

97 年公務人員特種考試身心障礙人員考試試題

等別：四等考試

類科：教育行政

科目：教育測驗與統計概要

一、建民最近參加教育行政網路模擬會考，其中三科的原始分數及其所對應的百分等級 (PR) 如下：國文72分 (PR=16)；英文60分 (PR=50)；教育測驗與統計52分 (PR=84) 依據上述資訊並參酌下列不同標準分數與百分等級對照表，回答下列四個問題：

不同標準分數百分等級對照表

Z分數	百分等級	T分數	新量尺分數
-2	2	30	70
-1	16	40	85
0	50	50	100
1	84	60	115
2	98	70	130

(一)建民在那個科目的表現較好？請說明你的理由。

(二)建民在英文的T分數為何？

(三)建民在教育測驗與統計的新量尺分數是在常模樣本平均數之上或之下的幾個標準差處？

(四)已知三科的信度都是0.91，那一個科目在新量尺分數的測量標準誤最小？理由為何？

【擬答】：

(一)1.建民的教育測驗與統計表現較好。

2.因為百分等級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。例如建民的教育測驗與統計得分52分在團體中100人勝過84人，而國文與英文勝過人數較少。

(二)建民英文的Z分數=0 (因為PR=50)， $T=10Z+50=10 \times 0+50=50$ 。

(三)教育測驗與統計的PR=84其Z=1，表示量尺分數是在常模樣本平均數上1個標準。

(四)英文科的測量標準誤會最小。當信度固定時，因為其PR=50，分數位於常模位置其測量誤差會較小，而國文與教育測驗與統計因為分數位於常態分配兩端其測量標準誤會較大。

二、已知某個選擇題測驗共有20題，每題各有四個選項，但只有一個正確答案。假設某位考生只能正確回答其中14題，其他6題則以隨機方式猜測作答。今以X表示該考生在此測驗答對的題數。

(一)請以公式表示該考生答對題數 (X) 之分配。

(二)試問X大於等於19的機率為何【 $P[X \geq 19] = ?$ 】

(三)試問X的平均數為何？

(四)試問X的變異數為何？

【擬答】：

(一)1.考生答對題數之分配為二項式機率函數：

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

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

2.二項分配的期望值與變異數：

(1)期望值 $E(X) = np$

(2)變異數 $V(X) = npq$

$$(二) P[X \geq 19] = C_5^6 \left(\frac{1}{4}\right)^5 \left(\frac{3}{4}\right)^1 + C_6^6 \left(\frac{1}{4}\right)^6 \left(\frac{3}{4}\right)^0 = \frac{18}{4096} + \frac{1}{4096} = \frac{19}{4096}$$

$$(三) E(X) = np = 6 \times \frac{1}{4} = 1.5$$

公職王歷屆試題 (97 身心障礙特考)

$$(四) V(X) = npq = 6 \times \frac{1}{4} \times \frac{3}{4} = 1.125$$

三、請問你是否同意以下各敘述？若不同意請說明理由。

- (一)描述統計可用於母群資料。
- (二)推論統計目的在瞭解樣本資料的特性。
- (三) α 越高犯第一類型錯誤的機率越大。
- (四) β 越高犯第二類型錯誤的機率越大。
- (五)檢定力 (Power) 越高犯第二類型錯誤的機率越大。

【擬答】：

(一)不同意。

因為敘述統計學 (Descriptive Statistics) 不只探討母群資料亦可探討樣本的個別特性，對研究資料之處理主要在蒐集、整理劃記、描述與表現結果，即將一群資料加以整理、摘要、組織與簡化，使讀者容易明瞭其中意義與傳遞訊息。內容包括集中趨勢、變異趨勢、圖形表示、相對地位量數等。

(二)不同意。

係探討母體與樣本之間的關係，根據樣本的資料推論未知母體性質 (非樣本性質)。當研究者欲瞭解母體某些特性，但母體群過於龐大，無法將母群體全部蒐集，最好辦法為經由隨機 (Random) 方式，自群體抽取若干具有代表性 (Representative) 個體為樣本，再依機率原理，經由假設、估計、考驗步驟，使用樣本統計量從事估計與考驗母體的方法。內容包括機率分配、抽樣方法、估計、考驗、分配、相關與迴歸。

