《教育測驗與統計》

一、李老師編製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：（10分）

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。（6分）

（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？（4分）

（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。李老師在編製完成40題的新測驗後，以該校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？（5分）

試題評析	庫李信度是常用的信度估算法之一，基於其屬於內部一致性信度的一種，影響因素與意義的解讀都相當容易作答；在真正分數信賴區間的計算上，也是班上練習不少的重點考題，命題有創意，難度不高，先編增題後的信度估算與解釋是拿分重點。
考點命中	《教育測驗與統計》第三回，傅立葉編撰，Ch.15，頁11-13。

答：

(一)1.該信度的基本假定是僅需進行一次施測，而且是以所有測驗題目的反應一致性爲基礎，適用於是非題或非對即錯的記分方式（二元計分）。0.64的庫李信度所代表的意義爲：真正分數的變異佔實得分數變異的比例僅有64%，而實得分數變異中還有36%屬於誤差分數的變異。

2.真分數與觀察分數的相關即是信度的意涵，可以利用以下公式計算：

$$r_{KR20} = \left(\frac{k}{k-1}\right)\left(1-\frac{\sum pq}{s^2}\right) = 0.64$$

3.這種項目間一致性會受到內容抽樣與抽樣之行爲領域的異質性等兩種誤差變異的影響。

(二)依題意，小英與小明在該數學推理測驗的實得分數，可以利用庫李信度係數計算其68%信心區間如下：

小英：

$$20 \pm Z_{0.16}(6)\sqrt{1-0.64}$$
$$= 20 \pm 0.92(6)(0.6)$$
$$= (16.688, 23.312)$$

小明

$$25 \pm Z_{0.16}(6)\sqrt{1-0.64}$$
$$= 25 \pm 0.92(6)(0.6)$$
$$= (21.688, 28.312)$$

計算結果表示，在68%的信心水準下，並考量包括信度在內的測量標準誤，兩人真正數學推理能力的測驗分數將由各自區間加以正確估計；在相同信心水準下，兩人的真正分數信賴區間具有相同的寬度，且當然是實得分數25分的小明表現較佳。

(三) $20 \pm Z_{0.16}(6)\sqrt{1-0.7}$

$= 20 \pm 0.92(6)(0.548)$

$= 20 \pm 3.025$

$= (16.975,23.025)$

(四)利用KR-20信度可計算得知，題數由20題加倍至40題後，X測驗的信度爲：

$$\frac{2(0.64)}{1+(2-1)(0.64)} = \frac{1.28}{1.64} = 0.78$$

以200位學生進行預試結果，0.68的信度未能如預期提高至0.78，可能原因有：

1.加入之20題新題目與原有題目的同質性有狀況。

2.加入新題目的難度太高或太低，造成猜答或粗心錯誤。

二、(一)什麼時候適合使用配合題？其主要的限制為何？（10分）
(二)配合題的命題應注意那些事項？（10分）

試題評析	本題命題實屬偏頗，放眼目前各大考試幾乎不見配合題型，考生只能憑藉豐富的考場經驗與印象嘗試作答。如果未能順利作答，也是非戰之罪！相信絕大多數考生都面對一樣的問題，因此並不影響上榜與否。
考點命中	參考「國家教育研究院雙語詞彙、學術名詞暨辭書資訊網」。

答：

(一)配合題是由選擇題變化而來的一種型式，適用於測量概念與事實之間的關係。此種題目在結構上包括一爲問題項目；另一爲反應項目。通常係由後者中選出與前者相適合之項目，由於項目間之性質難求一致，易提供不適當之暗示，故標準化測驗較少採用，此爲其主要限制。

基於配合題的優缺點，此種題型的適用時機爲：

1.想利用較小試卷篇幅，短時間內測量記憶性資料

2.測驗目的在於想測知受試者低階的認知能力程度時

(二)在一般教師自編測驗中，如能顧及以下命題要領，此種題目仍有其測量學生成就之價值：

1.各題目及各反應（即答案）在性質上應力求相近，且按邏輯次序排列；

2.題目與答案數不宜相等；

3.配對項題目不可過多或過少，以五項至十項爲宜；

4.作答的方法必須予以明確的規定和說明；

5.同一組題目宜印在同一頁上，以免造成作答的困擾。

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。（10分）

試題評析	莖葉圖是教育行政公職考試常考的重要統計圖表，組成成分與特性在上課時就再三利用考古題進行充分練習，本班考生回答上應無太大困難與阻力。唯一挑戰與得分關鍵在於以圖說明的舉例。
考點命中	1.《教育測驗與統計》第一回，傅立葉編撰，Ch.2，頁7-10。 2.《高點教育測驗與統計》總複習教材，傅立葉編撰，頁9。 3.《高點103高普考重點題神：教育》，＜教育測驗與統計＞，傅立葉編撰，十一(一)。

答：

(一)莖葉圖又稱「枝葉圖」，它是將數字資料組中的數字經整理列示後，進行描述與比較的一種統計圖表。其依數字的大小將基本不變或變化不大的位數作爲一個主幹（莖），而將變化大的位數作爲分枝（葉），列

在主幹的後面，如此可以清楚地看到每個主幹（枝）後面的數字資料個數與資料分布情形。因此，枝葉圖的組成成分主要有三：

1.最左邊的一行數字，資料經大小排列與分組後，除最大資料組與最小資料組以個別組之資料筆數標示，以及中位數所在組（帶括弧）以該組資料筆數標示外，其他數字表示各組（枝）向上或向下往資料中心值組累積的資料筆數；

