

101年公務人員特種考試原住民族考試試題

等 別：三等考試

類 科：教育行政

科 目：教育測驗與統計

一、下表三因子變異數分析摘要表為四年級、五年級和六年級男、女學生重複實施兩種不同測驗型式的統計分析結果，其中有些空格內的數字待你完成。回答時，請先繪製此表於試卷上，然後填入所有空格內該有的正確數值（均四捨五入至小數點第二位），並逐一分別陳述各交互作用、主要效果和事後比較的顯著性結果及其詮釋。（表內空格分數計20分；各交互作用、主要效果和事後比較的顯著性結果及其詮釋計10分；合計30分）

變異來源	SS	df	MS	F	p-value	事後比較
組間	326.36	11				
測驗（重複測量）	<input type="text"/>	<input type="text"/>	5.203	3.343	.068	
年級（獨立樣本）	<input type="text"/>	<input type="text"/>	100.201	13.332	.000	六>四 五>四
性別（獨立樣本）	<input type="text"/>	<input type="text"/>	85.392	11.362	.001	男>女
測驗×年級	<input type="text"/>	<input type="text"/>	0.970	<input type="text"/>	<input type="text"/>	.537
測驗×性別	<input type="text"/>	<input type="text"/>	1.645	<input type="text"/>	<input type="text"/>	.305
性別×年級	<input type="text"/>	<input type="text"/>	13.315	<input type="text"/>	<input type="text"/>	.171
測驗×年級×性別	<input type="text"/>	<input type="text"/>	2.575	<input type="text"/>	<input type="text"/>	.192
組內	3955.41	872				
受試者間S	3276.81	436	7.516			
殘差（A×S）	678.60	436	1.556			
全體Total	<input type="text"/>	<input type="text"/>				

【擬答】：

(一)編製 ANOVA 表

變異來源	SS	df	MS	F	p-value	事後比較
組間	326.36	11				
測驗（重複測量）	<input type="text" value="5.203"/>	<input type="text" value="1"/>	5.203	3.343	.068	
年級（獨立樣本）	<input type="text" value="200.402"/>	<input type="text" value="2"/>	100.201	13.332	.000	六>四
性別（獨立樣本）	<input type="text" value="85.392"/>	<input type="text" value="1"/>	85.392	11.362	.001	五>四
測驗×年級	<input type="text" value="1.94"/>	<input type="text" value="2"/>	0.970	<input type="text" value=".623"/>	.537	男>女
測驗×性別	<input type="text" value="1.645"/>	<input type="text" value="1"/>	1.645	<input type="text" value="1.057"/>	.305	
性別×年級	<input type="text" value="26.63"/>	<input type="text" value="2"/>	13.315	<input type="text" value="1.772"/>	.171	
測驗×年級×性別	<input type="text" value="5.15"/>	<input type="text" value="2"/>	2.575	<input type="text" value="1.655"/>	.192	
組內	3955.41	872				
受試者間S	3276.81	436	7.516			
殘差（A×S）	678.60	436	1.556			
全體	<input type="text" value="4281.77"/>	<input type="text" value="883"/>				

(二) 1. 變異數分析摘要表中可知，A因子（年級）與B因子（性別）的主要效果均達顯著水準。故不同四年級、五年級和六年級的依變項有顯著差異，不同性別的依變項有顯著差異。進行主要效果後六年級高於四年級、五年級高於四年級，男生高於女生。

2. 無論AB、BC、AC、ABC均無交互作用，無須進行單純主要效果的顯著性考驗。

公職王歷屆試題 (101 原住民特考)

二、甲、乙兩人在各科成績如下表所示，依據全部考生在各科的平均數 (\bar{X}) 與標準差 (SD)：1 請計算甲、乙兩人在各科的標準分數。2 若要在甲、乙兩人中擇一錄取，請問應該錄取誰？因應不同的選才需求或特殊考量而可能有不同的選擇結果，請依各種可能的考量或選才需求，分別陳述錄取的理由為何？

