

98 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、試從主試者、受試者、測驗情境、測驗材料、計分程序五方面說明影響測驗分數的主要因素，以及如何控制或排除這些不利的影響，以維持測驗分數的信度？

【擬答】：

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，所得結果一致性。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。影響信度主要因素分述如下：

(一)主試者

1. 主試者測驗的實施程序為影響信度的重要因素，例如：測驗實施時未遵守測驗指導語、時間限制、給予學生權限外協助、座位安排不當、非標準化的測驗程序、主試者的偏頗與暗示、評分的主觀性等等。
2. 主試者評分過程產生月暈效應、評分前後寬鬆不一、邏輯誤差、偏見等因素進而影響評分一致性，透過嚴謹、適當主試者訓練，提高評分者間一致性。

(二)受試者

1. 團體分數變異量（樣本同質性）
在其他條件相等下，團體變異性愈大，其信度愈高。從相關係數來看，當團體內個別差異愈大，測驗分數分散範圍愈大，其相關係數也愈大，其測驗信度亦較高。由信度定義公式實得分數變異量愈大，表示團體間個別差異也愈大，若誤差分數變異量不變，則測驗結果信度就愈高。
2. 受試者身心狀況因素亦會影響測驗結果，包括動機、意願、情緒、態度、受測者的身心健康狀況、注意力、持久性、作答態度等變動。
3. 美國「全國專業教學標準協會」（NBPTS）提出當受試者充分地瞭解計分規準、知道評分者如何評定其作業時，有助於增加測驗評分一致性。

(三)測驗情境

1. 測驗情境中間隔時間長短會影響測驗信度，如重測信度與複本信度二次施測間隔時間愈短，信度愈高，反之，間隔時間愈長，其他變項介入可能性增加，信度降低。
2. 信度係數與所選用估計信度方法有很大關係，須視其試題間相關係數大小決定。選用信度估計方法不同，就有不同的測量誤差來源與不同大小估計值產生。
3. 測驗情境因素包括燈光、通風、安靜、氣氛、聲音、桌面、空間因素等也會對信度有直接與間接影響。因此施測前注意測驗環境因素，包括場所大小必須考量施測人數多寡，一般團體測驗場所應有足夠大的空間，而個人測驗其空間應較小；無論團體或個人測驗有充足光線、流通空氣、舒適溫度、避免噪音干擾等因素，使受試者在舒適環境中施測。

(四)測驗材料

1. 試題太難或太易均會使大多數學生答錯與答對，會產生二種極端情形（如正偏分配與負偏分配），使其測驗分數的變異程度較小，信度會較低。而測驗題目難易適中，可使測量特質的範圍較大，提高信度。

公職王歷屆試題 (98 地方政府特考)

2. 測驗之信度取決於試題的編製，是否有依雙向細目表來編製試題，如測驗的取材、測驗的長度（題數）、試題的難度、鑑別度以及編排方式、指導語不夠清楚、題意不清、題型等均會影響信度之高低。測驗本身任何缺點均會影響試題所欲測量的功能，而降低信度。
3. 測驗前應注意實施測驗時所有物件是否準備充分，包括答案紙、指導手冊、測驗題本、作答工具、計時設備、施測材料與工具，事先應檢查其完整性避免遺漏；且應事先熟悉如何使用材料與設備，避免臨場揣摩使用影響受試者心情。

(五) 計分程序

1. 測驗的評分如評分者個人主觀意見與判斷的影響，則不同評分者評分的結果含有很大評分差異，如此，測驗的信度係數則較低。因此客觀測驗如是非題、選擇題、配合題應有較高信度係數。
2. 計分程序中因不同計分規準界定特質的具體或明確程度不同，例如分析式計分規準叫整體式計分規準更為明確與具體，會導致評分者間一致性受到影響。

最後測驗的「標準化」亦為影響信度主因，包括(一)測驗發展程序的標準化—測驗編制過程有一致性規範，包括測驗目標、雙向細目表、設計與修審試題、抽樣與預試、分析試題以及建立測驗的信度與效度等。(二)施測過程的標準化，施測過程必須完全依照測驗指導說明手冊程序進行，有固定而且明確的指導語以及施測程序，係指亦即施測者不可任意更改規定的時間、手續、步驟、指導語，使每個受試者都能在相同的條件與情境之下進行測驗。(三)測驗計分、使用與解釋的標準化，是指分數的計算方法與結果的解釋原則，應該要遵循一定的要領與方法。測驗標準化的主要目的是為了使測驗能有較好的測量品質，降低測量的誤差，增加測驗分數的效度與信度。

