

## 113 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

梁昕/吳迪老師

一、研究者想從學生的閱讀自信預測其閱讀表現，蒐集了 402 名學生的資料，得到下表的描述性資料。閱讀自信與閱讀表現的共變量為 53.58。

	最小值	最大值	平均數	標準差	變異數
閱讀自信	3.03	13.53	9.45	1.89	3.57
閱讀表現	297	753	558	70	7900

(一)檢驗閱讀自信與閱讀表現是否有顯著的相關，寫出統計檢驗步驟並根據結果下結論。(15 分)

(二)以閱讀自信 (X) 來預測閱讀表現 (Y)，其迴歸方程式為何？(10 分)

《考題難易》★★

《破題關鍵》考簡單迴歸與相關分析，此為近幾年常考的題型，只要熟記公式應可拿分。

《命中特區》吳迪著“教育統計”3-3 及 6-10

【擬答】

(一)  $\begin{cases} H_0: \text{閱讀自信與閱讀表現無顯著相關} \\ H_1: \text{閱讀自信與閱讀表現有顯著相關} \end{cases}$

$$r_{xy} = \frac{S_{XY}}{S_X S_Y} = \frac{53.58}{\sqrt{3.57} \sqrt{4900}} = 0.4051$$

$$\text{設 } \alpha = 0.05$$

$$\text{拒絕域 } C = \{t \mid t > 1.96 \text{ 或 } t < -1.96\}$$

$$\text{其中 } t_{0.975}(400) \doteq 1.96$$

考驗統計量

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.4051\sqrt{402-2}}{\sqrt{1-0.4051^2}} = 8.86$$

結論： $t=8.86 > 1.96$  落入拒絕域  $\Rightarrow \text{Re}H_0$

有證據顯示閱讀自信與閱讀表現有顯著相關。但仍有 5% 之機率犯型 I 錯誤

(二)  $\hat{Y} = a + bx$

$$1. b = r \frac{S_Y}{S_X} = 0.4051 \times \frac{\sqrt{4900}}{\sqrt{3.57}} = 15.01$$

$$2. a = \bar{Y} - b\bar{X} = 558 - 15.01 \times 9.45 = 416.16$$

$$\Rightarrow \hat{Y} = 416.16 + 15.01X$$

二、研究者探討性別 (男、女) 與教學介入 (新教學法、傳統教學法) 對 39 名學生學習表現的影響。以二因子變異數分析資料，結果得到，總平方和 (SSy) 為 534.5，組別、性別、組別 X 性別三者的平方和加總為 243.5，組別與組別 X 性別兩者之平方和加總為 183.5。組別 X 性別的平方和為 10.5。請寫出二因子變異數分析表 (含變異來源、平方和、自由度、平均平方和與 F 值)。二因子變異數分析交互作用顯著的意義為何？結果出現顯著的交互作用時，該如何進一步分析與解釋資料？(25 分)

《考題難易》★★

《破題關鍵》考二因子變異數分析表及交互作用的意義。

《命中特區》吳迪著“教育統計”8-5

【擬答】

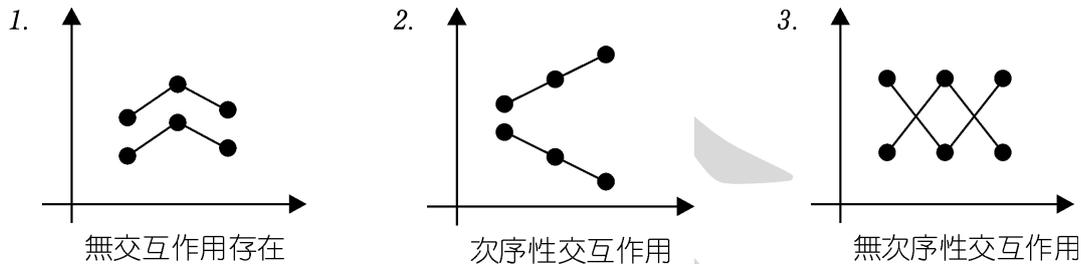
(一) ANOVA 表

來源	SS	df	MS	F 值
組別	173	1	173	$F_1=20.82$
性別	60	1	60	$F_2=7.22$
交互作用	10.5	1	10.5	$F_3=1.26$

