項」（試題分成答對與答錯）與一個「連續變項」（外在效標）。其常用方法：

① 點二系列相關：適用試題二分變項（答對與答錯、錄取與不錄取、通過與未通過），而效標屬於連續變項（0~100分）。

當 r_{pb} 為正值且越接近 1 鑑別度越佳， $r_{pb}=0$ 代表題目鑑別度是 0。

② 二系列相關法：適用情況於學生在試題反應呈常態分配，卻因某種需求與原因，以人為方式將其分為答對與答錯二種情形。

3. ϕ 相關亦可使用 ϕ 相關 (Phi Correlation)，適用在試題與效標均為二分變項情況。

$$\text{公式：} \phi = \sqrt{\frac{2}{n}}$$

三、(一)請舉例說明正偏態(positively skewed)的意義。

(二)請舉例說明矩形分配(rectangular distribution)的意義。

(三)請舉例說明二項式分配(binomial distribution)的意義。

(四)請舉例說明二項式分配母群(population)的平均數與標準差分別為何？

(五)變異數分析的結果若達到統計的顯著性，可接著做事後比較，事後比較包含簡單比較與複雜比較，請解釋何謂簡單比較？

【擬答】：

(一)偏態 (Skewness; SK, g_1 , β_1) 不對稱的分配即是有偏態，指群體中大部分的分數落於平均數那一邊；亦即次數分配不對稱或平均數與中位數落在兩個不同的點，此時重心失去平衡，曲線自左向右傾斜或自右向左傾斜。正偏態分配，即表群體中較多分數集中在低分方面，亦或平均數落於中位數右邊。

$$\text{公式：} SK = g_1 = \frac{m_3}{m_2 \sqrt{m_2}} = \frac{m_3}{S^3} > 0$$

(二)1. 均勻分配定義：設一連續隨機變數 X ，其值域為 $a \leq X \leq b$ ，且在區間 $[a, b]$ 中各數值發生機率均相同，則該分配為一均等連續分配。例如等車時間、飛機飛行時間、顧客排隊買票時間均屬於均勻分配。

2. 均勻分配機率函數：

$$\text{機率函數：} f(X) = \frac{1}{b-a} \quad a \leq X \leq b$$

3. 均勻分配的期望值與變異數：

$$\text{期望值 } E(X) = \frac{b+a}{2} ; \text{ 變異數 } V(X) = \frac{(b-a)^2}{12}$$

(三)1. 意義：若一隨機試驗試行 n 次，每次試驗只有成功與失敗二種結果，而且每次成功機率 p 均不變，我們稱此為二項試驗。投擲一枚均勻硬幣產生的機率分配。

2. 二項式機率函數：

$$\text{機率函數：} f(x) = C_x^n p^x q^{n-x} \quad x = 0, 1, \dots, n$$

公職王歷屆試題 (101 地方政府特考)

$$0 \leq p \leq 1, q = 1 - p; n \text{ 為正整數}$$

(四)二項式分配母群的期望值 $E(X) = np$ ；變異數 $V(X) = npq$

※ 舉例：16 題是非題完全猜測

$$\text{期望值 } E(X) = np = 16 \times \frac{1}{2} = 8; \text{ 變異數 } V(X) = npq = 16 \times \frac{1}{2} \times \frac{1}{2} = 4$$

(五)在探索性資料分析中，研究者通常無法知道處理實驗水準那幾對平均數有顯著差異，因此，在 F 考驗值達到顯著水準時，便須進行事後比較，以確定那幾對平均數間有差異存在。進行事後比較須針對所有實驗處理 C_2^k 個比較，大部分屬於非正交比較。簡單比較（每次進行二個平均數差異比較），進行方法有①薛費法（Scheffé Method）；②杜凱法（Tukey Method）；③紐曼—柯爾法（Newman-Keuls Method）。

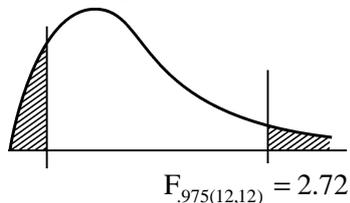
四、某研究者想瞭解 A、B 兩班的變異程度是否不同，而 A 班共有 13 人，A 班的標準差是 20，B 班共有 20 人，標準差是 10。請使用假設考驗的步驟，考驗 A、B 兩班的變異程度不同的假設，並解釋研究結果。（ $F_{.975(12,12)} = 2.72$ ）

【擬答】：

1. $H_0: \sigma_1^2 = \sigma_2^2$

$H_1: \sigma_1^2 \neq \sigma_2^2$

2. $\alpha = 0.05$



3. 計算： $F = \frac{\hat{S}_1^2}{\hat{S}_2^2} \times \frac{\sigma_2^2}{\sigma_1^2} = \frac{40}{10} \times 1 = 4$

4. 結論： $F = 4 > F_{.975(12,12)} = 2.72$ ，落入拒絕區，拒絕虛無假設，亦即有充分證據支持 A、B 兩班變異數有顯著差異。唯此一裁決犯第一類型錯誤概率 0.05。