

# 《統計實務概要》

試題評析	四等的統計實務考的內容與統計學、抽樣、迴歸有關，因此考統計人員的考生，若在這些考科有完整的準備，應該可以拿高分。
考點命中	1.《高點·高上統計實務講義》第三回，盛華仁編撰，頁7-8。 2.《高點·高上統計實務講義》第一回，盛華仁編撰，頁45。

一、某社區50歲以上居民共1000人，以65歲為切點，其年齡及性別分布如下表：

	年齡<65	年齡≥65	合計
男	320	160	480
女	320	200	520
合計	640	360	1000

為了解這1000人的健康狀態，採隨機抽出100人作為一組代表性樣本。試求：（每小題10分，共30分）

（一）請設計一個分層抽樣法（stratified sampling）選出所需之100人。

（二）請設計一個系統抽樣法（systematic sampling）選出這100人，且性別與年齡的比例結構要符合上表的比例。

（三）請設計一個群集抽樣法（cluster sampling）選出這100人，且性別與年齡的比例結構要符合上表的比例

**答：**

（一）第一層為男性年齡小於65歲

$$\text{抽出 } 100 \times \frac{320}{1000} = 32 \text{ 人}$$

第二層為男性年齡大於65歲

$$\text{抽出 } 100 \times \frac{160}{1000} = 16 \text{ 人}$$

第三層為女性年齡小於65歲

$$\text{抽出 } 100 \times \frac{320}{1000} = 32 \text{ 人}$$

第四層為女性年齡大於65歲

$$\text{抽出 } 100 \times \frac{200}{1000} = 20 \text{ 人}$$

（二）將這四個族群的人，分別給予編號且將每個族群的人分成十等份，先從第一個等份中以簡單隨機抽樣抽出1人，再每間隔10人抽出1人，此為系統抽樣法。

（三）將這四個族群的人，分別分成10個性質相同的群集，再從這10個群集中各抽1個群集做為代表性樣本，此抽樣法稱為群集抽樣。【版權所有，重製必究！】

二、新聞報導臺灣近年大腸癌發生率是第一名、肺癌則是死亡率第一名。試問：

（一）如何定義癌症的發生率和死亡率？（8分）

（二）大腸癌發生率第一但肺癌是死亡率第一，這是否有矛盾？請解釋之。（7分）

**答：**

（一）癌症發生率：每十萬年中人口數中，新發生癌症的比例。

癌症死亡率：每十萬年中人口數中，因癌症而死亡的人數。

(二)大腸癌發生率第一的意思是表示，當年度新得到大腸癌的人數最多，但不表示當年度因大腸癌死亡的人數最多，肺癌則是死亡率第一名，而是指因肺癌而死亡的人數最多，所以二者並無矛盾。

三、某政府單位統計部門想委託民意機構調查民眾對新施政的滿意度。設滿意的百分比為 $p$ ，今欲進行一個調查，希望在95%的信心水準下將抽樣誤差控制在3%左右。假設每施測100份問卷平均只能回收50份，且每份問卷的施測成本是200元，則大約要編列多少預算？（已知標準常態分布的2.5%分位點為-1.96、97.5%分位點是1.96；且根據一個較小型的前期調查， $p$ 約在60%上下。）（20分）

**答：**

$$e = Z_{0.025} \sqrt{\frac{P(1-P)}{n}} \Rightarrow 0.03 = 1.96 \sqrt{\frac{0.6 \times 0.4}{n}} \Rightarrow n = 1024.43$$

施測份數  $1024.43 \times 2 = 2048.86$

預算要編列  $2048.86 \times 200 = 409,772$ 元

四、政府在實施關於政策滿意度或民眾意向、觀感、需求等調查時，不管是經由網路、電話訪問或者面訪等，都不免需要先做好問卷的設計與效度、信度之測試。請舉出至少2種“信度”（reliability），並說明其意義。（15分）

**答：**

信度是指可靠性或一致性。信度好的指標在同樣或類似的條件下重複操作，可以得到一致或穩定的結果。信度有三種：

(一)穩定信度：

這是一種長期的信度。也就是指標在不同時間做測量時，可以得到同樣的結果。如果每次都得到同樣的結果，則此指標即有穩定信度。

(二)代表性信度：

代表性信度是橫跨各個次母體或群組的信度。也就是指標用於不同次母體或群組（如年齡、性別）時，可以得到同樣的結果。例如：對年齡的測量，應該是詢問不同年齡層時，都可得到一致性的資訊，不論此指標是正確的問到年齡，還是有同樣方向的偏誤，如以多報少。

(三)同等信度：

同等信度是應用在利用多重指標測量同一構念的情況。可以用如Cronbach's  $\alpha$ 這種統計方法來做此種信度分析。如果多重指標有相當的信度，就可利用這些指標來建構量表。另一種同等信度的特殊分析方法是做編碼者間信度（intercoder reliability）的分析。當用多位觀察者、評判者或編碼者時就可用此方法。其目的是檢視不同的觀察者或編碼者是否彼此間的意見一致。

五、根據過去數年的資料，某國家發現他們的國內生產毛額（GDP）變動率和失業率（UEM）之間有一個線性關係。在此，我們定義第 $t$ 年的GDP變動率（簡記為 $\Delta GDP(t)$ ）如下：  
（ $t=1, 2, \dots, 8$ ）

$$\Delta GDP(t) = \frac{GDP(t) - GDP(t-1)}{GDP(t-1)} \times 100\%$$

以下是該國過去8年間的GDP變動率和失業率（UEM(t)）的數值。

$\Delta GDP(t)$	-0.29	2.53	4.03	0.96	-1.68	3.43	2.85	1.76
UEM(t)	4.39	4.24	4.18	3.96	4.48	3.92	3.76	3.71

某人以 $\Delta GDP$ 當作自變數、以UEM當作因變數，進行單變數之迴歸分析後得到下表的分析結果：

參數估計(parameter estimation)					
變數	自由度	參數估計	標準誤	T值	p值
截距項 (intercept)	1	4.221	0.124	34.2	<0.0001
$\Delta GDP(t)$	1	-0.083	0.045	-1.66	0.021

試求相應於上表之迴歸分析模型，並根據此表結果解釋該國之國內生產毛額和失業率之關係。(20分)

**答：**

迴歸分析方程式：

$$UEM_t = 4.221 - 0.083\Delta GDP_t$$

根據迴歸分析報表的結果， $\Delta GDP_t$  的p值為0.021，小於顯著水準  $\alpha = 0.05$ ，表示拒絕虛無假設，表示  $\Delta GDP_t$  對  $UEM_t$  有顯著負向的影響，亦即  $\Delta GDP_t$  每增加1%，失業率會減少0.083%。

【版權所有，重製必究！】