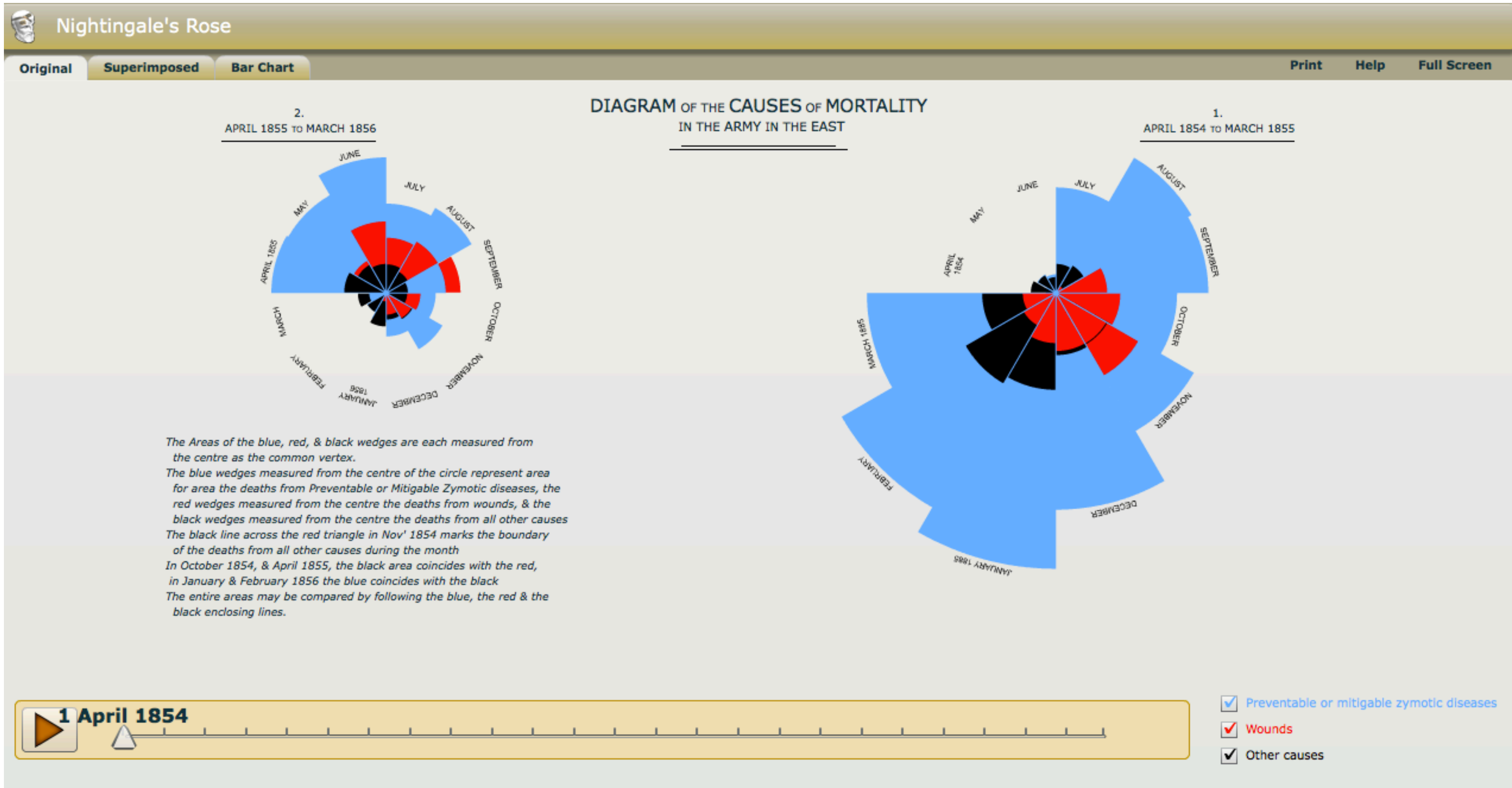
THE UNIVERSITY OF WESTERN AUSTRALIA

BUSINESS SCHOOL

MGMT5504

SEMESTER 2 2022 – LECTURE 3



Module 2 – Probability Distributions and Inference

Probability Essentials

- Probability – the chance that an uncertain event will occur (always between 0 and 1)
- Concept of probability is quite intuitive; however, the rules of probability are not always intuitive or easy to master.
  - An event with probability zero cannot occur.
  - An event with probability 1 is certain to occur.
  - An event with probability greater than 0 and less than 1 involves uncertainty, but the closer its probability is to 1 the more likely it is to occur.

Interpretations of Probability

Objective and Subjective

- **Classical** - Propensity – i.e. logic analysis (dice)
- **Empirical** - Relative frequency in the long run (mortality tables for calculating insurance premiums)
- **Subjective** – Personal and subjective judgement (reliability of witness testimony)

Probability Essentials (cont'd)

- Experiment – a process that produces outcomes for uncertain events
- Sample Space– the collection of all possible experimental outcomes
- Event – a subset of the sample space
  - Collectively Exhaustive events
  - Mutually Exclusive events
- Event A union Event B : $A \cup B$
- Event A intersect Event B: $A \cap B$
- Complement of Event A : A^C or $\bar{A}$

Example

A snowboarder athlete attempts to predict her chances of earning a medal in the women's halfpipe.

Sample Space S={gold, silver, bronze, no medal}

Event A: 'At least earning a silver medal'

Event B: 'At most earning a silver medal'

Event C: 'Winner of the competition'

$A \cup B = S$; $B \cup C = S$

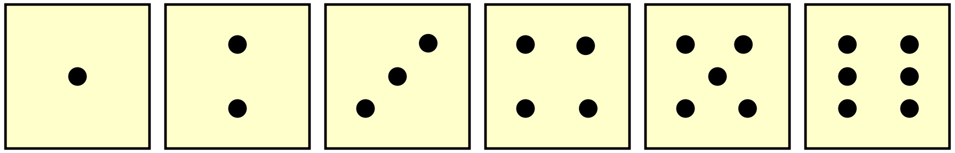
$A \cap B = \{silver\}$; $A \cap C = \{gold\}$

A^C={bronze, no medal}; B^C={gold}=C

Sample Space Examples

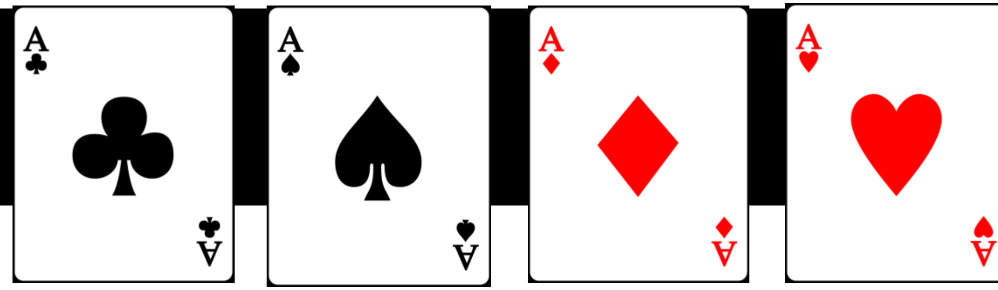
The Sample Space is the collection of all possible outcomes

