

103 年公務人員高等考試三級考試試題

類 科:教育行政

科 目:教育測驗與統計

一、李老師編制了一份數學推理測驗 X，下表為該份測驗的一些相關資料，請根據數據回答問題。

平均數	20
標準差	6
題數	20
KR-20	0.64

(一)以 KR-20 為例：

1. 從古典測驗理論真分數模式的角度，解釋該係數所代表之意義及其基本假定。
2. 計算觀察分數與真分數的相關
3. 說明此信度係數之高低受那些因素的影響。

(二)小英與小明在該份測驗上分別得到 20 分及 25 分，請據此建立 68%信心區間，並解釋及比較兩人的表現。

(三)若 X 測驗之兩週再測信度為 0.70，以 68%信心區間預測小英兩週後的再測成績為何？

(四)由於 X 測驗之信度不理想，李老師決定將測驗題數增加至 40 題，請根據 KR-20 之值估計 40 題測驗的信度。李老師再編制完成 40 題的新測驗後，請該校 200 位資優生進行預試，得到 KR-20 係數 0.68，說明係數未如預期提高的可能原因為何？

【擬答】：

(一)1. 測驗分數中由特質的真實差異與測量誤差所造成誤差比率多寡。如誤差成份高，則信度低；誤差成份低，則信度高。庫李信度 KR-20 基本假設：

- (1) 來自測驗內容抽樣的誤差。適用於記分非對即錯的標準化測驗。
- (2) 題目不會受到填答速度影響。
- (3) 題目間具有同質性，亦即測量一個共同因素。

$$2. r_{xt} = \sqrt{r_{xx}} = \sqrt{.64} = .8。$$

3. 庫李信度 KR-20 分析題目之一致性，此題目間一致性量數的高低受到兩種誤差變異量的影響，一為內容取樣誤差，一為所取樣之行為領域的異質性誤差。

$$(二) SE_{meas} = S_t \sqrt{1-r_{xx}} = 6\sqrt{1-.64} = 3.6$$

$$\text{小英 } X - Z \cdot SE_{meas} \leq t \leq X + Z \times SE_{meas}$$

$$20 - 1 \times 3.6 \leq t \leq 20 + 1 \times 3.6$$

$$16.4 \leq t \leq 23.6$$

$$\text{小明 } X - Z \cdot SE_{meas} \leq t \leq X + Z \times SE_{meas}$$

$$25 - 1 \times 3.6 \leq t \leq 25 + 1 \times 3.6$$

$$21.4 \leq t \leq 28.6$$

使用百分位數帶 (Percentile Band) 將原始分數上下各加減一個標準誤，當兩個測驗的帶狀範圍重疊，則表示測驗分數沒有顯著差異，小英與小明分數有部分重疊，表示兩人在數學推理測驗應不相上下。

$$(三) SE_{meas} = S_t \sqrt{1-r_{xx}} = 6\sqrt{1-.7} = 3.286$$

預測小英

$$X - Z \cdot SE_{meas} \leq t \leq X + Z \times SE_{meas}$$

$$20 - 1 \times 3.286 \leq t \leq 20 + 1 \times 3.286$$

$$16.714 \leq t \leq 23.286$$

(四)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。

公職王歷屆試題 (103 高普考)

二、(一)什麼時候適合使用配合題？其主要的限制為何？

(二)配合題的命題應注意那些事項？

【擬答】：

(一)1. 適用時機：配合題由二個部分組成，一是問題項目（前述項目），另一個是反應項目，測量簡單連結為基礎的事實性訊息，若有教材內容相關聯性很高，配合題有利於在短時間內測出大量相關聯的事實資料。配合題是一種改良式的選擇題，適用低年級學生或語文能力較低的學生施測。

2. 主要的限制：

- (1) 編製同性質題目較困難。
- (2) 僅能測量機械記憶的事實知識。
- (3) 配合題比其他題型，使用更多無關的線索。

(二)命題應注意事項：

1. 同一配合題中，各個項目性質應儘量相同。
2. 問題項目與反應項目的數量不宜相等，且不宜限制每個反應項目被選次數，否則容易造成學生猜測答案；問題項目最好少於反應項目且最好印在同一頁上。
3. 作答方式與計分方式在測驗指導語中清楚說明。
4. 反應項目宜按系統排列（如邏輯、時間、數字、字母）。
5. 問題項目與反應項目儘量力求簡短。
6. 問題項目避免使用不完全的敘述句。
7. 配對項目以不超過十項為原則。

三、請以圖說明「莖葉圖(stem-and-leaf plot)」的主要成分為何？並說明莖葉圖的特色

【擬答】：

(一)普林斯頓大學 Tukey 在 1977 年發展用以描述觀察簡便方法，最適合兩位數資料呈現，莖葉圖編製過程以 10（或 10 的倍數）為組距，係將每一觀察值切割成莖與葉兩部分，中間以垂直線區隔，數字中十位元數部分當作「莖」，而數字中個位數定義為「葉」，若將莖葉圖橫著看與直方圖功能類似。可討論數值中是否出現極端值，但如原始數據繁多，則難以計算。

