

111 年公務人員高等考試三級考試試題

等 別：三等考試

類 科：教育行政

科 目：教育測驗與統計

思澤老師

甲、申論題部分：

一、請說明以雷達圖 (radar chart) 來繪製統計圖表的適用時機，(12 分) 並說明解讀雷達圖資訊時要注意那些陷阱？(13 分)

1. 《考題難易》：★★★★

2. 《破題關鍵》：雷達圖並不常出現在普考教測統的命題範圍，今年特意出現，也是試圖更改過往重計算的作法

3. 《使用學說》：描述統計～圖表製作

【擬答】：

(一)雷達圖

雷達圖可比較多個定量變數，適用於視覺化哪些變數具有相似值或者變數之間是否存在離群值。雷達圖包含一系列分支，每個分支代表一個變數。雷達圖也適用於判定資料集中哪些變數評分較高或較低。換句話說，將三或更多定量變數的多元資料對應到一根軸上，看起來像一張蜘蛛網，其中心軸至少輻射出三個輻條，稱為半徑，資料的值就對應到這些輻條上。它旨在針對產品、服務或任何其他感興趣的項目，一目瞭然地顯示其間的相似性、差異和離群值。

(二)解讀雷達圖的陷阱

1. 難以判斷半徑長度

半徑上的距離很難用肉眼判斷。雖然人們或許知道並理解一個變數比另一個更長或更短，但實際上這可能很難量化。在這種情況下，在圖表上使用同心圓將有助於使半徑長度更容易判斷。如果對長度的理解很重要，還可以考慮在這種情況下使用折線圖。

2. 雷達圖可能會扭曲資料

一旦所有測量值都記錄在圖表上，整個區域可能被填滿，而這可能會扭曲資料，因為可能將陰影面積當做對大小和正面表現的視覺判斷依據。此外，如果一個圖表有五個變數在半徑上測量到 100，則測量結果為 90 的圖表面積會比結果為 82 的圖表大 10%。

3. 雷達圖可以連接彼此沒有關聯的變數

如果圖表上有五個變數，並且它們都標記在半徑上，那麼很容易認為這些並列的測量值之間存在某種關係。然而，以布朗尼為例，巧克力風味的質地和等級之間可能沒有關係，因為它們是兩個獨立的變數。

4. 雷達圖可能導致遮擋和混亂

如果變數太多或資料系列太多，圖表可能會變得混亂。如果在一張圖表上繪製多個資料系列，則此問題會更複雜，並且資料點可能會被遮擋。雖然保持顏色陰影透明有助於避免遮擋，但如果有多個系列都達到同一個變數測量值，則會更難提取清晰的資料全景。同樣地，形狀也會產生心理作用。人類可以很容易識別及辨認正方形、圓形和三角形等形狀的資料，而雷達圖形狀的隨機性使它不如可量化的已知形狀那麼好用。

參考來源：

林清山著。心理與教育統計。東華出版。<https://www.ibm.com/docs/zh-tw/spss-modeler/SaaS?topic=types-radar-charts>

二、請分別說明虛假相關 (spurious correlation)、淨相關 (partial correlation) 及部分相關 (part correlation) 的意義, (12 分) 並各舉一例說明其應用時機。(13 分)

1. 《考題難易》: ★★★

2. 《破題關鍵》: 相關一直是教育測驗與統計的重要概念, 所以多個變項間關係有不同類型, 考生們在本題應該可以順利取得不錯的分數

3. 《使用學說》: 相關迴歸

【擬答】:

(一) 虛假相關

所謂假相關 (spurious correlation), 是指兩變項之間的相關是表面的, 不代表兩變項間具有真的因果性 (causality) (因生果或互為因果)。

像是假相關可能在兩種情形下產生:

1. 兩變項間本無關係, 只因第三變項存在而造成兩變項的假相關。例如小學生的平均身高雖與其平均書寫辭彙間有相關關係。事實上兩者間彼此無關, 造成其相關的是第三個變項 (年齡或年級)。
2. 兩變項之間的相關不是直接的, 而是由於兩變項間所存在的第三個中介變項使然。例如某公司調查發現已婚員工平均子女數與其定期內平均請假缺席時數具有正相關關係。此種情形並非是假相關, 而是間接相關。原因是構成員工缺席的原因與他們子女多需要照顧有關。在心理學上通常採用淨相關 (partial correlation) 統計方法剔除假相關因素, 從而了解變項之間的相關關係。