二、試以條列方式列出標準化成就測驗的編製步驟，由「擬定測驗編製計畫」到「測驗出版社正式發行」，所列出的步驟要周延，先後順序要正確，但不必解釋執行的細節或目的。

【擬答】：

標準化成就測驗其編製過程步驟如下：

(一) 提出測驗發展計畫：

提出測驗發展計畫必須包括說明編製測驗目的與理由、標準化測驗描述（測驗特質、施測對象年齡與年籍、施測方式、施測時間）、試題、基本市場、競爭對象等。

(二) 評估測驗發展計畫：

出版商接到編製者發展測驗計畫，必須進行產品完整性、行銷市場、本出版社研究發展能力、生產能力等項目評估，是否有能力編製此標準化測驗。

(三) 確定測驗目的：

標準化成就測驗目的確定學生學習技能（如數學、語文……）、瞭解學生單一科目成就、一般成就、診斷學生學習困難。

(四) 設計雙向細目表：

依據教學目標與教材內容，同時參考學者、測驗專家、課程專家的意見，詳細研究分析，做為編製試題依據。

(五) 選擇題型與編擬試題：

題目需經預試，從中淘汰不良試題，編製試題應比預定題數多。此階段必須考量測驗涵蓋年齡與年籍、測驗長度與版本數、編寫試題者能力、試題題型。

(六) 預試和評鑑：

公職王歷屆試題 (98 地方政府特考)

預試對象最好採用將來正式測驗的應試群體，取樣需注意受試者代表性，一般大約 400 人，預試前必須考量預試經費、受試者年齡與年級範圍、預試樣本大小、施測方式，預試時施測環境、施測程序、時間限制……等儘量與正式施測一致；同時施測過程對受試者反應、作答時間、題意不清均需詳加記錄，提供日後參考。

(七)編輯測驗—建立信度、效度、常模：

預試完畢後，根據蒐集資料、分析題目品質優劣（古典測驗理論法—試題難度、鑑別度、誘答力或試題反應理論—鑑別參數、難度參數與誘答力參數），做為淘汰不良試題依據。建立信度、效度、常模，供日後施測參考。

(八)編製指導說明手冊：

將優良試題編製成測驗題本，同時編製指導手冊，如實施過程指導說明、答案紙作答時間限制、計分方式均需詳細說明。

三、請解釋下列名詞：

(一)評分規程(Rubrics)

(二)t 檢定(t test)

(三)統計考驗的接受區(Acceptable region)

(四)盒鬚圖(box and whisker plot)

【擬答】：

(一)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）計分準則（Scoring Rubric）的具體程序或步驟：

1. 蒐集學生相關作品。
2. 將學生作品分為不同程度的組別（如：高、中、低三組），並分析各組的表現特質。
3. 對表現特質進行層面的分類，並寫下分類的原因。此階段的目的是在建立教師與學生共通的語言，學生透過參與，能對規準及成功的標準有更深一層的瞭解。
4. 以清楚明確的語言界定層面。
5. 列出每一層面不同等級行為表現的描述語，根據各程度的描述語，編製檢核表或等級評量表。
6. 選取各個層面不同等級具有代表性的樣本作品。
7. 修訂與改進，試用訂定的標準，若無法有效描述不同程度的表現，可在過程中不斷修訂與改進。

(二)t 檢定(t test)

1. 意義：

二個獨立分配，一為常態分配 $Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$ ，另一為 $\chi^2_{n-1} = \sum_{i=1}^n \frac{X_i - \bar{X}}{\sigma} = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2}$ ，自由度為 $n - 1$ 的卡方分配，若 Z 與 χ^2 獨立，則令經過無數次隨機抽樣，形成無數個 t 值，這些 t 值成分配，即 t 分配。適用於連續變項資料 t 分配用於假設考驗即為 t 考驗。

$$\text{公式} : t = \frac{Z}{\sqrt{\frac{\chi^2}{df}}} = \frac{\frac{\bar{X} - \mu}{\sigma / \sqrt{n}}}{\sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2 / \sigma^2}{df}}} = \frac{\bar{X} - \mu}{\hat{S} / \sqrt{n}}$$

2. 來源：

西元 1908 年 W. S. Gossett 以筆名 Student 發表，後來被 R. A. Fisher 蒐集在他的統計表中，將這小樣本抽樣分配稱為 Student's t Distribution，簡稱為 t 分配；常態機率曲線是 Karl F. Gauss 所提出的，故又稱高斯分布。

3. t 分配特性：

- (1) 當 $df > 2$ 時，t 分配平均數 $E(t) = 0$ ，變異數 $V(t) = \frac{df}{df-2} > 1$ (當自由度 ≤ 2 時，t 分配沒有變異數)。
- (2) 每一個自由度必對應一條 t 分配曲線。
- (3) 當 $df = 1$ (即只有二個群體時)，此時 $t^2 = F$ 。
- (4) 峰態係數 $\beta_2 = \frac{3df-2}{df-4}$ ， $df > 4$ ； $\beta_2 > 0$ ，故 t 分配為高狹峰分配。