誤差	291	35	8.31	
總變異	534.5	38		

(二)交互作用顯著的意義為不同組別會對於不同性別有不同的影響。例如新教學法可能比較適合男生，而傳統教學法可能比較適合女生。所以當出現顯著的交互作用時須做單純主要效果考驗。以了解其交互作用的種類。交互作用的種類可以利用每一個細格的平均數畫出圖形表示。其種類有：

1. 兩條線平行或接近平行代表二個變項間無交互作用存在。
2. 兩條線不平行且無交點稱為次序性交互作用。
3. 兩條線有交點稱為無次序性交互作用。



(參考:林清山著”心理與教育統計學”，2010年元月，東華書局)

三、下表是 205 位學生在一份由 10 道選擇題構成試卷的試題分析情形。表格內呈現的為學生在四個選項的選擇人數百分比，數值右上方有星號的是正確答案；另有試題的區辨力 (Dic) 和男 (M) 女 (F) 生的答對率。每個試題的權重相同。整份試卷的難度為何？那道题目的變異數最大，其數值為何？那道題目區辨力不理想？試解釋其原因。那道試題會有公平性的疑慮？說明理由。評述此份試卷之誘答選項設計。(25 分)

題目	A	B	C	D	Dic	F	M	題目	A	B	C	D	Dic	F	M
1	2.9	16.6	69.3*	11.2	.49	68.8	69.6	6	3.0	23.7	62.6*	9.6	.47	66.3	59.6
2	15.6	54.6*	4.4	25.4	.33	54.8	54.5	7	96.0*	2.0	0.5	1.0	.26	97.8	94.5
3	4.9	19.0	9.3	66.3*	.46	64.5	67.9	8	0.0	8.6	12.1	78.3*	.59	79.8	77.1
4	75.1*	4.9	6.8	13.2	.55	76.3	74.1	9	0.5	64.6*	23.7	10.1	.46	65.2	64.2
5	1.0	92.2*	2.9	3.4	.39	90.3	93.8	10	2.6	5.1	6.6	85.2*	.52	92	79.6

《考題難易》★★★

《破題關鍵》本題為試題分析的解釋與判斷，作答時必須搭配數據結果一併說明；而第 4 小題，測驗公平性的判斷，可從相同的題目，但男女答對的比例確有大幅差異的題目切入說明。

《使用學說》試題分析、測驗公平性

《命中特區》「113 教育測驗正規課講義」-p143-148；p318-320。梁昕編著

【擬答】

(一)本測驗的難度評估

所謂難度 p 值即為答對人數所佔的比例，p 值越大，表示題目越容易，反之，若 p 值越小，則表示題目越困難。本測驗各題的難度 p 值介於 0.546 至 0.960 之間，顯示題目難度大多介於中等至容易程度，且平均難度為  $(0.693+0.546+\dots+0.646+0.852)/10=0.744$ ，也顯示整體而言題目偏容易。

(二)本測驗中第二題的變異數應較大

變異數表示答題結果的分散程度，從作答表現情形可以看出，第二題的作答結果較為分散，並未集中於某一特定選項，因此該題的變異數應該較大；而該題的變異數可透過衡量每個選項對於變異數的貢獻加以計算，其公式為： $\sum_{i=1}^k \frac{p_i \times (1-p_i)}{n}$ ，其中 i 為第一個選項到第四個選項， $p_i$  表示各選項的選取率。將各題分別計算後，第二題的變異數確實為最大，其變異數為：

$$\frac{0.156 \times (1-0.156)}{205} + \frac{0.546 \times (1-0.546)}{205} + \frac{0.044 \times (1-0.044)}{205} + \frac{0.254 \times (1-0.254)}{205} = 0.0298$$