(二) 淨相關

兩個變項在它們與一個或多個變項的共同解釋力被排除之外, 所剩下的純相關程度以 $r_{12.3}$ 表示。

例如: 國語成績和數學成績的相關極高, 但是國語、數學都會受到智力的影響, 因此將智力自果與、數學成績中分別排除之後所剩下的純相關即是「淨相關」。

$$r_{12.3} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{1-r_{13}^2}\sqrt{1-r_{23}^2}}$$

(三) 部分相關

指「第一變項」和「排除第三變項共同解釋力後的第二變項」之間的相關程度。

以 $r_{1(2.3)}$ 表示。

例如: 國語成績和排除致力因素後的數學成績間的相關。

$$r_{1(2.3)} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{1-r_{23}^2}}$$

參考資料:

林清山著。心理與教育統計。東華出版。

志光
保成
學儒

112年 虛實整合

重聽OK
旁聽OK

多元學習新型態

突破傳統上課形式 **5大方式彈性又便利**
| 面授學習 | 直播學習 | 在家學習 | 視訊學習 | Wifi學習 |

◆學習◆
零時差

同類科各班別
皆可同步直播上課

◆服務◆
零死角

服務緊貼需求
隨時掌握學習狀況

線上
課業諮詢

老師
申論批閱

雙師資
雙循環

多元
補課方式

上榜生
經驗親授

時事
專題講座

歷屆試題
練習

班導師
制度

各班服務略有不同，詳情請洽全國志光、保成、學儒門市

三、受試者回應歷程 (examinee response processes) 資料的探究為當前評量的關注焦點之一，試說明其意義、資料來源類別及其對測驗效度的貢獻。(25 分)

1. 《考題難易》：★★★★★
2. 《破題關鍵》：本題目所涉及效度觀念常出現在教育測驗與統計的考古題，但是這一次卻是探討新近效度證據建立觀點，難度頗高
3. 《使用學說》：效度

【擬答】：

(一)受試者回應歷程 (examinee response processes) 的意義和資料類別來源

受測者回應歷程係指受測者在閱讀、解釋和形成解法的受測期間，所表現的思考歷程、策略、取向和行為等。蒐集受測者的回應歷程資料可以透過 Think-aloud protocols 和認知面談等方式，由受測者自己口頭說明自己的思考歷程。在以電腦為基礎的施測，也可以蒐集眼動類型、凝視持續和凝視停留時間等，透過操弄測驗作業特性，前述反應表現也會跟著變化，進而推論受測者面對測驗 (作業) 時內在心智歷程。前述變化可以透過口語報告和視覺化圖形，瞭解受測者的解決問題和錯誤類型、凝視電腦螢幕位置、受測者處理不同題目凝視時間和題目呈現後受測者掌握線索，並引出正確知識來解決題目。

資料類型涉及受測者如何閱讀和解釋題目、受測者如何解決問題的步驟和引用策略、受測者的知識和能力如何對應題目要求、受測者是真實理解題目要求去回答而非透過猜測方式和測驗有效的提供線索和資源搭配等皆是可以透過受測者回應歷程蒐集到的資料類別。

(二)對測驗效度的貢獻

1. Messick (1990)並不認同傳統效度的定義和證據建立方式，認為效度證據建立應該考量下列五種形式，分別是：

(1)基於測驗內容的證據

與內容效度的觀念雷同，重點在針對測驗內容進行邏輯性的評估與解釋。

(2)基於反應歷程的證據

針對受測者在受測反應或是受測後透過面試的方式來瞭解受測者採用心智歷程如何影響答題表現。

(3)基於內在結構的證據

如同建構效度的觀念，基於理論架構採用在測驗發展過程，可以採用的適當統計分析進行解釋。

(4)基於與其他變項的相關

觀念雷同效標關連效度與建構效度，測量相同建構的測驗間其預期的相關結果與真實受測結果是否一致。

(5)基於受測結果的推論正確性為證據

受測結果的解釋其正確推論程度。

2. 1967 年 Neisser 撰寫「認知心理學」一書後，認知心理學家開始以訊息處理

(information processing) 的概念研究人類解決問題的歷程，Embretson (1983)將項目反應理論和認知心理學二個領域專業串連在一起，相對於傳統心理計量學研究所重視解題的最後結果，這種以基本認知歷程來分析測驗表現將有助於了解測驗所測量的究竟是什麼。它不僅有助於心理學家重新建構一個周延和精確的智力理論，也更能夠具體指出受測者的優點和弱點，進而增加測驗在診斷上的用途。