莖	葉
3	2
4	3367
5	12344679
6	578
7	23

圖 忠孝國小五年級數學成績莖葉圖

(二)一般次數分配圖將原始資料轉換成圖形過程中，資料部分訊息已經流失，即無法瞭解原始資料內容，為了克服此種缺失，讓圖形兼具數字和圖形的優點，保留原始資料。

四、(一)何謂抽樣誤差？抽樣誤差會受到那些因素的影響？

(二)什麼是統計檢驗力(power)？什麼因素會影響統計檢驗力？

【擬答】：

(一)1. 抽樣誤差：由母體平均數，變異數 σ^2 中，每次抽出若干個體為一組樣本，再由各組樣本求算出一個樣本平均數 (\bar{X})，重複此一步驟，反覆進行無數次抽樣，進行每一次的抽樣會得到一次抽樣誤差，無數次的抽樣會得到無數次的抽樣誤差，這些抽樣誤差形成的標準差即為抽樣分配標準差。為與母體之標準差區分，一般又稱標準誤 (Standard Error)；且樣本平均數的分散情形 ($\sigma_{\bar{x}}$) 顯然比母群的分散情形 (σ) 小很多。

$$\text{公式} : \sigma_{\bar{x}} = V(\bar{X}) = \frac{\sigma_x^2}{n}$$

2. 抽樣誤差會受到那些因素的影響：包括母群的分散情形 (σ) 越大則抽樣誤差越大，兩

公職王歷屆試題 (103 高普考)

者成正比；樣本人數 (n) 越大則抽樣誤差越小，二者成反比。

(二) 1. 統計檢驗力(power)：在固定 α 下，經統計考驗而正確拒絕 H_0 之機率值為 $1 - \beta$ ，亦即當我們拒絕 H_0 ，而 H_0 的確是錯誤的。 $1 - \beta$ 代表其檢定能力的大小，只在研究假設為真時才能計算，如果研究人員進行一項研究來探討某個「現象」或是某個實驗效果，這個現象或效果真的存在的機率就是統計考驗力。具有積極意義的正確決定。符號 $\text{Power} = 1 - \beta = P(\text{接受 } H_1 | H_1 \text{ 為真})$

2. 影響統計考驗力因素：

- (1) 樣本大小 n：統計考驗力會隨樣本大小增加而增加，因為樣本大小 n 增加時，會使得抽樣誤差 σ_x 變小，同時使 α 、 β 變小，而 $1 - \beta$ 變大。
- (2) 顯著水準 α ：顯著水準與統計考驗力之間的關係是正向的，統計考驗力會隨顯著水準增大而增加（因為 α 增大時，拒絕區將增大，比較容易拒絕 H_0 ，犯型 II 錯誤會減少，相對地 $\text{Power} = 1 - \beta$ 會增加）。
- (3) 假如其他條件不變，統計考驗力會隨著真正參數值 H_1 離 H_0 假設參數值愈遠而增加。
- (4) 單側考驗與雙側考驗：單側考驗比雙側考驗的臨界區還大，所以單側考驗比較容易拒絕 H_0 。由於單側考驗拒絕 H_0 機率較高，單側考驗犯型 II 錯誤機率較低，同時統計考驗力相對地提高。
- (5) 效果值 (Effect Size)：效果值只被檢定的效果項，實驗研究中實驗處理效果大小，效果值愈大，研究偵測到所要探討的現象的可能性也就愈大，表測驗得分的平均數差距愈大，統計考驗力相對地提高。

$$\text{效果值} = \frac{c - \bar{c}}{\sigma}$$

(6) 測驗工具之信、效度，信效度愈高統計考驗力愈高。

五、陳老師根據一個具有代表性的學生樣本測得智力測驗分數(X)的平均數為 105，標準差為 15，而同時也知道該樣本的學業成就(Y)之平均數為 80，標準差為 6。智力測驗分數與學業成就之間的相關為 0.8。根據這些資料回答下列五個小題。

- (一) 學生的學業成就分數的變異中有多少%是無法以智力測驗分數加以解釋的？
- (二) 若將學業成就的分數加以標準化之後，再以 $Z=50+10z$ 的方式來加以直線轉換成 Z 分數，那麼轉換後的學業成就與智力測驗分數的相關係數變為多少？
- (三) 如果陳老師想利用智力測驗的分數來預測學業成就的分數，則陳老師所得到的非標準化直線迴歸的方程式為何？
- (四) 根據(三)所計算得到的迴歸方程式，則估計標準誤將等於多少？
- (五) 如果有一位學生的智力測驗分數為 90，則該生實際的學業成就分數有 95% 的機會將落在那一個區間之內？

【擬答】：

(一) 疏離係數的平方 = $1 - r^2 = 1 - .8^2 = .36$ ，即表「學生的學業成就分數的變異中，無法被智力測驗分數解釋到變異量部分為 36%」。

(二) 相關係數 = .8。二變項經直線轉換或四則運算後，其相關係數是不會改變的。

(三) 非標準化直線迴歸的方程式 $\hat{Y} - \bar{Y} = r \cdot \frac{S_Y}{S_X} (X - \bar{X})$

$$\therefore \frac{\hat{Y} - 80}{6} = .8 \times \frac{X - 105}{15} \rightarrow \hat{Y} = 46.4 + .32X$$

(四) 預測標準誤 $S_{Y \cdot X} = S_Y \sqrt{1 - r_{XY}^2} = 6\sqrt{1 - .8^2} = 4.8$

(五) $X=90$ 代入 $\hat{Y} = 46.4 + .32X = 46.4 + .32 \times 90 = 75.2$

$$\hat{Y} \pm Z_{.975} \cdot S_{Y \cdot X} \leq Y \leq \hat{Y} + Z_{.975} \cdot S_{Y \cdot X}$$

$$75.2 \pm 1.96 \times 4.8 \leq Y \leq 75.2 + 1.96 \times 4.8 \quad (95\% \text{ 信賴區間})$$

$$65.792 \leq Y \leq 84.608$$

推估實際學業成績落在 65.792 至 84.608 之間可以信賴程度為 95%。但仍有犯錯機率為 5%。