(三)本測驗的鑑別度評估

## 公職王歷屆試題 (113 高考三級)

鑑別度是指試題能區分能力優劣的程度，一般而言，鑑別度大於 0.4 是較為理想的情況，若鑑別度小於 0.4 或是鑑別度為負，都表示該題的鑑別功能較不理想。根據表格結果，第七題的鑑別度僅有 0.26，進一步看該題的作答反映情形，有 96% 的同學均答對了，顯示無論能力高低，均能夠正確回答該題，顯示該題的鑑別功能較不理想；此外，第二題的鑑別度為 0.33，在鑑別度的表現也不盡理想。

(四)根據測驗結果，在第十題中，男性的答對率(92%)大幅超過女性的答對率(79.6%)，表示此題可能有測驗公平性的問題，對於男性是較有利回答的；此外，第六題中，男性的答對率(66.3%)也超過女性答對率(59.6%)，雖差距不如第十題來的大，但也出現了一定程度的差距，也可能有測驗公平性的隱憂。

(五)本測驗的誘答力分析

測驗的誘答力是指在選擇題中，錯誤的選項必須具有一定程度的似真性，吸引一知半解的學生選此錯誤選項，降低因猜測而答對的機率。根據作答結果顯示：

1. 第五題與第七題的三個錯誤選項誘答力均不佳：第五題與第七題的答對率分別為 0.922 與 0.96，表示無論能力高低，多數同學均未選擇錯誤選項，顯示這些選項的誘答功能較弱。
2. 第八題與第九題的 A 選項誘答力不佳：第八題 A 選項的選擇率為 0%，第九題 A 選項的選擇率為 0.5%，表示多數同學均未選擇這兩個錯誤的選項，未能發揮吸引一知半解的同學選擇此錯誤選項的誘答力。

### 志光.學儒.保成

教育行政 近三年 **17 狀元 8 榜眼 7 探花**

<b>狀元</b> 112 高考教育行政 林○娟	<b>狀元</b> 112 普考教育行政 許○沂	<b>狀元</b> 111 高考教育行政 陳○兒	<b>狀元</b> 111 地特四等教育行政 盧○(花東區)	<b>狀元</b> 111 地特三等教育行政 劉○勇(竹苗區)
<b>狀元</b> 111 地特三等教育行政 邱○青(南投區)	<b>狀元</b> 111 地特三等教育行政 吳○婕(花東區)	<b>狀元</b> 111 地特三等教育行政 陳○俞(新北市)	<b>狀元</b> 111 地特三等教育行政 陳○杰(台中市)	<b>狀元</b> 111 地特四等教育行政 王○威(新北市)
<b>狀元</b> 110 普考教育行政 吳○書	<b>狀元</b> 110 地特三等教育行政 吳○書(台中市)	<b>狀元</b> 110 地特三等教育行政 劉○威(彰投區)	<b>狀元</b> 110 地特四等教育行政 陳○君(新北市)	<b>狀元</b> 110 地特三等教育行政 白○欣(台北市)
<b>狀元</b> 110 地特三等教育行政 洪○婷(台南市)	<b>狀元</b> 110 地特三等教育行政 謝○芸(基宜區)	<b>榜眼</b> 111 地特三等教育行政 李○廷(竹苗區)	<b>榜眼</b> 111 高考教育行政 沈○君	<b>榜眼</b> 111 普考教育行政 甘○葵
<b>榜眼</b> 111 地特三等教育行政 蔡○儒(新北市)	<b>榜眼</b> 111 地特三等教育行政 黃○寧(台北市)	<b>榜眼</b> 110 地特四等教育行政 黃○蓉(花東區)	<b>榜眼</b> 110 地特三等教育行政 蕭○崑(台中市)	<b>榜眼</b> 110 地特三等教育行政 李○慧(台北市)
<b>探花</b> 112 高考教育行政 陳○媛	<b>探花</b> 112 普考教育行政 王○萱	<b>探花</b> 111 地特三等教育行政 簡○雯(台中市)	<b>探花</b> 111 地特三等教育行政 黃○諺(竹苗區)	<b>探花</b> 111 地特四等教育行政 黃○蓉(花東區)
<b>探花</b> 110 高考教育行政 莊○婷	<b>探花</b> 110 普考教育行政 黃○鈺			

