

應用統計講義

第一回

302200-1



社團法人 **考友社** 出版發行

應用統計講義 第一回



| | |
|---------------|----|
| 第一講 基本概念..... | 1 |
| 命題大綱..... | 1 |
| 重點整理..... | 2 |
| 一、緒論..... | 2 |
| 二、統計繪圖..... | 11 |
| 三、敘述統計..... | 28 |
| 精選試題..... | 68 |

第一講 基本概念



一、緒論

- (一)統計學之意義
- (二)統計學之分類
- (三)母體與樣本
- (四)資料相關名詞
- (五)資料蒐集方法
- (六)衡量尺度
- (七)變數類別

二、統計繪圖

- (一)屬質性資料次數分配
- (二)屬量性資料次數分配
- (三)莖葉圖
- (四)長條圖
- (五)圓餅圖
- (六)直方圖
- (七)折線圖
- (八)肩形圖

三、敘述統計

- (一)概述
- (二)中央傾向
- (三)離散度
- (四)偏態與峰態
- (五)百分等位與百分位序
- (六)盒鬚圖

*
* **重點整理** *
*

一、緒論

(一)統計學之意義：

1. 現今人類所面對的各種現象已愈趨複雜多樣、雜亂紛陳，爲了從這些大量、紛亂、具有不確定性的現象裡，找出其中有用的訊息與原則，一種專爲處理不確定問題的科學方法便應運而生，這就是所謂的統計學（Statistics）。
2. 簡而言之，統計學就是一種有系統的科學方法與研究工具。它運用蒐集、整理、分析、解釋資料等程序，將其轉化爲簡潔有序、脈絡分明的資訊；同時藉由嚴謹、科學化的推論與預測，可在不確定狀況下做出明智的決策。
3. 目前統計學已廣泛應用在許多領域，如個人或家庭的收支預算或保險理財、企業的經營決策或產品定位、國內外經濟問題的研究、政府公共政策的釐定、全球人口成長的結構分析…等。由此可見，統計學並非針對某一特定領域探究，而是一種可用於許多學門的研究方法。
4. 其他像是溫室效應與聖嬰現象對全球造成的影響爲何？人民對政府官員的滿意度爲何？調漲大學學費是否會影響升學意願？投資股票是否會賺錢？抽菸、嚼檳榔和癌症之間是否有直接關聯？核四廠的興建是否真的有其迫切性？…等，也都可以透過統計的方法來分析、解釋、甚至預測。
5. 雖說統計學的應用領域和應用方法不斷地擴展，但統計方法的應用並非毫無限制。茲歸納應注意的兩點如下：
 - (1)資料應正確。
 - (2)方法應正確。

(二)統計學之分類：

1. 依分析順序：

(1)敘述統計（Descriptive Statistics）：

又被稱爲資料分析（Data Analysis），是統計方法的初步程序。其主要的目的是將原始資料作一番整理、組織和表述，並且利用各種統計圖表及統計量的計算，使統計的結果清楚展現出來。這部分的過程相

當繁瑣，但卻是後續分析、推論最至關緊要的基礎。

(2)推論統計 (Inferential Statistics) :

主要目的則在於由樣本資料推估母體參數，以作為策略擬定之決策依據。例如，飲料廠商想瞭解所開發的產品甜度是否適中，則可藉由適當的抽樣，抽取代表性的消費者，進行口味測試，並進而推估全市場對甜度接受的情況。

2. 依變量多寡：

(1)單變量統計 (Univariate Statistics) :

探討的變量只有一個，是統計學領域中最根本、也最易理解的基礎分析，如單因子變異數分析 (One-Way ANOVA)、簡單迴歸分析、複迴歸分析、時間序列分析…等。

(2)多變量統計 (Multivariate Statistics) :

為單變量統計的延伸，所探討的變量則有兩個或兩個以上，如多變量迴歸分析、多變量變異數分析 (MANOVA)、鑑別分析、主成份分析、因素分析、集群分析、多元尺度分析…等。

3. 依研究範疇：

(1)理論統計 (Theoretical Statistics) :

為統計的學理基礎，著重在公式的推導及定理的證明，故較偏向數理統計方面的討論，須善加活用微積分、線性代數…等工具。例如，貝氏定理、常態分配理論、中央極限定理…等，均涉及較深入的數學推演，屬理論統計的範疇。

(2)應用統計 (Applied Statistics) :

用以解決真實世界的諸多問題，其目的在善用理論統計中的各種定理或公式，支援決策之制定。例如，變異數分析、迴歸分析、時間序列預測…等，均以理論統計為基礎，應用到諸多研究領域。以企業管理而言，不論作業管理、行銷管理、人資管理、資訊管理、財務管理，只要稍微深入的分析，都離開不了統計的應用。