Example 1: All 6 faces of a die



Example 2: All 52 cards of a deck of cards

Visualizing Events



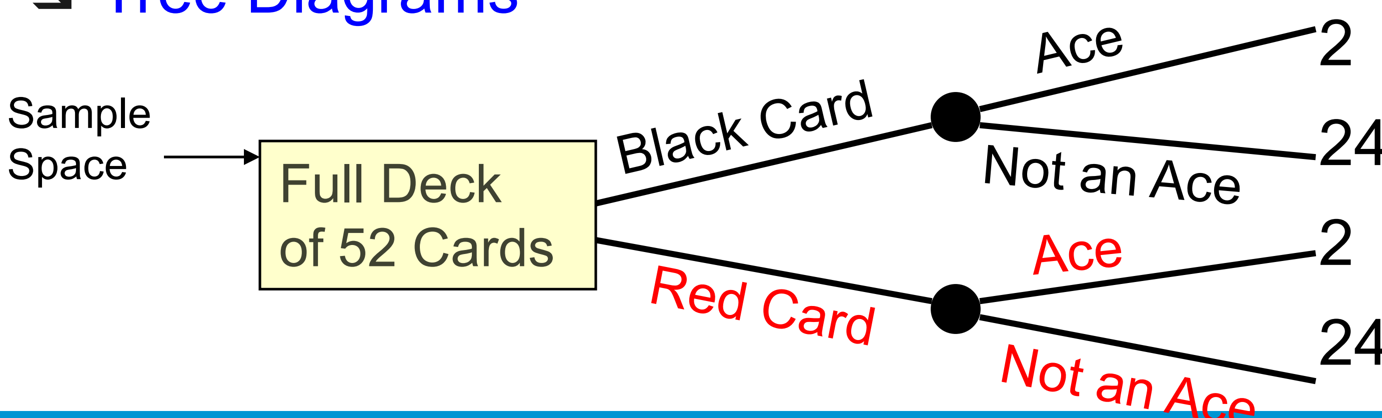
Club Spade Diamond Heart

- Contingency Tables

	Ace	Not Ace	Total
Black	2	24	26
Red	2	24	26
Total	4	48	52

Sample Space

- Tree Diagrams



Example (Cont'd)

A snowboarder athlete attempts to predict her chances of earning a medal in the women's halfpipe.

Outcome	Probability
Gold	0.1
Silver	0.15
Bronze	0.2
No Medal	0.55

Event A: 'At least earning a silver medal'

Event B: 'At most earning a silver medal'

Event C: 'Winner of the competition'

$P(A)=P(\{Gold\})+P(\{Silver\})=0.1+0.15=0.25$

$P(B)=P(\{Silver\})+P(\{Bronze\})+P(\{No\ Medal\})$

$= 0.15+0.2+0.55=0.9$

$P(C)=P(\{Gold\})=0.1=1-P(B)$

$P(A\cup B)=P(A)+P(B)=0.25+0.9$ *???? What's wrong?*

Probability – Characteristics

Probabilities must lie between 0 and 1

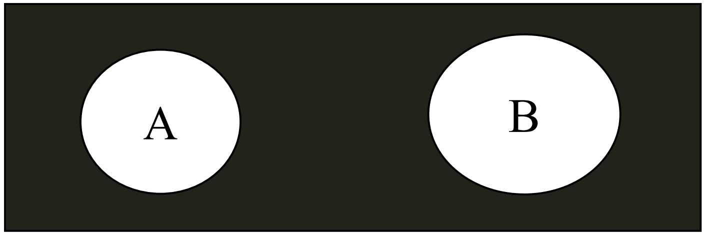
Probability = a number between 0 and 1 that measures the likelihood that some event will occur (e.g., P(A)).

Probabilities must add up

If two outcomes for a decision problem are mutually exclusive (at most one of them can occur) and also collectively exhaustive, then the probability that either outcome occurs =1, the sum of the individual probabilities (one or the other occurs).

P (A or B) = P(A) + P(B)=1

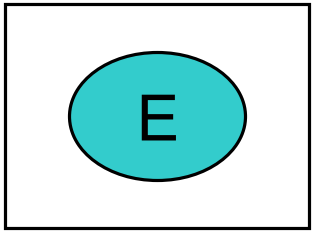
$\Sigma p_i=1$



Rules of Probability

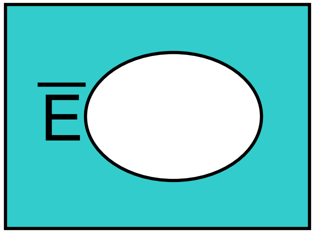
Complement Rule (simplest)

- The complement of an event E is the collection of all possible elementary events **NOT** contained in event E. The complement of event E is represented by $\overline{E}$ or E^C.



- Complement Rule:

e.g., If there is a 40% chance it will rain, there is a 60% chance it won't rain!



$$P(\overline{E}) = 1 - P(E)$$

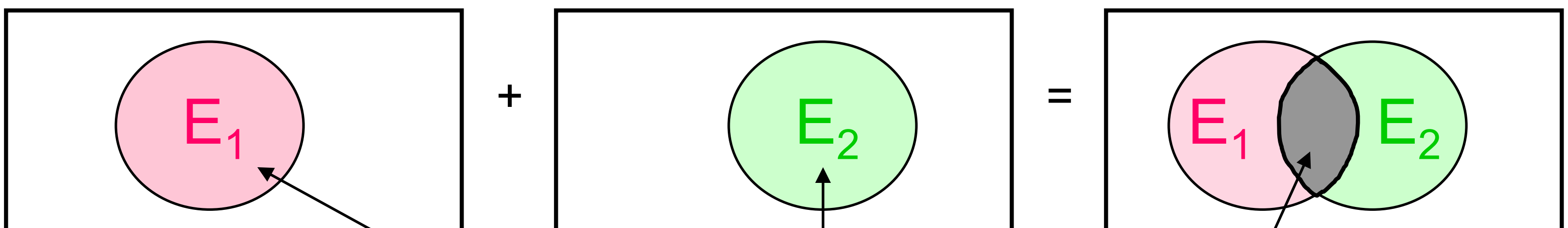
or

$$P(E) + P(\overline{E}) = 1$$

Addition Rule for Two Events

Addition Rule (when events are NOT mutually exclusive):

$$P(E_1 \text{ or } E_2) = P(E_1) + P(E_2) - P(E_1 \text{ and } E_2)$$



$$P(E_1 \text{ or } E_2) = P(E_1) + P(E_2) - P(E_1 \text{ and } E_2)$$

Don't count common elements twice!

***NOTE:* "or" indicates addition of events**

***joint* probability**

Example

Anthony feels that he has a 75% chance of getting an A in statistics and a 55% chance of getting an A in management economics. He also believes he has a 40% chance of getting an A in both courses.