參考來源：

郭生玉著。教育測驗與評量。精華出版。

Ercikan & Pellegrino (2017) Validation of Score Meaning for the Next Generation of Assessments

四、請說明 Q 技術的適用時機與實施歷程。(25 分)

1. 《考題難易》：★★★

2. 《破題關鍵》：Q 技術甚少出現在教育測驗與統計的考古題中，在心理與教育研究上，針對自我概念等相關研究議題，可以提供適當測量方式，進行研究

《使用學說》：實務評量

【擬答】：

(一) Q 技術與研究例子

此種技術為 Stephenson 於 1953 年所創，主要為評量自我觀念或其他情感變項。研究者首先根據研究的目的與需要，設計一些描述行為特質的詞句如「快樂的」、「努力的」等。再將這些詞句分別各寫在一張卡片上，卡片的數目最好是介於六十張至九十張之間。像是人本論取向的 Rogers 為了測量自我概念，修正 Stephenson(1953)發展的測驗。測量方式可以分成二個情境，在「真實我」與「理想我」前提下，受測者基於施測者所提供的卡片中，描述與性格特徵有關的語句，進行評估；評估方式則是將卡片區分從「最像我」到「最不像我」間分成程度差異的數堆，用以表示卡片描述符合程度。

這樣的測量方式將可以量化「真實我」與「理想我」間不一致的情況，例如：Block 與 Rubins(1993)研究自尊隨年紀變化的情形，研究中定義自尊為知覺自我與理想我的相似程度（高相似程度為高自尊），結果發現男生跟女生在 14 歲左右自尊程度差異不大，但是隨著年紀增加，性別差異逐漸擴大，男生自尊程度逐漸上升；女生自尊程度逐漸下降。另一方面研究者也使用面談資料探討高自尊程度的歸因性別差異，發現女性重「親密關係」；男性重「控制性」，並有情緒上疏離（性別刻板印象）。

(二)實施歷程

Q 分類的實施，是請受試者或了解受試者的其他人，依據這些卡片上詞句符合自己或受試者的程度，做若干等級的分類，其等級從非常符合（同意）到非常不符合（同意），通常以九至十一個等級為宜。受試者依據常態分配，在每一等級放入適當數目的卡片。例如在

九個等級時，可根據常態分配的標準九(stanine)，依次為百分之四，百分之七，百分之十二，百分之十七，百分之二十，百分之十七，百分之十二，百分之七，百分之四分別投入卡片。

Q 分類常被用來研究個人的心理問題，受試者可以 Q 分類分析自己在不同情境，例如家庭、學校或其他社會情境中的狀況，也可以分析對所認識的人如父母、教師的看法。在自我觀念研究方面，Q 分類可用來分析受試者自己認為的真實我(real self)，他認為別人對他的看法（社會我，social self）及他的理想我(ideal self)。Q 分類可用來研究在心理治療過程中，個人自我觀念改進的情形，在治療之後，個人自我觀念應會變得比較積極或接近個人的理想我。

參考來源：

郭生玉著。教育測驗與評量。精華出版。

<https://terms.naer.edu.tw/detail/1315797/>

公職



想上榜嗎?其實你只需要做到這件事

加入志光保成學儒

學費省很大 全年課程不間斷，一次繳清學費輔導至考取。 (每年僅需繳交換證教材費)	課程最完整 完整課程循環，基礎班→正規班→專題課程→總複習等，全部擁有。	上榜賺獎金 報名考取班第一年考取同職等考試，頒發獎學金。	學習最便利 輔導期間可依自己時間選擇面授或視訊學習，提高學習效率。
師資最多元 重點科目安排多元師資，雙循環教學，可旁聽加強弱科，強化上榜實力。	加選最超值 輔導期間加選其它科目增加考試機會，另享專案優惠。	榜單最實在 年年榜單見證，錄取人數最多，錄取率最高，奪榜實例全國第一。	公約有保障 考取班簽訂公約，保障您的權利與義務至考取為止。