

102 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、測驗是研究者用來收集受試者作答反應資料的工具，請問：

(一)何謂測驗的信度(reliability)?

(二)常模參照測驗(norm-referenced test)的信度係數可分成那幾種?

(三)承(二)，每一種信度係數的主要誤差來源各為何?

(四)常在公務人員考試中使用測驗式試題，如果考試後想進行試題分析的話，它最適合使用那一種信度係數作為分析指標?

【擬答】：

(一)信度意義

1. 信度是指相同受試者在不同時間，使用相同測驗測量（或複本測驗測量多次）或在不同情境下測量，所得測驗結果穩定性；不同測驗分數一致性、不同評分者間評分一致性、評量結果精確性，藉以反應真實量數程度的一種指標。二次測量結果相當一致，就表示測量分數可靠性高、穩定性高或具有可預測性。一份優良教育測驗信度至少要在.80 以上的信度係數才有使用價值。

2. 定義公式： $r_{xx} = \frac{S_t^2}{S_x^2}$ 或 $r_{xx} = \frac{S_x^2}{S_x^2} - \frac{S_e^2}{S_x^2} = 1 - \frac{S_e^2}{S_x^2}$ (信度 = 1 - 誤差)

(二)最常用常模參照測驗信度之估計方法有四種：①重測信度；②複本同時信度與複本間隔信度；③內部一致性信度（折半信度、庫李信度、 α 係數）；④評分者信度。(三)重測信度—時間誤差；複本同時信度—試題取樣誤差；複本間隔信度—試題取樣誤差+時間誤差；內部一致性信度（折半信度—試題取樣誤差、庫李信度—試題取樣誤差+試題異質性誤差、 α 係數—試題取樣誤差+試題異質性誤差）；評分者信度—評分者誤差。(四)內部一致性信度（折半信度、庫李信度、 α 係數）。

二、某一位研究員在針對某次測驗後進行試題分析時，獲得下表有關某一試題的各選項選答資料。

組別	選項				
	*A	B	C	D	未答
高分組	20	10	50	20	0
低分組	10	25	25	40	0

註：細格中的數字表示「人數」，*表示「正確答案的選項」

請參考上表，逐一回答下列問題：

(一)請計算該試題的難度指標值？

(二)請計算該試題的鑑別度指標值？

(三)請評估該試題是否具有理想的選項誘答力？

(四)請評鑑該試題的優劣，並說明你的結論。

(五)假如該名研究員發現該試題的答案可能是錯誤的。基於該測驗中每一試題都具有優良試題特質的前提下，請問該試題的答正確答案應該是那一個選項才較為合理？請說明你的理由。

【擬答】：

$$(一) P = \frac{P_H + P_L}{2} = \frac{\frac{20}{100} + \frac{10}{100}}{2} = 0.15$$

$$(二) D = P_H - P_L = \frac{20}{100} - \frac{10}{100} = 0.1$$

公職王歷屆試題 (102 地方政府特考)

- (三) 選項 B、D 有誘答力，因為選擇錯誤選項的低分組學生人數應比高分組學生還多；選項 C 無誘答力需修改，因為選擇錯誤選項的低分組學生人數應比高分組學生還少，也是不佳題目（吸引高分組學生作答、可能題目不清、爭議答案、作答粗心）。
- (四) 此題 P 值愈小，難度愈高；美國測驗專家 Ebel 在 1979 年提出一套鑑別度評鑑標準，當 $D = .1$ 表示試題劣、應該淘汰。

(五) 答案應為 C。因為 $P = \frac{P_H + P_L}{2} = \frac{50}{100} + \frac{25}{100} = 0.375$ ， $D = P_H - P_L = \frac{50}{100} - \frac{25}{100} = 0.25$ ，題目中優良試題特質前提其鑑別度 0.4 以上、難度為 0.5 時，以選項 C 較為接近優良試題條件。

三、在教育統計學中，某研究員若想檢定「數學考試焦慮是否存在著性別差異」的假設，他該怎麼做，才比較恰當？請說明你的檢定步驟。

【擬答】：

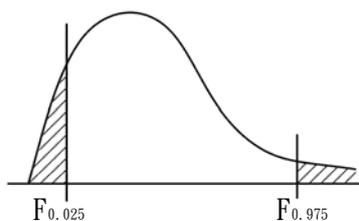
同學解此類考題宜特別注意。題目中母群體 σ_1^2 與 σ_2^2 均未知，且樣本 n_1, n_2 也為給條件，且二群體為獨立樣本（性別）。同學進行考驗需分二步驟進行(1)先考驗 $\sigma_1^2 = \sigma_2^2$ 。(2)再決定使用考驗平均差方法。