1. What is the probability that he gets an A in at least one of these courses?
2. What is the probability that he does not get an A in either of these courses?

$P(A_S \cup A_M) = P(A_S) + P(A_M) - P(A_S \cap A_M)$=0.75+0.55-0.4=0.9

$P\left((A_S \cup A_M)^C\right) = 1 - P(A_S \cup A_M)$ =0.1

Conditional Probability

P(A|B)

- Probabilities are always assessed relative to the information currently available. As new information becomes available, probabilities often change.
- A formal way to revise probabilities on the basis of new information is to use *conditional probabilities*.
- Let A and B be any events with probabilities $P(A)$ and $P(B)$. Typically the probability $P(A)$ is assessed without knowledge of whether B does or does not occur. However if we are told B has occurred, the probability of A might change.

Conditional Probability Example

Of the cars on a used car lot, 70% have air conditioning (AC) and 40% have a CD player (CD), and 20% of the cars have both.

- What is the probability that a car has a CD player, **given that** it has AC ?

i.e., we want to find $P(CD \mid AC)$

Conditional Probabilities

	CD	No CD	Total
AC	0.2	0.5	0.7
No AC	0.2	0.1	0.3
Total	0.4	0.6	1.0

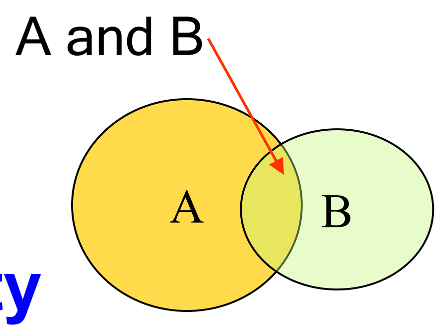
Of the cars on a used car lot, 70% have air conditioning (AC) and 40% have a CD player (CD). **20%** of the cars have both.

$$P(CD \mid AC) = \frac{P(CD \text{ and } AC)}{P(AC)} = \frac{0.2}{0.7} = 0.2857$$

Probability Characteristics – Cont’d

Joint probability (Compound events)

If two different events A and B occur, the probability is a joint probability and can be calculated as a product.



Conditional probability

Probabilities are always assessed relative to info currently available.

$$P(\mathrm{A} \mid \mathrm{B}) = \frac{P(A \text{ and } \mathrm{B})}{P(B)}$$

Multiplication Rule

- In the conditional probability rule the numerator is the probability that both A and B occur. It must be known in order to determine $P(A|B)$.
- However, in some applications $P(A|B)$ and $P(B)$ are known; in these cases we can multiply both side of the conditional probability formula by $P(B)$ to obtain the **multiplication rule**.

$$P(A \text{ and } B) = P(A|B)P(B)$$

- The conditional probability formula and the multiplication rule are both valid; in fact, they are equivalent.

Summing Up - Conditional Probabilities

$$P(\mathrm{A} \mid \mathrm{B}) = \frac{P(A \text{ and B})}{P(B)}$$

So - joint probability P(A and B) is

$$P(A \text{ and B}) = P(\mathrm{A} \mid \mathrm{B})P(B) = P(\mathrm{B} \mid \mathrm{A})P(A)$$

Independent and Dependent Events

- Two events are *independent* if the occurrence of one event does not affect the probability of the occurrence of the other event.
- Two events are independent *if and only if*

$$P(A \mid B) = P(A) \quad \underline{or} \quad P(B \mid A) = P(B)$$

- The main advantage of knowing that two events are independent is that the multiplication rule simplifies to

$$P(A \text{ and } B) = P(A)P(B)$$

Example

A stockbroker knows from the past experience that the probability that a client owns share is 0.6 and the probability that a client owns bond is 0.5. The probability that the client owns bonds if she already owns share is 0.55

1. What is the probability that the client owns both of these securities?
2. Given that the client owns bonds, what is the probability that the client owns share?

$P(A) = 0.6,\ P(B) = 0.5,\ P(B|A)$ =0.55

1. $P(A \cap B) = P(B|A) \times P(A)$=$0.55 \times 0.6 = 0.33$
2. $P(A|B) = P(A \cap B)/P(B) = 0.33/0.5 = 0.66$

Example

Samantha, a UWA student in year 3, contemplates her future immediately after graduation. She thinks there is a 25% chance that she will get a job at E&Y and a 35% chance that she will enroll in a full-time honours program in Perth. These are mutually exclusive outcomes.

$$P(A \cup B) = P(A) + P(B) = 0.25 + 0.35 = 0.60$$

- What is the probability that she does not choose either of these options?

$$P\left((A \cup B)^c\right) = 1 - P(A \cup B) = 1 - 0.60 = 0.40$$

Break

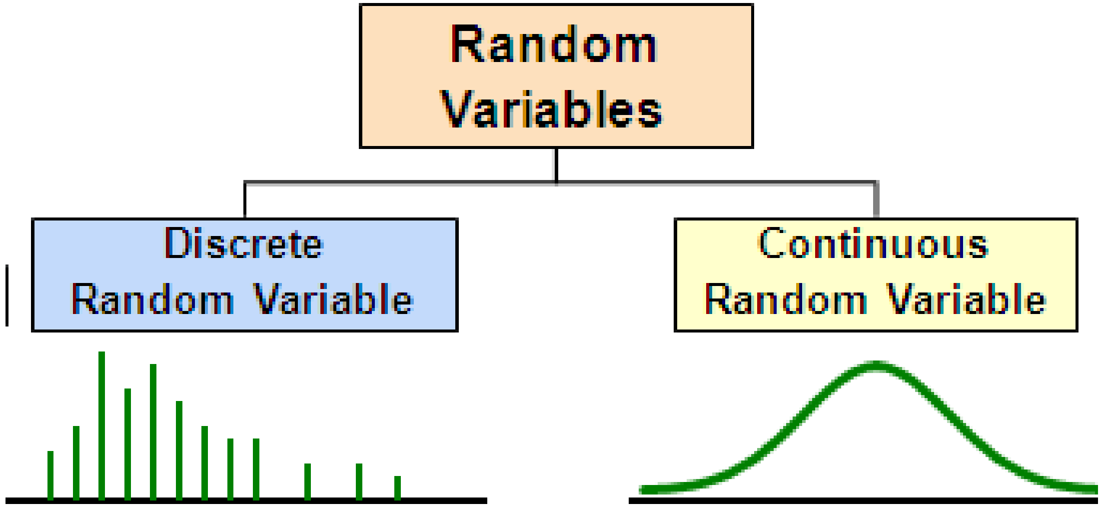
Theoretical Distributions

- Random variable
- Two types:
  - Discrete (has only a finite number of possible values)

Example: The number of defective light bulbs in a sample of five;

  - Continuous (has a continuum of possible values)

Example: The flight time between Perth and Sydney



Terminology

- Every random variable is associated with a **probability distribution** that describes the variable completely.
  - A **probability mass function** is used to describe discrete random variables.

P(X=x) x refers to a possible outcome

  - A **probability density function** is used to describe continuous random variables.