**系統式高效學習 112 年 11 月 12 日 楊○翰**

老師上課模式是先在黑板列出重要考點架構表，並配合課本內容進行講解，這樣的方式讓學生得以在每個單元形成系統性的知識，遇到題目馬上就能將腦海中對應的知識運用於解題。

四、評量的信度為何重要？一個測驗分數之變異數為 25，測驗發展者以奇數與偶數題進行折半，兩者的相關為 .87。已知斯皮爾曼-布朗公式 (Spearman-Brown formula) 為  $r_{kk} = \frac{kr_{xx}}{1+(k-1)r_{xx}}$ ，此一測驗的信度為何？其測量標準誤為何？解釋測量標準誤的意義。除了上述估計信度的方法外，還有那些方法可以估計信度，試述其特點？(25 分)

《考題難易》★★

《破題關鍵》本題為信度的基本考題，只要掌握信度的基本概念，折半信度的校正還原、測量標準誤的意涵與計算、各個信度類型的意涵與特性，便可完整取分。

《使用學說》信度、測量標準誤

《命中特區》「113 教育測驗正規課講義」-p70; p73-80; p88。梁昕編著

【擬答】

(一)信度的意涵與重要性

1. 信度是指，相同的受試者，以相同的測量工具或複本，測量結果的「一致性」(consistency)。

如果兩次測驗結果有很高的一致性，就表示測量分數具有高度的「穩定性」(stability)、「可靠性」(dependability)以及「可預測性」(predictability)。

2. 從測量誤差的角度，信度是指測驗分數中，「真實分數」與「測量誤差」的程度，整體而言，測驗分數越一致，受誤差的影響越小；反之，測驗分數起伏變動大，表示受誤差影響大。換言之，由測量誤差的變異量占總變異的比例，即可求得信度，也就是測驗分數反映出真實分數的程度或沒有誤差的程度。

(二)折半信度的校正還原

以斯布公式進行折半信度的校正還原：

$$r_{kk} = \frac{k r_{xx}}{1 + (k-1) r_{xx}} = \frac{2 \times 0.87}{1 + (2-1) \times 0.87} = 0.93$$

(三)測量標準誤的意義與估算

1. 測量標準誤為又稱為分數的標準誤差，是指測量結果的誤差範圍，表示測驗分數可能的變動範圍。當我們進行測量時，由於測量工具並非完美一致，所以難免會有測量誤差發生，亦即觀察分數與真實分數總會有差距，而當測量一個人無限多次時，這些大大小小的測量誤差將構成常態分配，此時這個誤差分配的標準差即為測量標準誤。
2. 測量標準誤的公式為： $SEM = S \sqrt{1 - r_{tt}}$ ，S 為測驗分數的標準差， $r_{tt}$  為測驗的信度，將題目給定的數據帶入： $SEM = 5 \sqrt{1 - 0.93} = 1.32$

(四)其他信度估計方法與特點

1. 再測信度

估計測量信度最直接的方式，將同一測量題目針對同一對象前後進行兩次，根據兩次測驗總分求其相關係數，此相關係數即為再測信度，也可稱為穩定係數；若相關係數高，就表示再測信度高，即表示題目具有穩定性，經得起時間考驗。

