

《教育測驗與統計概要》

一、丁老師想要瞭解亞洲四小龍國二生的回家作業完成時間（以小時為單位）與數學成就差異，從TIMSS 2011取得資料，經過分析之後結果如表所示。請回答以下問題。

(一)四個國家國二生完成作業時間，依國名從最高到最低排列為何呢？（4分）

(二)丁老師運用那一種統計方法檢定呢？此方法檢定步驟為何呢？（9分）

(三)在完成作業時間與數學成就，四個國家是否都有明顯差異呢？為什麼？（6分）

(四)請解釋數學成就的事後比較一欄的意義。（6分）

表：亞洲四小龍國二學生回家作業完成時間及數學成就的差異情形

變項	國家	平均數	標準差	F 值	Scheffe法事後比較
完成時間	臺灣(a)	2.19	0.90	$F_{(3,20541)} = 589^{**}$	$a > b^{**}$; $a < c^{**}$
	南韓(b)	1.96	0.84		$a < d^{**}$; $b < c^{**}$
	香港(c)	2.47	0.90		$b < d^{**}$; $c < d^{**}$
	新加坡(d)	2.69	1.14		
數學成就	臺灣(a)	586.14	98.08	$F_{(3,21674)} = 44^{**}$	$a < d^{**}$; $b < d^{**}$
	南韓(b)	587.54	82.93		$c < d^{**}$
	香港(c)	587.13	70.08		
	新加坡(d)	601.31	78.18		

** $p < .01$

試題評析	本題以分析報表精簡列出統計分析結果，所以考驗的是考生對於探討或研究議題的適當統計方法的正確採用。第一小題問的是完成作業時間平均數大小排序的敘述統計，第三小題考的是完成作業時間與數學成就檢定結論的推論統計；第二小題則是變異數分析的進行步驟，最後小題則為檢定達顯著後的事後比較分析。整體而言，題目難度不高，答題時應能具體寫出本題所採用的顯著水準0.01以加強閱卷委員的青睞。
考點命中	《高點·高上教育測驗與統計講義》第二回，傅立葉編撰，第13章，頁49-50，55。

答：

(一)依據題目所附利用TIMSS 2011所取得資料的分析報表可知，亞洲四小龍國二生完成作業時間從最高到最低，依國名排序可以從平均數(時數)排列為：新加坡、香港、台灣、南韓。

(二)丁老師運用檢定的統計方法為變異數分析的F檢定，此方法的檢定步驟如下：

1. 設立虛無與對立假說

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_k$$

$$H_1 : \text{至少一個 } \mu \text{ 不等}$$

2. 指定顯著水準

本題為0.01

3. 決定適當的檢定統計量(teststatistic)

4. 決定拒絕域與F臨界值

5. 進行ANOVA的F檢定結論

6. 當獲致顯著性結論，表示平均數不全等，應進行事後比較以找出導致拒絕虛無假設之確實組別團體的顯著差異。

(三)依題目所附分析報表顯示，四個國家的國二生在完成作業時間與數學成就都存在顯著差異。理由是兩個變異數分析的F值，589與44，分別在0.01的顯著水準下達顯著。

(四)數學成就後方欄位之Scheffe法事後比較顯示，經兩兩比較的檢定結果，台灣(a)、南韓(b)、香港(c)高二生的數學成就分別明顯低於新加坡(d)。足見，數學成就之F檢定所以達顯著的主因是新加坡高二生的數學成就

顯著高於其它三個國家。

二、研究者想要瞭解40位學生在學期初及學期末，對數學課喜歡程度改變情形是否具有顯著差異。學期初及學期末對40名學生調查，得到資料如表。請依據檢定步驟，寫出研究假設、選用統計方法與計算、以 $\alpha = .05$ 檢定、進行裁決與解釋。

(顯著水準 $\alpha = .05$ ，自由度1、2、3、4查表各為 $X^2_1 = 3.84$ 、 $X^2_2 = 5.99$ 、 $X^2_3 = 7.81$ 、 $X^2_4 = 9.49$)。(25分)

表：學期初及學期末的40名學生調查資料

		學期末		
		喜歡	不喜歡	總和
學期初	喜歡	8	16	24
	不喜歡	12	4	16
	總和	20	20	40

試題評析	本題命題以相依樣本設計為基礎，但是對應喜歡與不喜歡數學課的類別資料，所以考生須能夠正確選用卡方統計法之改變顯著性檢定，才能在後續的每個檢定步驟正確計算、裁決與解釋。題目難度相對較低，本班考生能夠輕鬆解答。
考點命中	《高點·高上教育測驗與統計講義》第二回，傅立葉編撰，第12章，頁39。

答：

針對研究者所感興趣之學生在學期初與學期末對於數學課程喜歡情形是否顯著改變的研究問題，以40位學生的樣本資料分析結果之附表資訊，依檢定步驟進行統計分析如下：

(一)研究假設

H_0 ：期初與期末對於數學課喜歡情形無顯著改變

H_1 ：期初與期末對於數學課喜歡情形有顯著改變

(二)適用的統計分析方法為卡方統計法的改變顯著性檢定

(三)計算

卡方統計量 = $(16-12)^2 / (16+12) = 0.571$

在0.05的顯著水準下，自由度為1的卡方臨界值為3.82

(四)檢定結論與裁決

由於0.571的統計量小於3.84的臨界值，檢定結論為：利用該40名學生的樣本資訊無法提供充分證據足以結論學期初與學期末學生對於數學課的喜歡程度改變情形有顯著差異。