P(a<X<b) a and b refer to values of a specific interval

  - A **cumulative distribution function** may be used to describe both discrete and continuous random variables

$P(X \leq x)$ x refers a possible outcome/value in discrete/continuous random variable variable

Discrete Random Variable

- Expected value and the standard deviation

$$E(x) = \sum xP(x)$$

- Example: Toss 2 coins,
  x = # of heads,
  compute expected value of x:

x	P(x)
0	0.25
1	0.50
2	0.25

E(x) = (0 x 0.25) + (1 x 0.50) + (2 x 0.25)
= 1.0

Expected value may not equal to most probable value!

- Standard Deviation of a discrete distribution

$$\sigma_x = \sqrt{\sum \{x - E(x)\}^2 P(x)}$$

where:
E(x) = Expected value of the random variable
x = Values of the random variable
P(x) = Probability of the random variable having the value of x

Standard Deviation for Tossing 2 Coins

- Recall that the expectation E(x) = 1

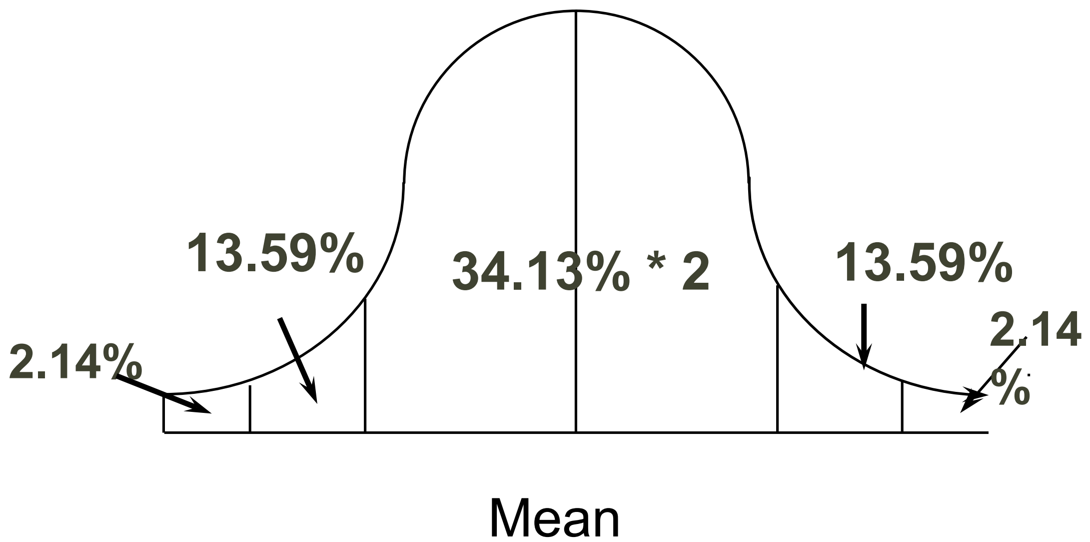
$$\boxed{\sigma_x = \sqrt{\sum \{x - E(x)\}^2 P(x)}}$$

$$\sigma_x = \sqrt{(0-1)^2(0.25) + (1-1)^2(0.50) + (2-1)^2(0.25)} = \sqrt{0.50} = .707$$

Possible number of heads
= 0, 1, or 2

Continuous distribution-Normal Distribution

- Gauss curve
- Bell shaped, symmetric, asymptotic
- Almost all of its values are within plus or minus 3 standard deviations
- I.Q., height = examples



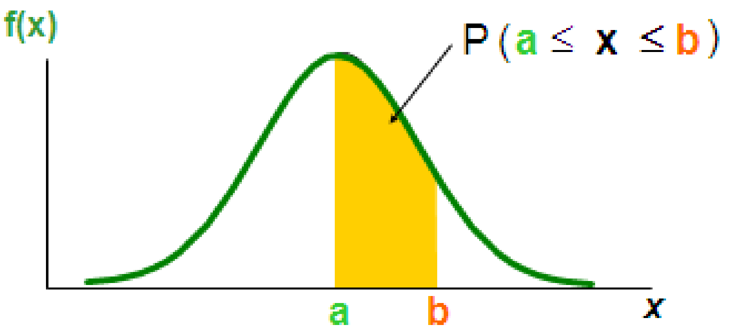
Standardised Values

- Used to compare an individual value to the population mean in units of the standard deviation

Standardize (value, mean, stdev)

$$z = \frac{x - \mu}{\sigma}$$

Probability is measured by the area under the curve



- Necessary to compare distributions for the NORMDIST (x,μ,σ,1/0) or NORMSDIST
- N_{DADM}=(65,20); N_{OB}(75,5)

- Z_{DADM}=(85-65)/20=1
- Z_{OB} =(85-75)/5 = 2

Examples N Distribution – Excel Functions

- The weekly rent for apartments and houses in Nedlands is N distributed with a mean of \$425 and a variance of \$90^2.
- What is the chance to find a house under \$320 per week?
  - P (x < 320) = ?
  - Z=STANDARDIZE(320,425,90)=-1.667
  - P (x < 320) =P (Z<-1.667) = NORMDIST(320,425,90,TRUE) = NORMSDIST(-1.667) = 0.1217
- What is the chance that you find a house with rent between \$450 and \$550/week?
  - =NORMSDIST(1.389)- NORMSDIST(0.278)=0.308

Statistics

- ✓ Statistics: Data, analysis, and interpretation
- ✓ Descriptive statistics vs Inferential statistics
- ✓ Cross-sectional and time-series/longitudinal

Recap

What Does Statistics Mean?

↘ *Descriptive statistics*

Collecting, presenting and describing data

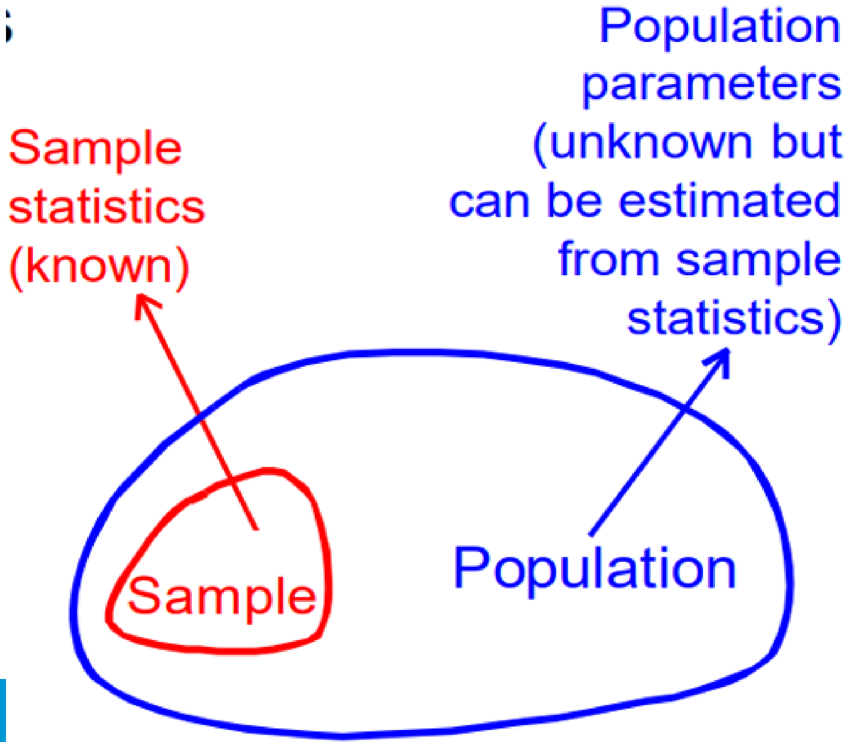
- E.g., number of people, trends in employment
- Numerical and graphical tools