2.中間一行數字表示所欲分析描述之資料的十或以上位數的「莖」，其表示分析資料的組數；

3.右邊的則是歸入各組的資料個位數部分，也就是所謂的「葉」，其表示各組中的資料筆數。

莖葉圖的組成成分可以以下圖示說明：

	Stem	Leaf
3	2	2 5 9
10	3	1 1 1 2 5 7 8
18	4	0 2 3 4 4 5 7 9
(13)	5	0 0 0 1 4 5 5 5 6 6 7 8 9
19	6	1 3 3 4 4 6 7 7 8
10	7	2 4 5 5 8 9
4	8	1 4 4 8

(二)因為具備直方圖的雛形，可用以了解資料的分配是否呈現對稱或偏態情形；在電腦輸出報表中各枝的左側會提供各組葉(次)數或累積葉(次)數，可提供判斷並檢視在處理過程中是否遺漏任何單一的一筆資料。為唯一保留所有原始觀測值的統計圖。

四、(一)何謂抽樣誤差？抽樣誤差會受到那些因素的影響？（10分）
(二)什麼是統計檢驗力（power）？什麼因素會影響統計檢驗力？（10分）

試題評析	本題著重於推論統計的兩個基本且重要名詞概念。除個別名詞的意義外，影響因素也是答題得分關鍵。整體而言，對於本班考生或是有準備的考生而言，難度不高，拿高分易如反掌！
考點命中	1.《教育測驗與統計》第二回，傅立葉編撰，Ch.9，頁12。 2.《高點教育測驗與統計》總複習教材，傅立葉編撰，頁89。

答：

(一)抽樣誤差指的是舉凡涉及抽樣方法與過程所產生的誤差，例如：抽樣方法的選用不當、抽樣樣本過小所導致其母體代表性不足……等。

原則上，抽樣誤差可概分為系統性與非系統性誤差兩大類。其中系統性誤差指的是研究者可以操弄、控制、甚至預期、測量其數量大小的誤差，例如：在一項資料收集的活動中，受訪者統一接受某項暗示或激勵而一致的影響受訪者的回答所造成被收集資料之偏差；至於非系統性誤差則指的是隨機誤差，例如：在同一抽樣中，每位受訪者因個人對特定問題的解讀與認知所致在回答時刻意扭曲其答案，此類誤差通常不是研究者所能掌控、預期、與解釋的。

以單一母體平均數推論統計之信賴區間估計為例，影響抽樣誤差的因素全取決於計算公式的後半段：

$$\overline{X} \pm Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

也就是影響因素有：1.樣本大小　2.顯著水準　3.標準差　4.樣本平均數的抽樣分配型態(t或Z分配)

(二)假設檢定結論會有犯型I或型II錯誤的可能，結論為拒絕虛無假設且正確拒絕虛無假設之機率，即為統計檢驗力，以$1-\beta$表示。影響因素有五：

1.顯著水準的嚴謹與否

2.樣本大小

3.採取之單尾或雙尾的檢驗類型

4.樣本統計量的抽樣分配型態

5.樣本統計量與假設之母體參數真實值的差值大小。

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。（每小題5分，共25分）

(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？

試題評析	本題以簡單迴歸分析爲命題標的。在估計標準誤、非標準化直線迴歸方程式等的計算都相當容易。美中不足的是，最後一個子題的計算或因資訊不足，更可能因典試命題委員一時不察，誤以爲可以用簡單的信賴區間估計進行個別受試學生依變項的估計，而造成資訊不足，無法求最終數字解！
考點命中	《教育測驗與統計》第一回，傅立葉編撰，Ch.8，頁50-53。

答：

(一)依題意與題目提供之統計量資訊，可以利用決定係數的計算結果加以轉換而得：

$$1-0.8^2=0.36$$

因此，學生的學業成就分數之變異中，有36%是無法利用智力測驗分數加以解釋的。

(二)因爲相關係數沒有單位，且不受X、Y之單位的影響，可稱爲標準化係數。因此，即使以T分數進行直線轉換，所獲得的相關係數維持不變，也就是題目原本提供的0.8。

(三)依題意與題目提供的統計量資訊，可以分別計算非標準化迴歸方程式的斜率與截距如下：

$$b_1=0.8\frac{6}{15}=0.32$$

$$b_0=80-0.32(105)=46.4$$

因此，可得非標準化直線迴歸方程式爲 $\hat{y}=46.4+0.32x$

(四)估計標準誤將可計算而得

$$s_{Y.X}=6\sqrt{1-0.8^2}=3.6\text{(分)}$$

(五)對於智力測驗90分的學生而言，可以先利用迴歸預測方程式估計學業成就測驗預測分數，計算得：

$$\hat{y}=46.4+0.32(90)=75.2$$

再代入以下公式，進行實際學業成就測驗的95%信賴區間：

$$\hat{Y}\pm t_{(n-2,2/\alpha)}s^2{}_Y\left[1+\frac{1}{n}+\frac{(x-\bar{x})^2}{\sum x^2}\right]=75.2\pm t_{(n-2,0.025)}3.6^2\left[1+\frac{1}{n}+\frac{(90-105)^2}{\sum x^2}\right]$$

然而，因題目在樣本人數等的資訊不足，無法獲得最後真確數字答案。