

99 年公務人員特種考試身心障礙人員考試試題

等 別：三等考試

類 科：教育行政

科 目：教育測驗與統計

一、何謂試題分析 (item analysis)？它的內容包括那些？試舉出一種可以進行試題分析的電腦軟體程式名稱？

【擬答】：

(一)試題品質則可藉由試題分析而提高，試題分析包括：①依試題內容與形式做品質分析；②依試題統計特性做量化分析：

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。
1. 分析題目特性：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。

3. 特性主要指標與計算方式：

(1) 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。

$$\text{公式：} \Delta = 13 + 4Z \text{ [} \Delta \text{ (Delta) 難度指標]}$$

(2) 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

公職王歷屆試題 (99 身心障礙特考)

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：點二系列相關、 ϕ 相關、二系列相關法。

(二) 試題分析的電腦軟體程式名稱

常見測驗邊製套裝軟體如美國 Assessment Systems Corporation 於 1985 年發展的 MicroCAT，主要是由 ITEMAN、RASCAL、ASCAL 等軟體組成。此系統功能包括提供古典測驗理論與試題反應理論試題參數估計、建立題庫、IRT 受試者能力估計、自動編製測驗與建立適性測驗。

二、解釋名詞：

- (一) KR20 信度 (KR-20 reliability)
- (二) 同時效度 (concurrent validity)
- (三) 計分規準 (scoring rubrics)
- (四) 聚斂效度 (convergent validity)
- (五) 實作評量 (performance assessment)

【擬答】：

(一) 係由庫德 (G. F. Kuder) 和李查遜 (M. W. Richardson) 於 1937 年在《心理計量期刊》發表「測驗信度估計理論」，設計一種估計測驗項內部的一致性信度的方法，根據學生在所有試題作答反應，分析其試題間一致性，也只需實施一次測驗即可。僅適用答案非對即錯的二分計分測驗。分析題目之一致性，此題目間一致性 (Inter-item Consistency) 量數的高低受到兩種誤差變異量的影響，一為內容取樣，一為所取樣之行為領域的異質性，異質性愈低，則題目間之一致性也就愈高。

$$\text{定義公式：} r_{KR20} = \frac{n}{n-1} \left(1 - \frac{\sum pq}{S_x^2} \right)$$

(二) 效標關聯效度分為同時效度與預測效度二種，主要的區別方式在測驗使用目的不同與外在效標取得時間的不同。指測驗分數與外在效標取得在同一時間內連續完成，計算此二種資料的相關係數即為同時效度。目的在使用測驗分數估計個人在效標方面的實際表現。例如新編一國中數學成就測驗實施於一群國中生，以他們在校數學月考成績為效標求其相關，如同時效度適當，則可使用測驗分數瞭解學生的學習表現。

(三) 評分規程為實作評量評分方式之一，屬於分析式評分方式。學者專家發展各學科評分規程，增進實作評量效率；評分規程由二個向度構成表格，一個向度為一系列判斷規準，另一個向度則由三至四個層次精熟水準，二個向度交互細格中以文字說明每一精熟水準狀況。(歐滄和，民 93)

(四) 學者 Campbell 與 Fiske 在 1959 年提出一份良好的構念效度，須符合下列二條件聚斂效度高與區別效度低；聚斂效度 (輻合性效度) (Convergent Validity) 一份測驗分數必須和相同構念其他測量的分數間存在高相關，亦即不同方法測量相同特質。如美勞性向測驗的分數與在校美勞科目成績有高相關，即為聚斂效度。

(五) 屬於多元評量方式之一，實作評量是指一項工作被執行與完成，即評定學生在工作項目中實際行為表現，介於紙筆測驗與真實情境評量間，在模擬各種不同真實程度與直接性的測驗情境中，提供教師一套系統評量學生實作表現方式。可以是實際表現過程或表現後作品，或二者組合皆可，視實際表現領域的性質而定，又稱表現評量。在學科領域內屬於技能部分的學習，例如美術創作、音樂表現、戲劇表演、詩歌朗誦、體能狀況等方面，較適合用實作評量方式來檢測學習成效。

三、在教育統計學中，研究者常會使用迴歸分析 (regression analysis) 方法，作為建立預測方程式的依據。但在迴歸方程式中，迴歸係數 (regression coefficients) 是使用什麼方法估計出來的？該等迴歸係數具有何種特色 (至少列出三種)？

【擬答】：

公職王歷屆試題 (99 身心障礙特考)

(一)迴歸係數使用估計方法

最小平方法 (Least Square Method) 當 X 與 Y 變項有相關時，找出一條直線通過各點附近，將研究蒐集資料散布圖中的各點到這直線之平行於縱軸，且垂直於橫軸的距離平方和為最小的一種統計方法。這條稱為最適合線 (Best-fit Line) 或迴歸直線。即 $\sum(Y - \hat{Y})^2 = \sum(Y - a - bX)^2$ 為極小值。

求出 a 、 b ，使得誤差項平方最小， $\sum e^2 = \sum(Y - \hat{Y})^2 = \sum(Y - a - bX)^2$ ， e 為誤差。

(二)迴歸係數具有特色

1. 意義

根據 X 變項來預測 Y 變項的「迴歸係數」(Regression Coefficient)，斜率 b 表示迴歸線的傾斜度和方向，亦即效標變項 Y 的改變量與預測變項 X 改變量的比值。也可以 X 變項與 Y 變項之共變項除以 X 變項之變異數。

2. 特色

迴歸係數表示迴歸方程式預測方向性，如值為正值表示自變項對依變項影響為同方向，反之則相反；迴歸係數帶有單位非標準化統計量，可以反應自變項對依變項影響的數量，但由於單位差異無法進行相對比較。

四、假設某研究人員已獲得 10,000 名受試者的資料，並以該筆資料中的自變項 (X ，社經地位) 預測依變項 (Y ，學業成就)，且建立起一條標準化迴歸方程式如下：學業成就 = 0.50 (社經地位)。請問：

(一)學業成就和社經地位之間的相關係數是多少？

(二)社經地位能解釋學業成就的變異數百分比有多少？

(三)假設某生的社經地位標準分數為 2.0，那麼，他的學業成就標準分數約略可以贏過多少百分比的受試者？

【擬答】：

$$(一) \hat{Z}_Y = \beta \cdot Z_X = rZ_X = .5Z_X$$

學業成就和社經地位之間的相關係數 = .5

(二)即表「在學業成就的變異數中，可被社經地位解釋到變異量部分為 $.5^2 = 25\%$ ，即社經地位預測學業成就預測正確部分高達 25%」。

(三) $\hat{Z}_Y = \beta \cdot Z_X = rZ_X = .5 \times 2 = 1 \therefore PR = 84$ ，他的學業成就標準分數約略可以贏過 84% 百分比的受試者。