	國文		英文		數學		理化		史地		總分
全部考生	\bar{X}	SD	\bar{X}	SD	\bar{X}	SD	\bar{X}	SD	\bar{X}	SD	
	78	10	70	16	40	25	65	20	80	12	
甲	80		82		55		70		92		379
乙	68		54		90		90		80		382

【擬答】：

(一) 1. 甲生

國文

$$Z = \frac{X - \bar{X}}{S} = \frac{80 - 78}{10} = .2$$

英文

$$Z = \frac{X - \bar{X}}{S} = \frac{82 - 70}{16} = .75$$

數學

$$Z = \frac{X - \bar{X}}{S} = \frac{55 - 40}{25} = .6$$

理化

$$Z = \frac{X - \bar{X}}{S} = \frac{70 - 65}{20} = .25$$

史地

$$Z = \frac{X - \bar{X}}{S} = \frac{92 - 80}{12} = 1$$

2. 乙生

國文

$$Z = \frac{X - \bar{X}}{S} = \frac{68 - 78}{10} = -1$$

英文

$$Z = \frac{X - \bar{X}}{S} = \frac{54 - 70}{16} = -1$$

數學

$$Z = \frac{X - \bar{X}}{S} = \frac{90 - 40}{25} = 2$$

理化

$$Z = \frac{X - \bar{X}}{S} = \frac{90 - 65}{20} = 1.25$$

史地

$$Z = \frac{X - \bar{X}}{S} = \frac{80 - 80}{12} = 0$$

(二) 1. (1) 原始分數錄取乙生。因為乙生總分 382 分大於甲生 379 分。

(2) 原始分數的優點即是以受試者作答情形或答對題數直接作計算，簡單易懂，大部分人都明瞭計分方式；缺點即沒有考量測驗試題具有不同難度的事實，因而計分容易受到測驗難度的影響，而出現不公平、不精確的現象。原始分數測驗結果遴選資優班學生，可能會因為學生在某一科目分數拿取高分，導致測驗原始分數總分較高而被錄取，但未必是資優學生。

2. (1) 標準分數錄取甲生。因為甲生 $Z=2.8$ 分大於乙生 $Z=1.25$ 分。

(2) 標準分數不具有單位，優點即是可以顯示每位受試者在團體中的相對地位量數和排序大小，大多數的受試者尚容易理解該分數的意義；但其缺點即是容易受到極端分數 (Outliers) 的影響，而有扭曲次數分配的形狀之嫌，對個別受試者的表現會有過度放大或過度縮小之影響。標準分數測驗結果遴選資優班學生，能錄取學生在每一科目分數均為相對地位量數較高，較原始分數能正確篩選資優學生。

三、常使用的統計考驗的顯著性水準 (α) 有 0.05 與 0.01， α 值的設定將影響統計考驗的後果 (consequence)。試以「新款藥品欲推廣上市前，進行人體實驗，收集成效數據」為例，以第一類型錯誤和第二類型錯誤分別分析 α 值設定為 0.05 或 0.01 的可能不同後果與考量。

【擬答】：

(一) 1. 型 I 錯誤 (Type I Error) 或第一類型錯誤：

當虛無假設 H_0 為真，但依據統計考驗卻拒絕 H_0 ，認為 H_1 是對的，稱為型 I 錯誤。其中 α 風險係指犯型 I 錯誤之機率大小，在考驗中又稱顯著水準的風險。

$$\alpha = P(\text{拒絕 } H_0 \mid H_0 \text{ 為真})$$

2. 型 II 錯誤 (Type II Error) 或第二類型錯誤：

當虛無假設 H_0 為假，卻因統計考驗結果接受了 H_0 ，認為 H_0 是對的，稱為型 II 錯誤。其中 β

公職王歷屆試題 (101 原住民特考)

