

# 《教育測驗與統計概要》

一、請試述下列名詞之意涵：

- (一) 標準化測驗 (standardized tests) (8分)
- (二) T分數 (T-score) (8分)
- (三) 常態化標準分數 (normalized standard scores) (9分)

<b>試題評析</b>	本題具備解釋名詞的表徵，回答時卻須以簡答題的心態與方式應對。教育測驗與教育統計皆有命題，屬於綜合式命題。其中，「標準化測驗」的回答宜與常模參照測驗做連結；「T分數」與「常態化標準分數」則為統計學常用以衡量個別受試者表現情形的相對位置量數。
<b>考點命中</b>	1.《高點教育測驗與統計講義》第一回，傅立葉編撰，頁16-17。 2.《高點教育測驗與統計講義》第三回，傅立葉編撰，頁16-17。 3.《高點教育測驗與統計考猜》，傅立葉編撰。

**答：**

(一)標準化測驗：

標準化(心理)測驗的定義指的是在標準化的情境下，測量個人心理特質的工具或歷程。其使用原則含下列四點：

- 1.標準化測驗之實施必須顧及施測步驟、計分及結果解釋的一致性。
- 2.在施測步驟上，舉凡指導語、時間限制、物理環境和受試者的身心特質都應力求標準化。
- 3.測驗的計分要力求客觀與正確，結果解釋則需同時考量個案可能存在的身心障礙特性，測驗結果才能真實的反應個案的能力。
- 4.必須要挑選能適當評量個案能力的測驗，例如：市面上智力測驗的種類很多，但不是每一種測驗都適合特定個案使用，因此需先詳讀測驗手冊之內容(包括適用對象、測驗目的、常模與信效度、測驗題型等)。

(二)T分數：

T分數是教育統計學名詞，為麥考(W.A. McCall)所創，為尊崇心理學家桑代克(E.L. Thorndike)及特門(L.M. Terman)二氏，故取其姓氏第一字母T，命名為T分數。其標準差設定為10，平均數設定為50。為最常用的一種標準分數，屬用以判斷測驗表現好壞的一種相對位置量數。計算公式如下： $T = 50 + 10Z$ 。

(三)常態化標準分數：

常態化標準分數是標準分數之一。標準分數是一種具有數理特性而適於統計分析的衍生分數，為晚近測驗學界及教育界所普遍採用。標準分數係以平均數為參照點，用離均差的數值表示個別測驗分數在所屬團體中的相對位置。依轉換的方式而分，有直線標準分數(linear standard score)和常態化標準化分數兩種。無論原始分數(raw score)呈何種分配型態，經常態化轉換後的衍生分數(derived score)，必然變成常態分配的標準分數。由於中間參照點所設定的數值不同及離均差的尺度大小有別，在常態分配曲線下，各點各段所占百分比及累積百分比不一致，故有各種不同的標準分數。

二、常模參照測驗與效標參照測驗有何不同？當一個老師說要當掉最差的10%的學生，是那一種測驗的概念？(25分)

<b>試題評析</b>	本題在比較兩類測驗的差異，難度不高。考生只要能夠詳細有條理道出主要差異所在，掌握分數容易。
<b>考點命中</b>	1.《高點教育測驗與統計講義》第三回，傅立葉編撰，Ch.14，頁5-6。 2.《高點教育測驗與統計》考猜，傅立葉編撰。

**答：**

(一)效標參照測驗與常模參照測驗最大的不同在於：前者以課程目標作為參考系統，後者以某一個群體的表現作為參考系統。換句話說，效標參照測驗的標準是絕對的，是根據課程目標所定的，學生的表現水準決定

於該生達成課程目標的程度，與其他學生的表現水準無關；相對地，常模參照測驗的標準是相對的，是根據某一個群體的表現所定，學生的表現水準決定於該生與該群體比較的結果。因此，效標參照測驗的目的不在區辨學生的個別差異，而在評估教學或學習的成敗；常模參照測驗的目的則在區辨學生的個別差異，其結果通常用於淘汰或選擇人才。另外，兩類測驗的差異，還可以就其他幾點構面，比較列示如下：

常模參照測驗	效標參照測驗
a.針對一般化技能	針對特定技能
b.為實際作業達到的平均成績	為理想上應達到的成績
c.分數的變異性愈大愈好	分數的變異極小
d.以百分等級、標準分數表示	以及格、過關與否表示
e.常採選擇題、題目難易差距大	題目傾向於相同難度且較容易
f.內容取樣廣而淺	內容取樣窄而深
g.用於安置性、總結性	用於形成性、診斷性
h.出現較早	提出較晚，使用上較不為人所熟知
i.一次評多種技能，但無法徹底	需要很多這類問題來評多種技能

(二)當老師說要當掉最差的10%學生，本質上應屬於常模參照測驗。因為，測驗結果分數高低有別，所謂「最差的10%學生」指的是相對於全體同學，成績分數最不理想的比例學生，是一種相對標準，而非絕對標準。