4. t 分配主要用途：

- (1) 母體平均數的推論 (母體標準差 σ 未知，且為小樣本 $n \leq 30$)。
- (2) 二個母體平均數差 ($\mu_1 - \mu_2$) 的推論。
- (3) 相關係數的考驗 (包括積差相關、點二系列相關、等級相關)。

(三) 統計考驗的接受區 (Acceptable region)

假設考驗 (Hypothesis Testing) 事先對母體未知的母數，建立一個合理的假設值，再由樣本資料用來考驗此假設是否成立，稱為假設考驗或顯著性檢定或檢定或考驗。假設考驗有研究假設與統計假設，而統計假設又分虛無假設與對立假設。

1. 統計考驗的接受區為統計考驗結果落在兩個臨界值間，稱為保留區 (Region of retention)，若研究者的假設考驗值落在保留區內，則接受虛無假設，而非表虛無假設為真，只是無充分證據推翻虛無假設，因此不用「接受」虛無假設，而改用「不拒絕」虛無假設。

2. 統計考驗的接受區表示有機率犯型 II 錯誤與 $1-\alpha$ 的機率。

(四) 統計學者 Tukey 提出，又稱盒形圖 (Box-whisker Plot) 或稱五個量數彙總圖，其繪製方法非常簡單，該圖形中箱子包含次數分配最中間 50% 次數，箱子二端分別為第一與第三四分位數 (即 Q_1 與 Q_3)，箱子二側則以最大值與最小值表示。盒形圖 (Box-and-whisker Plot) 處理與分析資料時有何功能：

1. 可以用來顯示資料中四分位差、中位數、全距。
2. 分配範圍，如長方盒長度越長與外延垂直線越長，代表資料越分散。
3. 檢查資料中是否出現極端值、偏離值，當極端值與偏離值點數越多，代表偏離情形越嚴重。
4. 瞭解分配型態，當中位數上下兩側延伸線越不相等，表示偏態越明顯。

四、當使用共變數分析 (ANCOVA) 來考驗教學實驗效果時，除了要符合變異數分析 (ANOVA) 的基本假定外，還要符合那些基本假定？如果使用兩個以上的共變項，應該依據那些原則選擇共變項？

【擬答】：

(一) 意義：

共變項分析 (Analysis of Covariance，簡稱 ANCOVA)，即使用一種統計控制的方法，利用統計手段把可能影響實驗正確性之無關干擾變項加以排除；利用直線迴歸方法，將上述

公職王歷屆試題 (98 地方政府特考)

無關因素或共變量從變異數中排除，再經調整求出變異數的不偏估計值。其目的在：1. 增加統計考驗力；2. 降低實驗誤差。

(二)基本假設：

1. 共變量不限一個，不必一定要在實驗開始前獲得，但有時為避免實驗處理影響，會在實驗前便先行蒐集。共變量為等距變項。
2. 符合母數統計學假設：
 - (1) 變異數同質性 (Homogeneity of Variance)：當進行母數統計學考驗，如：t 考驗與變異數分析，必須遵守變異數同質性，即 $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$ ；因為變異數不同質性，表示兩個樣本平均數差異另有誤差來源，致使變異數呈現不同質情況。
 - (2) 常態性 (Normality)：指樣本來自母群體在依變項分配呈常態分配。
 - (3) 獨立性 (Independence)：指每個樣本均需是獨立或各變異來源的總離均差平方和 (可以分成幾個部分加總)，且每個部分均互相獨立或不相重疊。
3. 共變項與依變項呈直線相關。
4. 每一實驗組內迴歸線假設平行，亦即組內迴歸係數 ($\beta_1 = \beta_2 = \dots = \beta_j = \beta_w$) 相同。(余民寧，民 92)
5. 隨機分派且實驗處理必須為固定效果。

(三)如果使用兩個以上的共變項，應該下列原則選擇共變項

1. 變量與共變量：假定除了實驗變項外，尚有其他干擾變項也會影響依變項，將使得實驗變項與依變項因果關係無法確定；此一其他變項即干擾變項，在共變數分析中即共變量 (Covariate)，而依變項叫做變量 (Variate)。共變量與依變項相關盡量高，共變項與自變項或實驗處理之相關要盡量低，否則易產生多元共線性現象，造成解釋上扭曲現象。
2. 共變項分析是變異數分析與直線迴歸合併使用。先利用直線迴歸將共變量的影響排除後，再利用變異數分析去考驗各組平均數之間差異是否存在。
3. 共變項不限一個，不必一定要在實驗開始前獲得，但有時為避免實驗處理影響，會在實驗前便先行蒐集。共變項為等距變項，共變項與依變項呈直線相關。
4. 共變項大都為實驗研究過程影響研究結果的無關干擾變項，例如教師能力、性別、智力、前測、教學進度、教材內容。