(三)同意。

在心理與教育統計中，習慣上採用 $\alpha = 0.05$ 或 $\alpha = 0.01$ 作為假設檢定的顯著水準， α 值大小主要視犯型 I 錯誤造成損失與後果，及研究者經驗態度過去文獻而決定，在某些情況下 α 亦可增減。有時亦稱危險區或拒絕區。犯型 I 錯誤機率愈小，即 α 愈小，研究結果顯著水準愈高；犯型 I 錯誤機率愈大，即 α 愈大，研究結果顯著水準愈低。

(四)同意。

因為當虛無假設 H_0 為假，卻因統計考驗結果接受了 H_0 ，認為 H_0 是對的，稱為型 II 錯誤。其中 β 風險是指犯型 II 錯誤機率大小，因此 β 越高犯第二類型錯誤的機率越大。

(五)不同意。

當其他條件不變情況下，統計考驗力越高，則第二類型錯誤 β 機率會愈小；因為 β 與檢定力 $(1 - \beta)$ 是在虛無假設為假的前提下討論，兩者前提相同因此兩者和一定為 1，兩者具有相互消長特性。

四、(一)請說明題目分析的功用何在？

(二)下表為某測驗的四個問題之題目分析結果，就題目分析的結果，回答下列問題：

1. 就難度而言，下列那一題最困難？那一題最簡單？
2. 下列那一題的鑑別度最高？那一題鑑別度最低？
3. 就選項分析而言，那一題最需要修改？理由為何？

第一題		選項1	選項2	選項3	選項4*	P	D
	高分組	.04	.04	0	.92	.57	.71
	低分組	.25	.15	.39	.21		
第二題		選項1*	選項2	選項3	選項4	P	D
	高分組	.96	0	.04	0	.87	.17
	低分組	.79	.07	.07	.07		
第三題		選項1	選項2	選項3	選項4*	P	D
	高分組	.00	.00	.04	.96	.72	.48
	低分組	.00	.00	.52	.48		

公職王歷屆試題 (97 身心障礙特考)

第四題		選項1	選項2	選項3*	選項4	P	D
	高分組	.00	.00	1.00	.00	.90	.20
	低分組	.09	.07	.80	.04		

註：選項中加*者表示其為正確答案

【擬答】：

(一)信度與效度是測驗二項最重要的特徵，二項特徵均須視試題品質優劣而定，而試題品質則可藉由試題分析而提高。試題分析包括：

1. 品質分析 (Qualitative Analysis)

可由試題的內容效度審查、有效命題原則評鑑、教學目標進行邏輯的分析（亦即符合編製測驗程序品質要求），即可達到目的。品質分析包括以主觀經驗與個人專業素養，主觀地判斷題目數是否合理、文字表達是否清晰、錯誤選項安排是否合適。

2. 量化分析 (Quantitative Analysis)

量化目的主要在分析每題試題所具備三大統計特徵—難度 (Difficulty)、鑑別度 (Discrimination)、誘答力 (Distraction)。所有心理測驗中的試題必須經質與量二方面分析，才能顯現其特性與有效程度高低，據此修改或刪除試題，則試題才是一套可靠有效的測驗。試題分析功能

- (1) 提供回饋給學生，作為學生改進學習方法的參考、激勵學生學習表現、了解教師教學目標與調整自我學習方法。
- (2) 提供回饋給教師，作為實施團體或個別補救教學的依據，卻並學生學習錯誤類型具有診斷作用。
- (3) 提供改進班級課程教學建議依據，依據試題分析結果調整教材內容順序，或降低教材難度。
- (4) 增強教師編製測驗的技巧，教師運用試題分析可以釐清試題題意是否不清、是否有暗示線索、誘答項目是否具有吸引力。
- (5) 增進測驗題庫運用的效能。
- (6) 可以縮短測驗長度，節省測驗時間，提高信度與效度。

- (二) 1. 第一題最困難，因為 P 值 = .57 最小。第四題最簡單，因為 P 值 = .9 最大。
 2. 第一題鑑別度最高，因為 D 值 = .71 最大。第四題最簡單，因為 P 值 = .9 最大。
 3. 第三題選項 1 與選項 2 最需要修改。選項分析判斷原則每一錯誤選項至少有一低分組學生選擇、選擇錯誤選項的低分組學生人數應比高分組學生還多，因為第三題選項 1 與選項 2 均無人作答，顯示其無誘答作用。