Recap

- *Inferential statistics*

Make an inference (draw conclusions and/or make decisions) about a population from a sample

- Population Parameter vs Sample Statistics
- Estimation
- Hypothesis testing



Recap

Sampling Method

Probability Sampling

- Simple Random sampling
- Stratified Random sampling
- Cluster sampling

Non-probability Sampling

- Snowball sampling
- Convenience sample

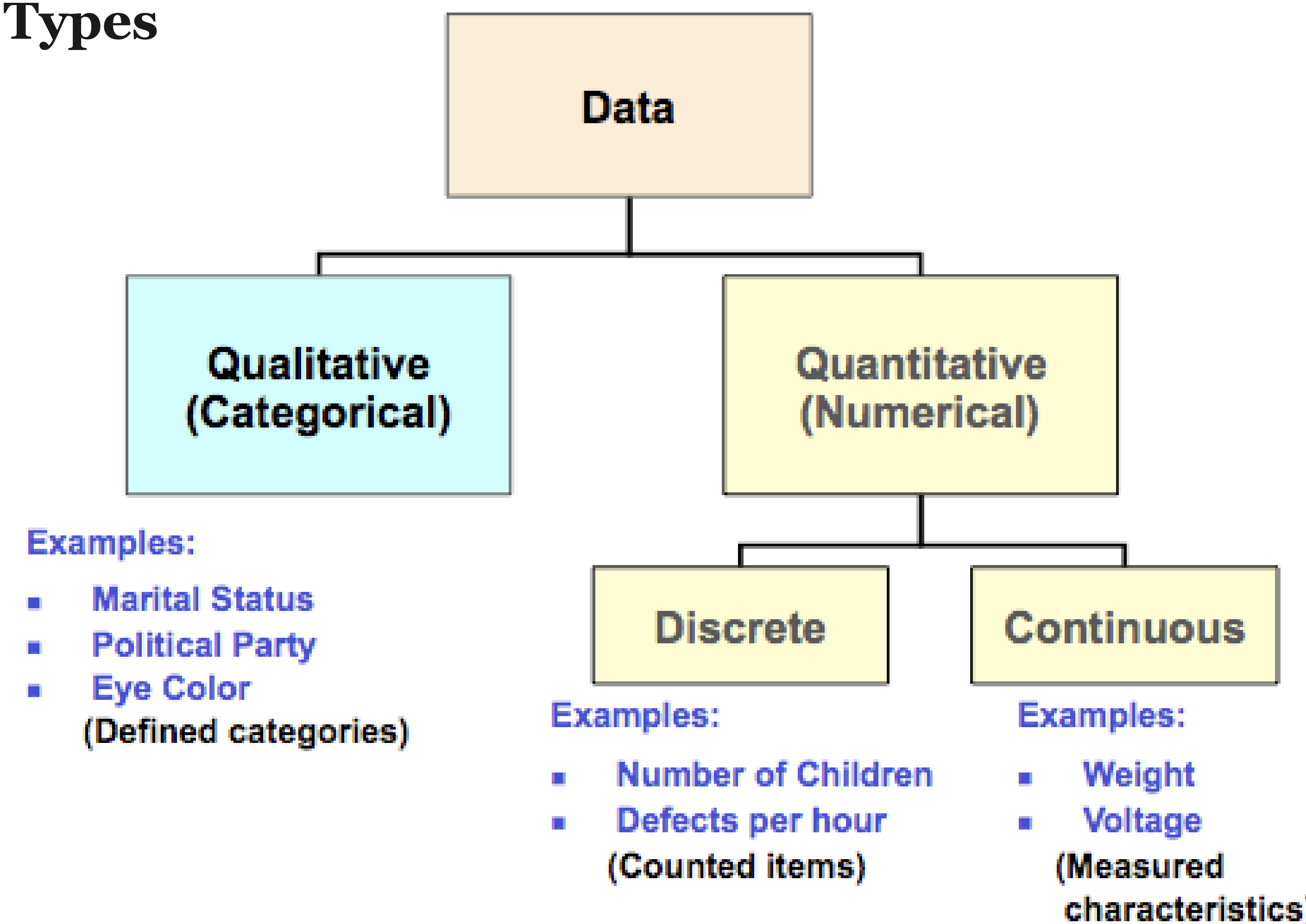
Recap

Types of Data/Variables

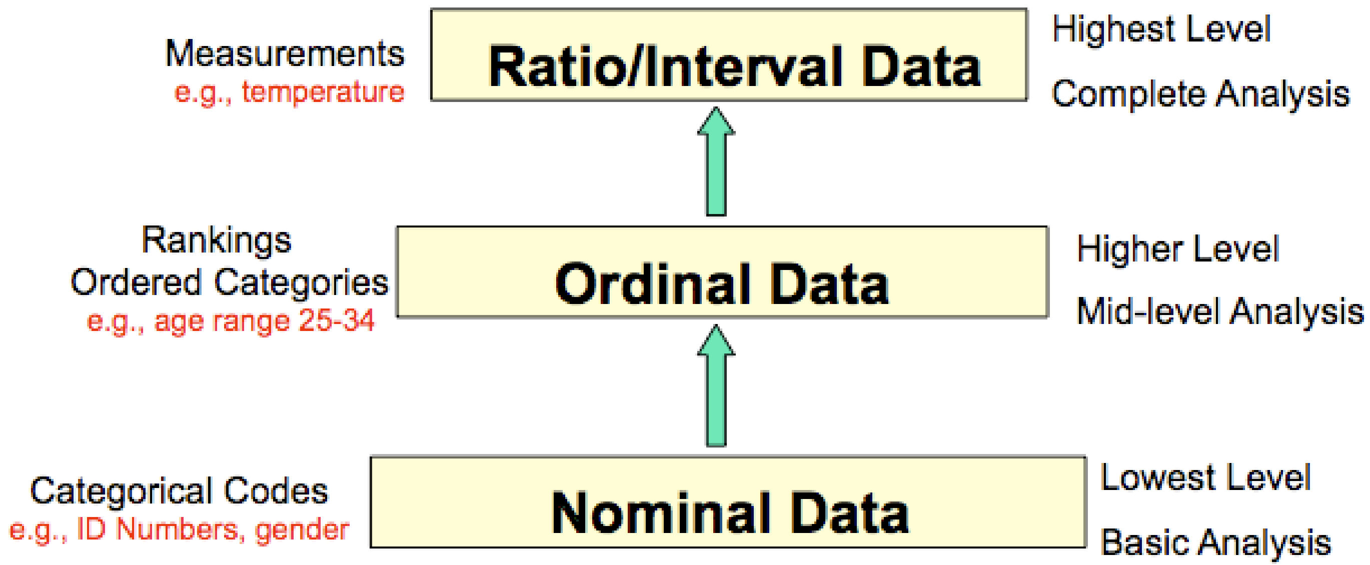
- Cross-sectional and time-series/longitudinal