4. 依母體假定：

(1)母數統計 (Parametric Statistics) :

在統計推論前，須先假定母體的分配型態。一般而言，母數統計適合用在像社會科學、管理科學等大樣本的研究。統計學領域中，幾乎大部分的探討主題都是母數統計。例如，單母體假說檢定、雙母體假說檢定、變異數分析…等，均假定統計量遵循常態分配，為典型的母數統計。

(2)無母數統計 (Non-Parametric Statistics) :

無須假定母體的分配型態。一般而言，無母數統計適合用於生命科

♥♥♥♥♥♥♥♥♥♥♥♥♥♥♥♥
♥
♥ **精選試題** ♥
♥♥♥♥♥♥♥♥♥♥♥♥♥♥♥♥

一、下表為甲班的基本資料和數學期中考成績：

| 姓 名 | 學號 | 性別 | 期中考成績 |
|-----|----|----|-------|
| 丁大維 | 10 | 男 | 95 |
| 王文中 | 25 | 男 | 68 |
| 張海婷 | 33 | 女 | 43 |
| 羅書豪 | 39 | 男 | 71 |
| 陳玉芬 | 51 | 女 | 82 |

- (一)試找出這個資料檔的變數。
 (二)任舉一觀察值，並說明其意義。

答：(一)變數為「姓名」、「學號」、「性別」、「期中考成績」。

- (二)選取「丁大維」和「期中考成績」兩者的交集，得到「95」，95 就是一個觀察值，它代表了丁大維在這次數學期中考得到 95 分。

二、一筆資料中取出 5 個觀察值，分別為 1.3、7.0、3.6、4.1、5.0，試回答下列各題：

- (一)算數平均數為何？
 (二)說明(一)所得之平均數為該筆資料的平衡點。

(三)以上列觀察值說明 $\sum_{i=1}^n (X_i - \bar{X}) = 0$ 。

(四)以上列觀察值印證 $\sum_{i=1}^n (X_i - \bar{X})^2 < \sum_{i=1}^n (X_i - A)^2$ ， $A \neq \bar{X}$ 。

答：(一)算數平均數 $\bar{X} = \frac{\sum_{i=1}^k X_i}{n} = \frac{1.3 + 7.0 + 3.6 + 4.1 + 5.0}{5} = 4.2$

- (二) 1. 比 4.2 小的數值有 1.3、3.6、4.1，距離和為：

$$|1.3 - 4.2| + |3.6 - 4.2| + |4.1 - 4.2| = 3.6$$

2. 比 4.2 大的數值有 7.0、5.0，距離和為：

$$|7.0-4.2|+|5.0-4.2|=3.6$$

3. 故知算數平均數是整個資料的平衡點。

(三)(1.3-4.2)+(7.0-4.2)+(3.6-4.2)+(4.1-4.2)+(5.0-4.2)=0，故得：

$$1. \sum_{i=1}^n (X_i - \bar{X}) = 0$$

$$2. \sum_{i=1}^n (X_i - \bar{X})^2 \\ = (1.3-4.2)^2 + (7.0-4.2)^2 + (3.6-4.2)^2 + (4.1-4.2)^2 + (5.0-4.2)^2 \\ = 17.26$$

(四) 1. 令 A 為不等於 4.2 之任意數，如 A=4.1，則：

$$\sum_{i=1}^n (X_i - A)^2 = 17.31$$

2. 故可印證：

$$\sum_{i=1}^n (X_i - \bar{X})^2 < \sum_{i=1}^n (X_i - A)^2$$

三、甲、乙、丙三班之學生人數分別為 50、55、48 人。設三個班級的平均體重分別為 60、56、59 公斤，試求三個班級的總平均體重。

答：(一)這是好幾組資料在給定各組平均數的情況下，求共同平均數。先計算出每班的總體重後相加，再除以三個班級的總人數，即：

$$\mu = \frac{n_1 \cdot \mu_1 + n_2 \cdot \mu_2 + n_3 \cdot \mu_3}{n_1 + n_2 + n_3} = \frac{50}{153} \times 60 + \frac{55}{153} \times 56 + \frac{48}{153} \times 59 \\ = 58.248$$

(二)其中，

$$W_1 = \frac{50}{153}, W_2 = \frac{55}{153}, W_3 = \frac{48}{153}$$

(三)故這三個班級的平均體重是 58.248 公斤。

四、假定某院學生分為甲、乙、丙三組上微積分課，甲組有學生 40 人，平均成績 80 分，標準差 8 分；乙組有學生 30 人，平均成績 75 分，標準