風險是指犯型 II 錯誤機率大小。

$$\beta = P(\text{接受 } H_0 \mid H_0 \text{ 為假})$$

(二)在心理與教育統計中，習慣上採用 $\alpha = 0.05$ 或 $\alpha = 0.01$ 作為假設檢定的顯著水準，代表所測量到關聯機率完全是因抽樣誤差產生，用以判斷虛無假設是否成立的機率水準， α 值大小主要視犯型 I 錯誤造成損失與後果，及研究者經驗態度過去文獻而決定，在某些情況下 α 亦可增減。有時亦稱危險區或拒絕區。犯型 I 錯誤機率愈小，即 α 愈小，研究結果顯著水準愈高；犯型 I 錯誤機率愈大，即 α 愈大，研究結果顯著水準愈低。

(三) H_0 ：新款藥品無效

H_1 ：新款藥品有效

1. 型 I 錯誤 (Type I Error)：「新款藥品無效是真的，卻因統計考驗結果，認為新款藥品有效」
2. 型 II 錯誤 (Type II Error)：「新款藥品無效是假的，卻因統計考驗結果，認為新款藥品無效」
3. 當其他條件不變情況下， α 與 β 具有互相消長關係，如 α 值增加，則 β 值減少，統計考驗力也會增加； α 值降低， β 值增加，相對統計考驗力會減少。二者間取捨必須由研究者取捨到最佳境界。
4. 在實驗研究領域中，研究者通常認為犯型 I 錯誤是比較嚴重且不可原諒，因此多數研究者願意犯較大型 II 錯誤而不願犯型 I 錯誤。例如：製藥廠推出新藥稱可治肝癌，此一推論只要犯一點點錯誤就會遭致嚴重結果。如服用該藥 100 人中，治好 99 人肝癌，卻存有治死 1 人的可能；因此我們寧願宣稱此藥無效，其錯誤並不嚴重，頂多只是該藥無法上市的錯誤。

四、請比較客觀式評量與多元評量兩者在評量理念、評量方式、信度與效度考量之異同。

【擬答】：

(一)客觀式評量與多元評量意義

1. 客觀式評量：評分方式較為客觀固定，不論由任一評分者，在不同情境下，其評分結果都一致，甚至可藉由電腦機械評分。
2. 多元評量：除認知評量（傳統紙筆測驗）尚須考量情意與技能評量，不同評量範圍與學習目標，其評量工具亦應不同，因此使用評量也應是多元的，多元評量是手段而非目的、並非廢除紙筆測驗、莫為多元而多元、並非一種評量；多元評量適合「過程知識」(Procedural Knowledge) 的檢驗，較重視思考歷程評量、學生學習精熟進步情形。

(二)異同

1. 評量理念：客觀式評量評定等級，具有鑑別功能分類編組功能；多元評量瞭解學生是否達到理想標準，幫助學生發展潛能與瞭解問題，達到最佳的學習標準。
2. 評量方式：客觀式評量教師自編測驗、標準化成就測驗、紙筆性向測驗、是非題測驗、配合題測驗與選擇題測驗皆屬之；多元評量只是一種通稱，又稱變通式評量 (Alternative Assessment)、另類評量、替代評量；包含內容有一實作評量、卷宗評量、真實評量、動態評量、生態評量、課程本位評量。
3. 信度與效度考量：
 - (1)客觀式評量評分客觀、公正、公開，題目編製容易大都有標準答案，因此信度與效度較高；客觀式評量將受試者得分之變異量分成個體間變異量及系統或隨機誤差兩種因素。
 - (2)多元評量評量受試者學習歷程，題目大都編製不易無標準答案，信度與效度較低或準確度可疑，尤其評分者信度更低。多元評量擴大為受試者 (Person)、試題 (Item)、評分者 (Rater)、測驗情境 (Occasion) 和相互間交互作用等多項因素。