- Categorical – nominal, ordinal
- Continuous/Numerical – interval, ratio

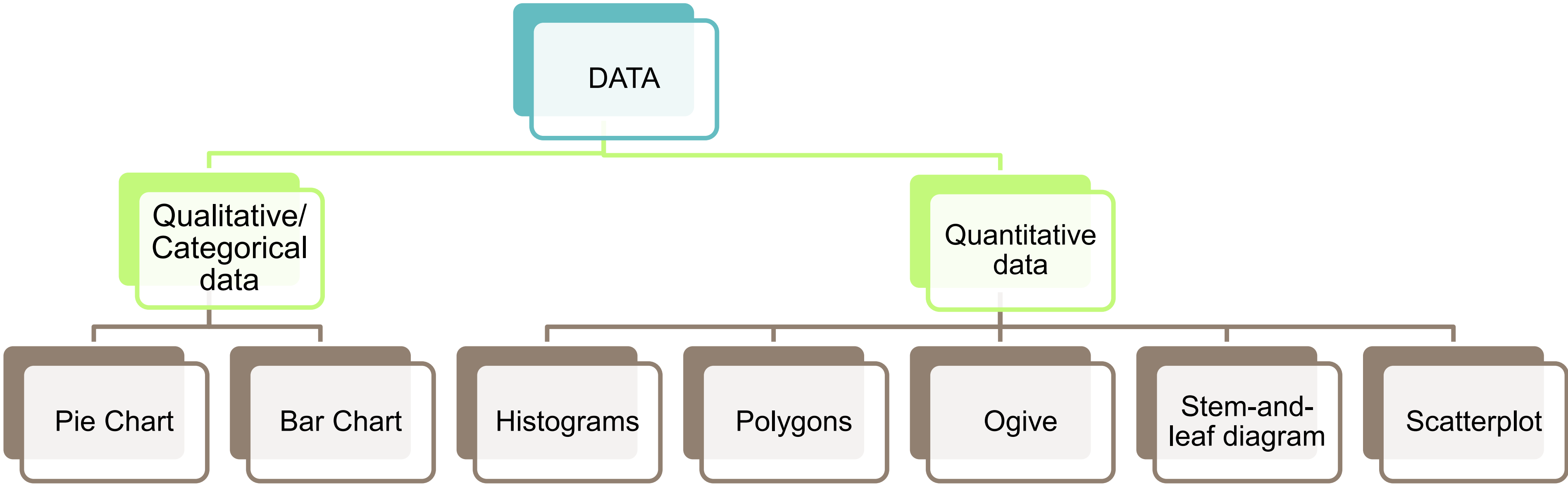
Data Types


Data
Qualitative (Categorical)
Quantitative (Numerical)
Discrete
Continuous
Examples:
Marital Status
Political Party
Eye Color
(Defined categories)
Examples:
Number of Children
Defects per hour
(Counted items)
Examples:
Weight
Voltage
(Measured characteristics)

Understand different types of measurement scales



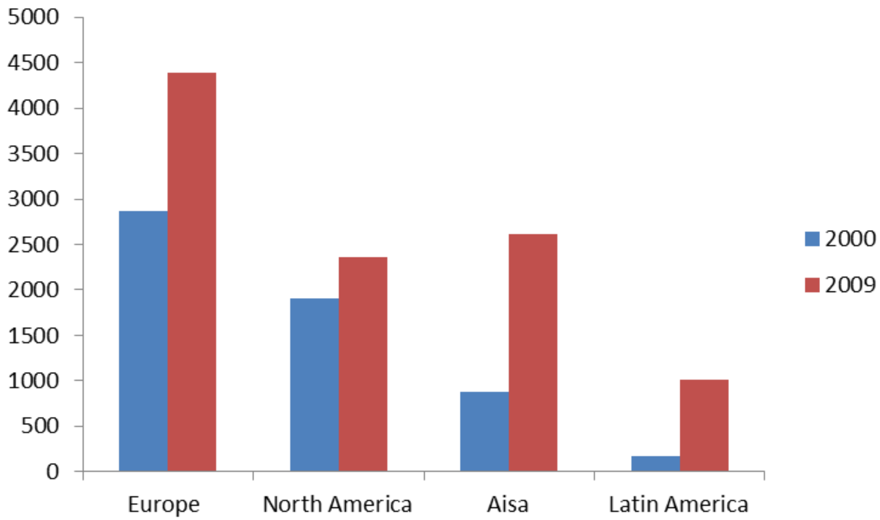
Descriptive analysis

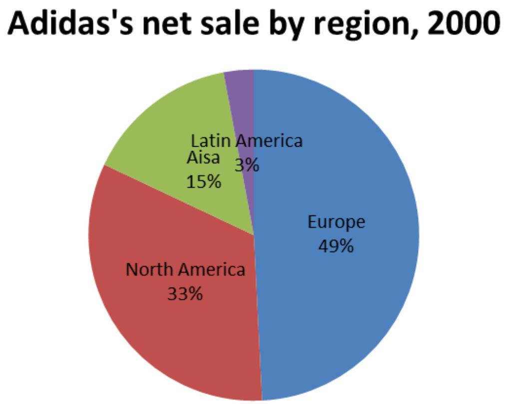


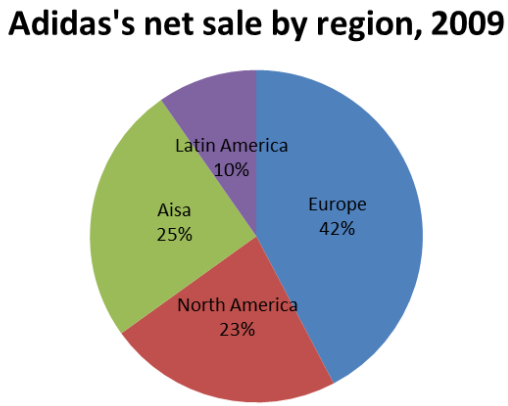
Example : Categorical variable

Adidas's net sales by region (millions of euros)

