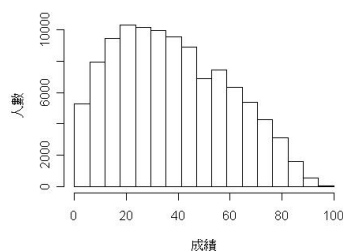


Introduction to Statistics

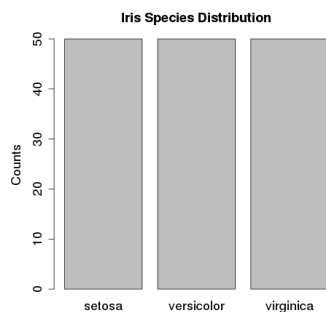
- Statistics: the science of handling data.
 - Descriptive statistics (敘述統計): to give a summary of observed data using numbers or graphs.
 - Inferential statistics (推論統計): to draw conclusions about certain quantities determined by the data generating rule based on the observed data.
- Examples of using descriptive statistics.
 1. Exam score data (91學年度指定科目考試英文成績)
 - Histogram (直方圖)



- 平均成績: 37.75619; median (中位數): 36.
 - 2. Species in the Iris flower data set (鳶尾花卉數據集的種類變數)
 - Relative frequency table (相對次數表)

Species	Counts	Relative frequency
setosa	50	1/3
versicolor	50	1/3
virginica	50	1/3
Total	150	1.000

- Bar chart (長條圖)



- The ways in which data are described/presented often depend on the data types.
- Types of Variables
 - Qualitative (質性的) Example: gender
 - Quantitative (數量型)
 - * Discrete (離散型). Example: counts
 - * Continuous (連續型). Example: temperature.
 - * Other
- Levels of measurement (scale)
 - Nominal (次序沒有意義).
 - Ordinal (次序有意義, 但同樣差距意義可能不同).
 - * Example: 滿意度 1-4 (4=非常滿意, 3=滿意, 2=不滿意, 1=非常不滿意)
 - Interval (同樣差距意義相同, 0無特殊意義)
 - Ratio (0=沒有, 比例是有意義的)
- Inferential statistics (推論統計): 根據觀察到的資料去推論資料生成的法則或由此法則決定的參數.
 - Example. 觀察 n 次丟銅板的結果, 正面為1, 反面為0. 則資料生成法則由銅板出現正面的機率決定. 根據 n 筆資料估計出現正面的機率即屬於推論統計問題.
 - Example. 有市長候選人3名, 隨機選取一些選民詢問他們所支持的候選人(結果以1,2,3表示). 則資料生成法則由全部選民數, 選取選民數以及3名候選人支持率決定. 根據受訪者回答結果估計3名候選人支持率也屬於推論統計問題.
- 觀察到的資料若為 $X_1, \dots, X_n$, 則 $(X_1, \dots, X_n)$ 稱為一組樣本(sample), 而 n 稱為樣本數 (sample size).
- Software R
 - R can be downloaded at <https://cran.r-project.org/>
 - R handout is available at
<https://stat.walkup.tw/teaching/statistics/handouts/r.pdf>