三、某班教師針對全班30位學生進行考試準備時間和微積分期中考的得分進行分析，結果發現考試準備時間（小時）與考試分數之間的相關為0.6。已知30位學生這次考試準備的平均時間為15小時，標準差為5；微積分期中考的平均分數為52分，標準差為16；

(一)請算出使用考試準備時間來預測微積分期中考分數的非標準化迴歸方程式。(5分)

(二)某生考試準備了25小時，你預測此人的微積分期中考分數應為多少分？(5分)

(三)此處迴歸情境中的估計標準誤是多大？(5分)

(四)針對所有準備25小時的人，他們的微積分期中考分數95%的可能性是在幾分與幾分之間。

(請以四捨五入取概數至整數即可， $\bar{x}$  值落在  $\mu \pm 1.96\sigma_{\bar{x}}$  1.96，之間者佔95%) (5分)

試題評析	本題為簡單迴歸分析的非標準化迴歸模式之計算題，雖然子題不少，除第四小題外，難度不高。題目在(四)最後括號內的補充文字說明，有可能誤導考生使用傳統的平均數信賴區間估計進行計算與作答。
考點命中	1.《高點教育測驗與統計講義》第一回，傅立葉編撰，Ch.8，頁50-53。 2.《高點教育測驗與統計總複習講義》，傅立葉編撰，頁14-15。 3.《高點教育測驗與統計考猜》，傅立葉編撰，四(四)。

答：

(一)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：

$$b_1 = 0.6 \frac{16}{5} = 1.92$$

$$b_0 = 52 - 1.92(15) = 23.2$$

因此，可得非標準化迴歸方程式為  $\hat{y} = 23.2 + 1.92x$

(二)某生考試準備了25小時，利用前述迴歸預測方程式可以預測，此人微積分期中考分數為

$$\hat{y} = 23.2 + 1.92(25) = 71.2 \cong 71(\text{分})$$

(三)此處迴歸情境的估計標準誤，可以計算得出：

$$S_{y.x} = 16\sqrt{1 - 0.6^2} = 12.8(\text{分})$$

【版權所有，重製必究！】

(四)針對所有準備考試25小時的人而言，其微積分期中考分數的95%信賴區間估計可以計算得

$$\hat{Y} \pm t_{(n-2, 2/\alpha)} s^2_Y \left[ \frac{1}{n} + \frac{(x - \bar{x})^2}{\sum x^2} \right] = 25 \pm t_{(28, 0.025)} \cdot 12.8 \left[ \frac{1}{30} + \frac{(25 - 15)^2}{25^2 (30)} \right] = 25 \pm 2.048(12.8)(0.0387)$$

$$= 25 \pm 1.014 = (23.986, 26.014)$$

因此，取四捨五入後，所有準備考試25小時的人，其微積分期中考分數有95%落在24分與26分之間。

四、如果你手中有某單位員工的出生地點、在家中的排行與年齡的資料，

- (一)請分別說明這三者分別屬於何種層次的變數？(10分)
- (二)有什麼集中量數與離散量數可以用來描述這三個變數？(10分)
- (三)這三個變數中有那些可以進行四則運算？(10分)

<b>試題評析</b>	本題命題與答題重點在於考生須先於第一小題正確辨識三種資料的變數層次或尺度，然後依功能不同的集中與離散兩類統計量數，回答適當使用以提供正確資訊的第二小題，最後說明四則運算可以計算的資料與提供資訊及決策。
<b>考點命中</b>	《高點教育測驗與統計講義》第一回，傅立葉編撰，Ch.1，頁3；Ch.3、Ch.4。

**答：**

(一)針對三種員工資料的特性，出生地點屬於名目尺度，在家中的排行屬於等級尺度，年齡屬於比率尺度。

(二)三個變數依其不同資料特性，分別可以使用加以描述之集中與離散量數有：

1.出生地點：

因為只是城市的不同，既無法排序，計算大小好壞差異，可使用的集中量數有：眾數(統計與描述公司員工最多來自之城市)；離散量數部分則無任何適用者。

2.家中排行：

適用的只有屬於集中量數的眾數。

3.年齡：

可使用的集中量數有：平均數(用以描述員工平均年齡)、中位數(用以描述全體員工年齡高低的中間值)、以及眾數(描述公司最主要員工年齡的族群)；至於離散量數則有：全距(描述公司最老與最年輕員工的年齡差異)、平均差、標準差、變異數、相對差、四分位距，以及四分差等用來描述公司員工年齡差異的有用量數。

(三)三個變數中，可以進行四則運算的只有屬於最高階資料層次尺度的年齡，由於具備真正原點，四則運算可以充分提供資料使用者有用的資訊，作為公司整體員工結構的詳細探討與必要檢討，以作為後續配合公司發展甚至轉型之員工招募活動的決策參考。

【版權所有，重製必究！】