三、問卷調查法是社會科學研究的重要方法之一。問卷調查需要設計適切的問卷題目，才能組成一份具有信度與效度的研究工具。請說明在進行社會科學研究時，編製問卷題目應掌握那些原則呢？(25分)

試題評析	本題為心理測驗範圍的考題，這要是藉由問卷調查法的工具編製，要求考生寫出題目編製的原則。由於問卷以選擇題型為主，李克特量尺的態度量表為搭配。所以如果能分開陳述應該更能夠清楚呈現答案的內容與鋪陳。
考點命中	《高點·高上高上教育測驗與統計講義》第三回，傅立葉編撰，第18章，頁40。

答：

進行社會科學研究時，當使用問卷調查法並要求具有一定的信效度，由於通常題目以選擇題為主，搭配部分態度量表，編製選擇題型的問卷題目的原則有幾點，分述如下：

(一)用字淺顯易懂

【版權所有，重製必究！】

題目的文詞應力求清楚明瞭，不要造成作答者對語意的誤解。而且用字也要簡單易懂，盡量節省作答者的作答時間。

- (二)每個問題只涵蓋一個觀念
一個句子只能提及一個觀念，以免作答者混淆。
- (三)避免主觀及情緒化的字眼
問卷的問題應該是採用客觀、中性的字眼，不應用會挑起作答者情緒的文字。
- (四)問題的選項應清楚界定
問卷裡各個問題的選項應界定清楚，特別是涉及年齡、收入等數字級距分類時。選項因為有重疊的情形出現，將會造成作答者無法正確作答，這是編製問卷者要避免的情形。
- (五)不用假設或猜測的語句
假設性的問題，因為作答者有太多的想像空間，以致於所得的結果不易歸納解釋。在實際應用上，價值並不高。
- (六)句子避免過長
通常作答者在填答一份問卷時，都不希望花太多的時間，假如問卷的題目簡單清楚，一目了然，作答者的配合度也會較高；反之，若題目複雜又冗長，作答者有可能會應付了事。
- 另外，針對使用李克特量尺的態度量表，編製原則有：
- (一)基本上以5、7或9等奇數的點數為主，當有強迫選擇的必要時，則採取偶數點數。
- (二)題目組成需要使用Cronbach的 α (Alpha)係數進行同質性探討並進行必要的試題修改或刪除替換。
- (三)可以準備適當的測謊量表。

四、請試述下列名詞之意涵：（每小題5分，共25分）

- (一) 第二類型錯誤 (Type II error)
- (二) 中央極限定理 (Central Limit Theorem)
- (三) 變異數同質性 (Homogeneity of Variance)
- (四) 驗證性因素分析 (confirmatory factor analysis, CFA)
- (五) 中介變項 (Intervening Variable)

試題評析	本題為綜合性的解釋名詞。其中，(一)至(三)題屬於統計學範圍，驗證性因素分析與中介變項的(四)(五)兩題則屬於研究方法範疇，雖然以普考的角度而言稍嫌困難，但是大致不影響考生的心態與分數，假設難度太高，導致幾乎所有考生都難以正常完整作答。
考點命中	《高點·高上高上教育測驗與統計講義》第二回，傅立葉編撰，第9章、第11章。

答：

- (一)第二類型錯誤
假設檢定的目的就是利用統計的方式，推測虛無假設是否成立。若虛無假設事實上不成立，但使用樣本資訊所做的統計檢定結論為缺乏充分證據的不拒絕虛無假設，這種檢定結論的錯誤稱為第二類型錯誤。其犯錯的機率通常以符號 β 表示。
- (二)中央極限定理
中央極限定理說明，在適當的條件下，大量相互獨立隨機變數的平均值，經適當標準化後依分布收斂於常態分配。是樣本平均數之統計量抽樣分配趨近於常態的一個重要基礎，也是在大樣本的條件下，針對母體平均數推論統計適用近似之常態化標準Z分配的理由。
- (三)變異數同質性
變異數分析的前提為假設各組變異數相等，亦稱各組必須具備同質性。主張這項假設成立的原因在於提升變異數分析結論的品質與決策效果。在許多統計分析軟體都有提供此一檢定，最常用的稱為「Levene's檢定」。當此一同質性的條件被違反時，變異數分析常須以適當或對應的無母數統計法取代。
- (四)驗證性因素分析
驗證性因素分析是結構方程模型中的一個環節，用於探討觀察指標與潛在變數之間的關係。相對於探索性因素分析(EFA)多為發展或編製量表時所用，用來協助判斷那些指標(題目)應該被選入或刪除，以及一個構念(潛在變數)項下分別具有那些向度(或可稱為分量表)等；驗證性因素分析(CFA)則多用於一個量表發展完

成後，用以檢定是否特定的指標(題目)都歸到理論預期的各個向度，其主要目的是理論驗證。

(五)中介變項

在從事社會科學，特別是心理學研究的領域，當探討「因」跟「果」關係的時候，有時自變項(因)與依變項(果)的因果關係或許不是這麼令人感興趣，因而常出現「第三者」，其中之一正是中介變項。此變項出現在自變項與依變項之間，通常是在比較複雜的因果關係中才會出現或由研究者經文獻探討而置入。中介變項既是自變項的依變項，也是依變項的自變項，因而可以中介自變項與依變項之間的因果關係。又可以分為完全中介與部分中介效應兩種。

高
點
·
高
上

【版權所有，重製必究！】