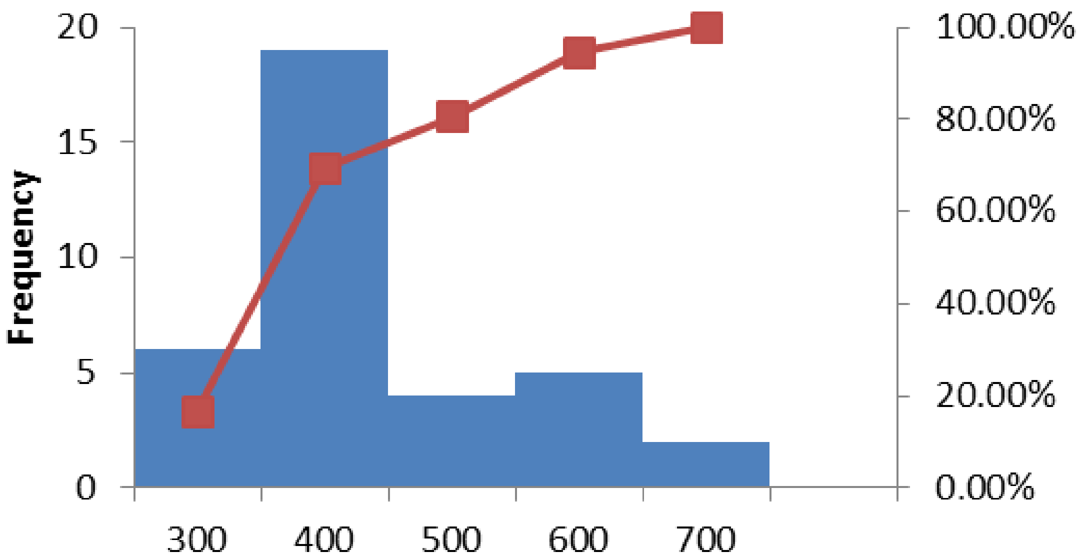
Region	2000	2009
Europe	2860	4384
North America	1906	2360
Aisa	875	2614
Latin America	171	1006
Total	5812	10364

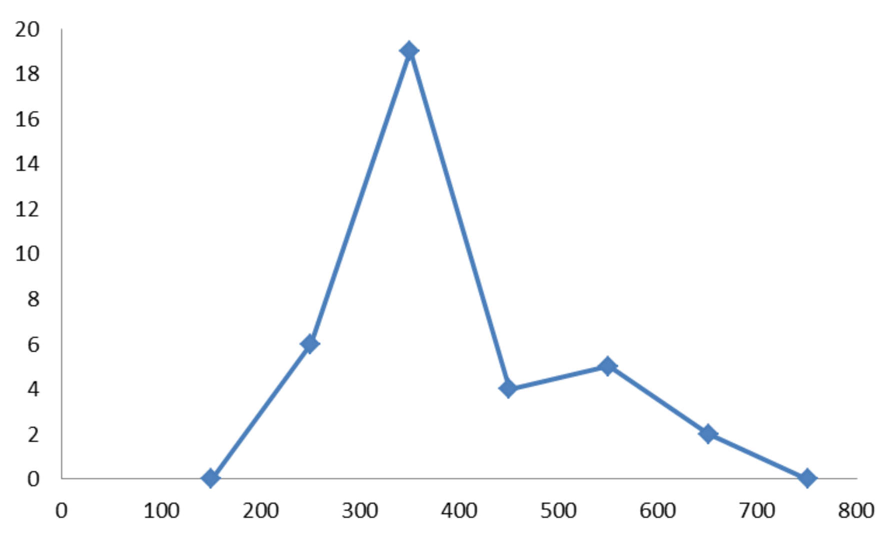






Example: Quantitate variable





House sales prices in Warrnambool, 2014

349	435	525	315
349	399.9	229	335
299	331.1	331.1	480
330	299.9	339	239
289.9	375	399.9	330
629	315	695	209
595	355	355	519
339.9	385	329	520
595	449.9	499.9	399.9

Stem and Leaf Diagram

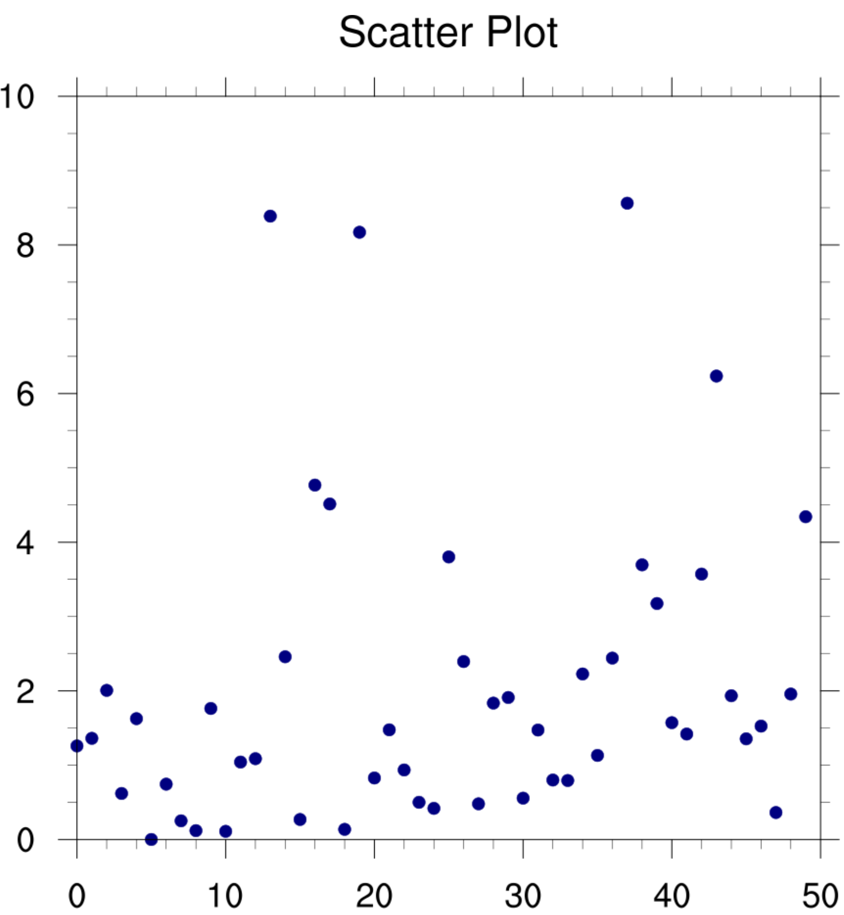
- A simple way to see distribution details for quantitative data

