

A. 心理與教育統計（共計 33%）

A-1. 證明：將變項 X 轉換成 Z 分數，則轉換後的 Z 分數之平均數為 0，變異數為 1。（3%）

A-2. 證明：樣本平均數抽樣分配（sampling distribution of the mean）的平均數為 μ ，變異數為

$$\frac{\sigma^2}{n} \quad (\mu \text{ 及 } \sigma^2 \text{ 為母群的平均數及變異數，} n \text{ 為抽樣數})。 \quad (6\%)$$

A-3. 證明： $s^2 \sim \frac{\sigma^2}{n-1} \chi_{n-1}^2$ ，即樣本變異數為自由度 $n-1$ 的卡方分配（ s^2 及 σ^2 分別為樣本及母群的變異數）。（6%）

A-4. 證明：當進行只有兩個類別的單向度卡方檢定，其統計量 $\chi^2 = \sum_{i=1}^2 \frac{(O_i - E_i)^2}{E_i} \sim \chi_1^2$ ，即自由度為 1 的卡方分配（ O_i 與 E_i 分別表示類別 i 之觀察及期望值）。（6%）

A-5. 證明：進行兩個獨立樣本平均數 t 檢定，假設兩樣本之變異數同質，因此混合（pool）起來算成 s_p^2 ，則此檢定之統計量 $t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \sim t_{n_1+n_2-2}$ ，即自由度為 n_1+n_2-2 的 t 分配（ n_1 及 n_2 分別為兩個樣本之人數）。（6%）

A-6. 證明：進行有 J 個類別之 one-way ANOVA 平均數 F 檢定，則此檢定之統計量

$$F = \frac{SSB / df_B}{SSW / df_W} \sim F_{J-1, N-J} \quad (\text{即自由度為 } J-1, N-J \text{ 的 } F \text{ 分配 (} N \text{ 為總人數；B=between, W=within)})。 \quad (6\%)$$

B. 心理測驗（共計 33%）

B-1. 個別差異（individual differences）之區辨為心理測驗的重要目的之一，請問在編製心理測驗時，研究者可如何進行分析以挑選具區辨（discrimination）力的題目？又，請問所有心理測驗在選題時，均需以區辨力為考量挑選題目嗎？為什麼？（10%）

B-2. 請說明測量標準誤（standard error of measurement）、信度（reliability）、效標關聯效度（criterion-related validity）間之關係。（7%）

B-3. 一研究者編製台灣國小學童注意力測驗，期能藉之發現在學習上可能需要協助的學童。在建構常模前，該研究者請教你宜如何取樣以組成常模樣本，請問你會建議如何選取常模樣本呢？並請說明你的理由。（8%）

B-4. 「國民中學學生基本學力測驗」是台灣當前社會一個重要的心理測驗。（a）請就你對此測驗的瞭解，論述此測驗的編製目的與類型，並（b）請說明與評論此測驗在目前台灣教育系統之功能。（8%）

見背面

C. 心理實驗法 (共計 34%)

C-1. 當受試者之間的異質性越高時，如果其他條件都一樣，則受試者內 (within subjects) 設計會比受試者間 (between subjects) 設計越更容易得到顯著結果。請你說明其道理。(5%)

C-2. 研究者想進行實驗，觀察操弄某一因子 A 可能產生的效果，但文獻也顯示另一因子 B 會產生穩定的影響效果。那麼，若其他條件都一樣時，採用 A x B 的二因子實驗設計，會比採用 A 的單因子實驗設計，更容易發現 A 的真正效果。請兼顧因果推論邏輯與統計分析的角度進行思考，以闡明其道理。(6%)

C-3. 古典心理物理學企圖尋找心理量與物理量之間的函數關係。但 Thurstone, L. (1927) 卻提出了比較判斷法則 (law of comparative judgment)，號稱可以不必仰賴物理向度的測量，也可以對不同刺激事件進行滿足等距尺度 (interval scale) 的心理度量 (scaling)，例如心理壓力程度或漂亮程度的測量。雖然已經經過 80 年，但翻查文獻，可以發現該方法仍然被大量使用。請簡單說明 Thurstone 比較判斷法根據的道理。(6%)

C-4. 心理物理學實驗古典閾值理論

(a) 在心理物理學中，閾值量測一直是相當重要的實驗工具。然而，Weber(1832)或 Fechner(1850)所指的閾值概念和現今科學家所使用的閾值概念有很大的差別。請比較古典及近代閾值概念的異同。(2%)

(b) 假設 A 先生恰好能區分 500 克及 550 克物體的重量差別。現在，A 先生手持一兩公斤的重物，根據韋伯定律 (Weber's law)，我們還需要增加多少重量在兩公斤重物上才能讓 A 先生察覺到重量的不同。(1%)

(c) Fechner(1850)提出了一個理論來解釋韋伯定律 (Weber's law)，試述此理論。(3%)

C-5. 心理物理學實驗測量方法

假設你想要知道在一個有雜訊的雷達屏幕上，可讓雷達操作員偵測到的最小目標大小。根據文獻，已知在一個清晰的屏幕上，以 50% 的正確率測量到人可感知到目標大小的閾值是 90 分視角 (visual angle)。因此，你預期在一個有雜訊的屏幕上閾值應該落在 90 分到 150 分之間，而屏幕解析度為 10 分。

(a) 假設你希望使用定值刺激法 (constant stimuli method) 來測量其閾值。你會如何著手測量？請寫出你的設計、程序、以及資料分析方法。(6%)

(b) 假設你想估計雷達操作員的反應偏誤 (response bias)，你將如何由 (a) 小題搜集的資料中計算反應偏誤？(提示：使用信號偵測理論)(2%)

(c) 其中一位同事建議你嘗試使用階梯法 (staircase method)。請問使用定值刺激法和階梯法的優缺點為何？(3%)

試題隨卷繳回