

Quiz 1 for Statistics (2011/10/18)

姓名: 學號:

1. Describe the difference between the descriptive statistics and inferential statistics in words and provide the examples for the descriptive statistics and inferential statistics, respectively (20%)

Ans:

Descriptive statistics: 藉由適當的圖形表格與數字將資料組織、彙整以及表達成方便讀取資訊的統計方法。(7%)

Examples: Bar chart、histogram、boxplot、frequency table、scatter plot、平均數、中位數、變異數……(3%)

inferential statistics: 研究如何根據樣本數據去推斷總體數量特徵的方法。它是在對樣本數據進行描述的基礎上，對統計總體的未知數量特徵做出以機率形式表述的推斷。(7%)

Examples: 估計、hypothesis testing……(3%)

2. Describe the following terminologies in words (30%).

◆ Population vs. Sample; Parameter vs. Statistic; Census vs. Sampling

Ans:

Population: 在研究或實驗目的下，所有可能研究對象所形成的群體。這個群體通常是非常地大，並且在實務上還可能會是無限地大。(5%)

Sample: 從母體以隨機或非隨機方式抽取部分資料。可能非常地大，但是小於母體。(5%)

Parameter: 為一數值型的數量(numeric quantity)，通常為未知數，用來描述特定研究母體之特徵(敘述統計量)。(5%)

Statistic: 為一數量(quantity)，通常由樣本資料計算求得，並用來估計未知的參數。(5%)

Census: 將整個母體納入觀察，並以全部母體作為調查對象。(5%)

Sampling: 以抽樣的方式從母體之中抽取一部分資料作為樣本，針對抽取的樣本進行調查，然後就其調查結果推論母體的特徵。(5%)

3. Describe the types of data in the real world, and provide the examples for these data types (20%).

Ans:

Interval data: 資料型式是數值型的且代表真實的數字，同時可進行所有的數學運算。該資料也可以被視為順序或名目資料來處理，只是會造成資料訊息的流失。(4%)

Examples: 身高、體重、所得和距離之類。(2%)

Nominal data: 以數值型態代表沒有存在任何順序之類別，且僅能藉由數的過程 (counting process) 來計算各種類別所發生之次數。資料不能被視為順序資料或區間資料來處理。(5%)

Examples: 使用編碼(codes)來記錄婚姻狀態：單身 = 1，已婚 = 2，離婚 = 3，鰥寡 = 4。(2%)

Ordinal data: 以數值型態代表具有順序的類別，立基於排序過程的運算是有效的。藉由數的過程 (counting process) 來計算各種類別所發生之次數，資料可以被視為名目資料但不可以被視為區間資料來處理。(5%)

Examples: 不好 = 1，普通 = 2，良好 = 3，很好 = 4，非常好 = 5 代表順序資料的編碼。(2%)

4. Use the interval data, {6,9,12,8,22,11,19,13,17,3}, to accomplish the following items.

(a) Mean, median, variance, coefficient of variation, 80th percentile (10%)

(b) Histogram, Boxplot (10%)

Ans:

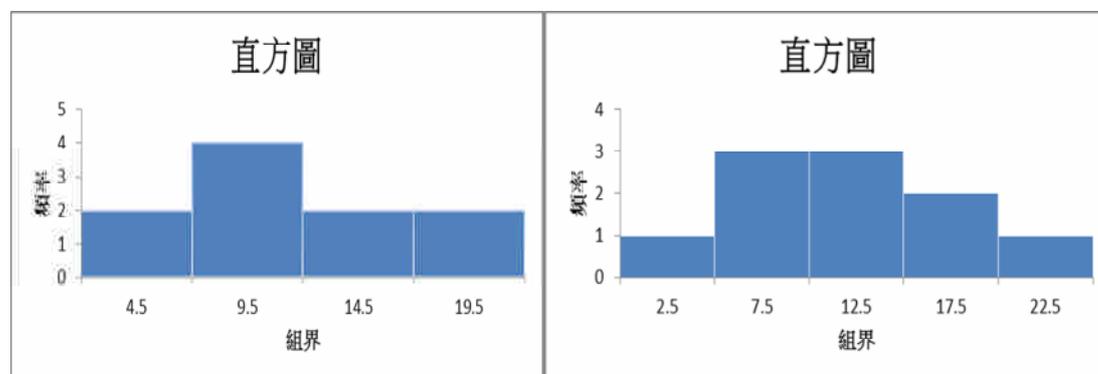
(a) Mean:12 ; Median=11.5 ; Population variance:31.8 ; Sample variance:35.3 ;

Coefficient of variation:0.468 or 0.495 ; 80th percentile:18.6

(b)Histogram:

組數： $1+3.3\log(10)=4.3$ 取 4 組

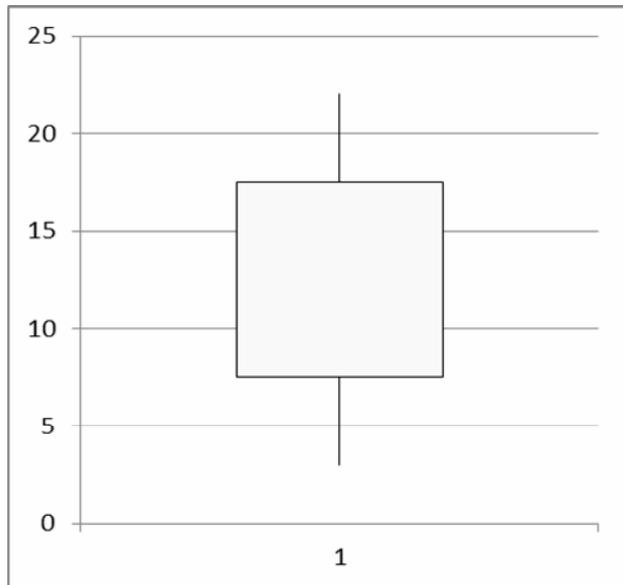
寬度： $(22-3)/4=4.75$ 取寬度 5



Boxplot:

Min:3 ; Q1:7.5 ; Median=11.5 ; Q3:17.5 ; Max:22

$1.5 \times (17.5 - 7.5) = 15$ 所以本題無離群值，只需延伸到最小最大值即可。



5. What is the statistics meaning to you? (5%)

Ans:

對我而言是一門可判斷數字的真偽，繼而做正確的決策的科學。(無標準答案)

6. Why do people need the statistics? (5%)

Ans:

人們可以藉由統計學來研究數據現象，這種現象如社會現象、自然現象...等。例如一年之間每一日的平均氣溫，或臺北市去年每一個市民的入息等等。(無標準答案)