2. 複本信度

- (1) 複本信度是指對受測者實施一份正本，一份複本的測驗，求這兩個測驗之間的相關，即為複本信度。複本測驗在內容、形式、題數、難易度等都必須正本相同，且測驗指導、時間限制、作答環境等都需一致。
- (2) 同時連續實施：若正本與複本在同一時間實施，此種複本信度可說明「內容取樣」所造成的誤差，稱為「等值係數」，或稱為「複本立即信度」，可用以檢驗測量歷經指標變動所呈現的等值性，當複本信度越高，表示所測量的題目，越能代表該範圍的測量內容。
- (3) 延後一段時間實施：若先測量正本，延後一段時間後再實施複本測量，除了可說明內容取樣的誤差外，亦可顯示出間隔一段時間後，受測者在不同測驗情境所造成的誤差量(即時間取樣誤差)，此種複本信度又稱為「穩定與等值」係數，或稱為「複本延宕信度」，可以同時檢驗測驗經歷時間變化與指標變動下的穩定性。

3. 庫李信度與 Cronbach's  $\alpha$

- (1) 庫李信度是由庫德(Kuder)和李察遜(Richardson)所發展出來的公式方法，該方法是用在一個測驗僅實施一次，而且不必將測驗分數分為兩半，是分析受測者對每個題目的作答反應，分析試題間的一致性，以確定測驗中的試題是否測量到相同特質的一種估計方法。適用於傳統的是非和選擇題，作答結果只分成「答對」或「答錯」的二分選項。K-R 法的測量誤差，主要來自於「內容取樣誤差」以及「內容異質誤差」。
- (2) 庫李公式的限制在於只能計算二元計分的測驗，若多元計分時(例如：李克特量表)，即無法採用庫李公式計算信度，而 Cronbach  $\alpha$  則能彌補這個缺失。Cronbach  $\alpha$  係依據全體受試者對測驗項目的反應，計算每一個項目以及總分的「變異數」，然後估計「項目變異數」占「總變異數」的比例，以此估計測驗分數的信度。此種方法可視為 K-R20 與 K-R21 發展而來，也可說是信度係數估計的一個通則，因為二元計分只是多元計分方法的一種特例。
- (3) K-R 法與 Cronbach  $\alpha$  可反應出來自於「內容取樣誤差」以及「內容異質誤差」。一般而言，當測驗中的每個試題都測量到相同能力或潛在特質時，即表示試題的測量功能間一致性越高，信度越大；反之，當測驗中的所有試題都測量到兩種以上的能力或潛在特質時，則表示試題的測量功能不一致，信度將越小。

# 志光×學儒×保成 多元學習模式 上課超便利

## 3 大學習式 你的最佳上榜學習式

NO1 互動性 面授學習

名師親臨授課，即時解答疑惑  
同儕交流互動，提升學習氛圍

NO1 超彈性 視訊學習

課程隨選隨看，無限重覆看課  
彈性跟課進度，自由調配時間

NO1 自主度 函授學習

免去舟車勞頓，連結雲端上課  
不分白天夜晚，自主規劃進度

## 2 大輔助利器 滴水不漏快取金榜

NO1 即時性 直播教學

線上按時登入  
現場直播跟課  
立即掌握進度  
無須等待視訊

NO1 人性化 在家學習

使用補課點數  
家裡就是課堂  
免去落課疑慮  
強化複習次數

# 全方位智能學習系統



志光×學儒×保成

虛實整合 引你入勝



### POINT 勝 上課方式最多元

多元學習 新型態 突破傳統上課模式  
學習不受環境影響

面授  
學習

直播  
學習

雲端  
複習

視訊  
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

### POINT 勝 考點掌握最全面

考試關鍵 不漏接 考前、考中及考後，皆享有  
志光、學儒、保成專業服務

考前叮嚀影片

考前重點下載

線上即時解答

考後影音解題

依各區規劃為主，請洽全國門市