(一) 假設考驗：【步驟一】變異同質性考驗：

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

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

2. 拒絕區 $\alpha = 0.05$



查表 $F_{0.975; df_1, df_2}$

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

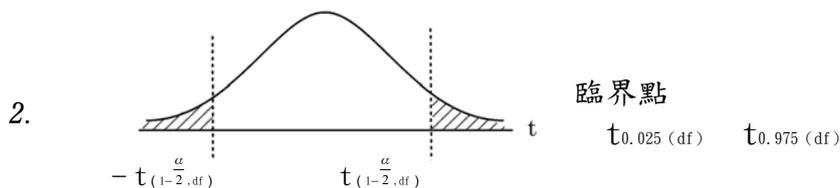
4. 結論： $F < F_{0.975; df_1, df_2}$ ，落入接受區，接受虛無假設，亦即接受二組變異數同質 $\sigma_1^2 = \sigma_2^2$ 。

反之 $F > F_{0.975; df_1, df_2}$ ，落入拒絕區，拒絕虛無假設，亦即接受二組變異數異質 $\sigma_1^2 \neq \sigma_2^2$

(二) 【步驟二】變異數同質 $\sigma_1^2 = \sigma_2^2$ 使用獨立樣本 t-test：

1. $H_0: \mu_1 = \mu_2$

$H_1: \mu_1 \neq \mu_2$



3. 計算：

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_P^2}{n_1} + \frac{S_P^2}{n_2}}}$$

4. 結果 $t > t_{.975 (df)}$ ，落入拒絕區，拒絕虛無假設亦即有充分證據支持不同性別在數學考試焦

公職王歷屆試題 (102 地方政府特考)

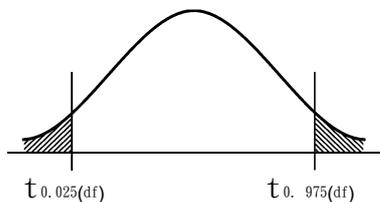
慮有顯著差異。惟結論仍有 0.05 機率犯型 I 錯誤。反之 $t < t_{.975(df)}$ ，落入接受區，接受虛無假設亦即無充分證據支持不同性別在數學考試焦慮有顯著差異。

(三)【步驟三】變異數異質 $\sigma_1^2 \neq \sigma_2^2$ 使用無母數統計學 t' -test：

1. $H_0: \mu_1 - \mu_2 = 0$

$H_1: \mu_1 - \mu_2 \neq 0$

2. $\alpha = 0.05$



3. 計算：
$$t' = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\hat{S}_1^2}{n_1} + \frac{\hat{S}_2^2}{n_2}}}$$

4. 結論： $t' > t_{.975(df)}$ ，落入拒絕區，拒絕虛無假設亦即有充分證據支持不同性別在數學考試焦慮有顯著差異。惟結論仍有 0.05 機率犯型 I 錯誤。反之 $t' < t_{.975(df)}$ ，落入接受區，接受虛無假設亦即無充分證據支持不同性別在數學考試焦慮有顯著差異。

四、參考下列某個實驗結果的電腦報表數據：

SV	SS	df	MS	F	p
Between group	1200	2	600	300.0	.000
Within group	50	25	2		
Total	1250	27			

請回答下列問題：

- (一)總共有幾名受試者參與實驗？
- (二)這些受試者共分成幾組進行實驗？
- (三)本實驗的虛無假設應該如何表示？
- (四)本實驗的統計檢定結果為何？
- (五)請你解釋這個實驗的結論。

【擬答】：

(一)28 人。

(二)3 組。

(三) $H_0: \mu_1 = \mu_2 = \mu_3$

$H_1: \mu_i$ 不全等

(四) $F = 300$ ， p 值 = .000 < $\alpha = 0.05$ 落入拒絕區。

(五)拒絕虛無假設，亦即有充分證據支持三種實驗處理對依變項之影響達顯著水準。惟此判決犯型 I 錯誤 5% 的機率。至於到底那二組平均數有差異，則需進行事後比較（使用薛費法 (Scheffé Method)、杜凱法 (Tukey Method)、紐曼—柯爾法 (Newman-Keuls Method)）。