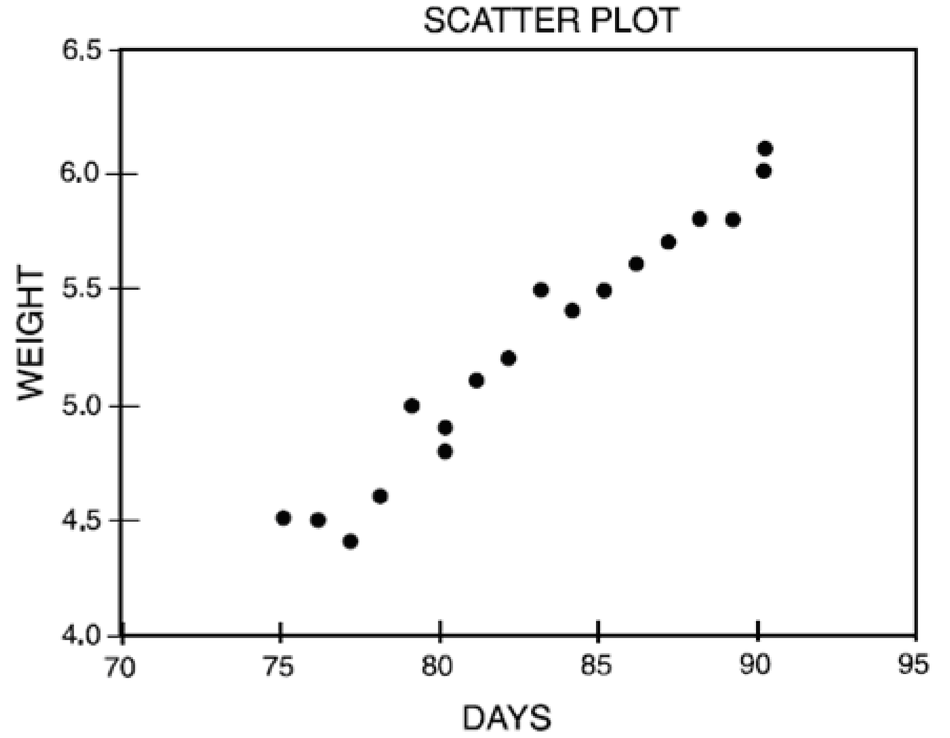
- Separate the sorted data series into leading digits (stem) and trailing digits (leaves)
- List all stems from low to high
- For each stem, list all associated leaves

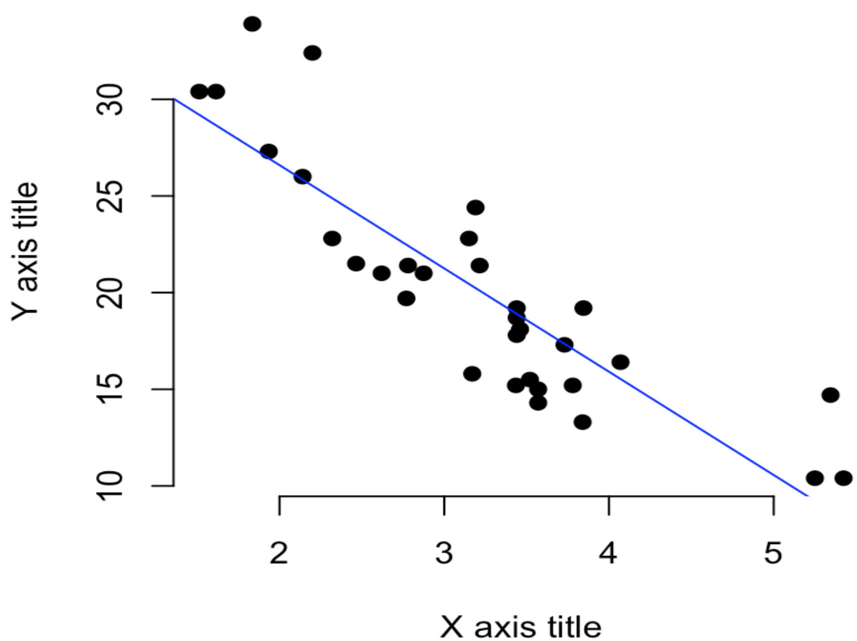
- Example: 12, 13, 17, 21, 24, 26, 27, 28, 30, 32, 35, 37, 38, 41, 43, 44, 46, 53, 58

Stem	Leaves
1	2 3 7
2	1 4 4 6 7 8
3	0 2 5 7 8
4	1 3 4 6
5	3 8

THE UNIVERSITY OF WESTERN AUSTRALIA





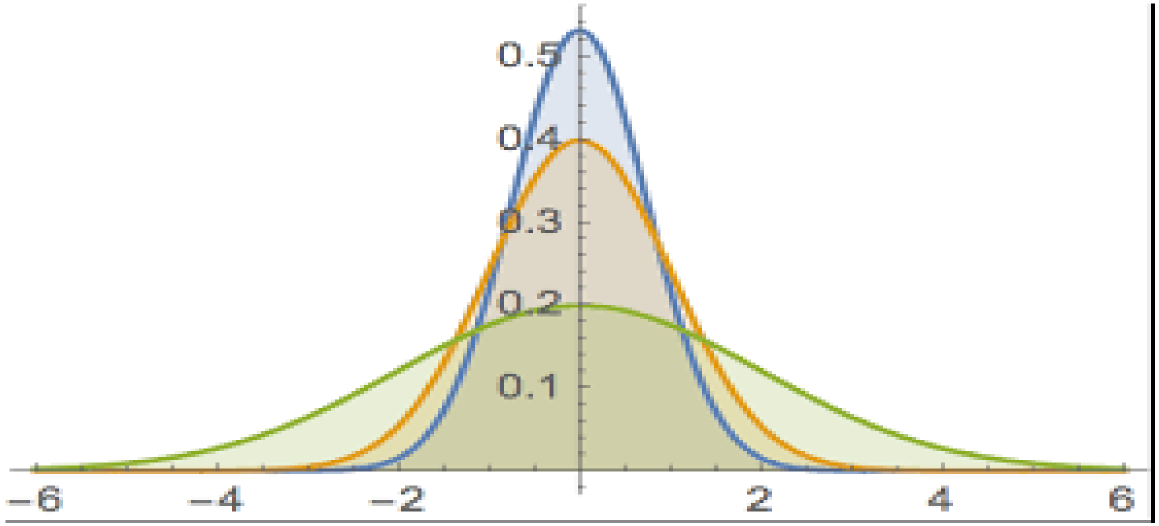


Descriptive Statistics

- Central tendency
  - Mean
  - Median
  - Mode
  - Percentile
  - Quartile

- Measures of dispersion
  - Range
  - Variance and standard deviation
  - Coefficient of variation

$$=\frac{standard\ deviation}{mean}$$



Calculation for central location

- Mean$=\frac{\sum x_i}{N}$

$=\frac{40000+40000+65000+90000+145000+150000+550000}{7}$=\$154,286

- Mode = 40,000
- Percentile (p^{th} percentile)

L_p=(n+1)$\frac{p}{100}$

- Median

L_{50}=(7+1)$\frac{50}{100}$=4 , thus Median=\$90,000

- Quartile (25th percentile, 50th percentile, 75th percentile)

L_{25}=(7+1)$\frac{25}{100}$=2 , thus 1st Quartile=\$40,000

L_{50}=(7+1)$\frac{50}{100}$=4 , thus 2nd Quartile=\$90,000

L_{75}=(7+1)$\frac{75}{100}$=6 , thus 3rd Quartile=\$150,000

Title	Salary (\$)
Administrative assistant	40,000
Research assistant	40,000
Computer programmer	65,000
Senior research associate	90,000
Senior sales associate	145,000
Chief financial officer	150,000
President (and owner)	550,000

Question: Calculate and interpret the 60th percentile salary of employees. $L_{60} = 4.8$, thus 60% of employees' salary is below \$134,000

Box-Whisker plot

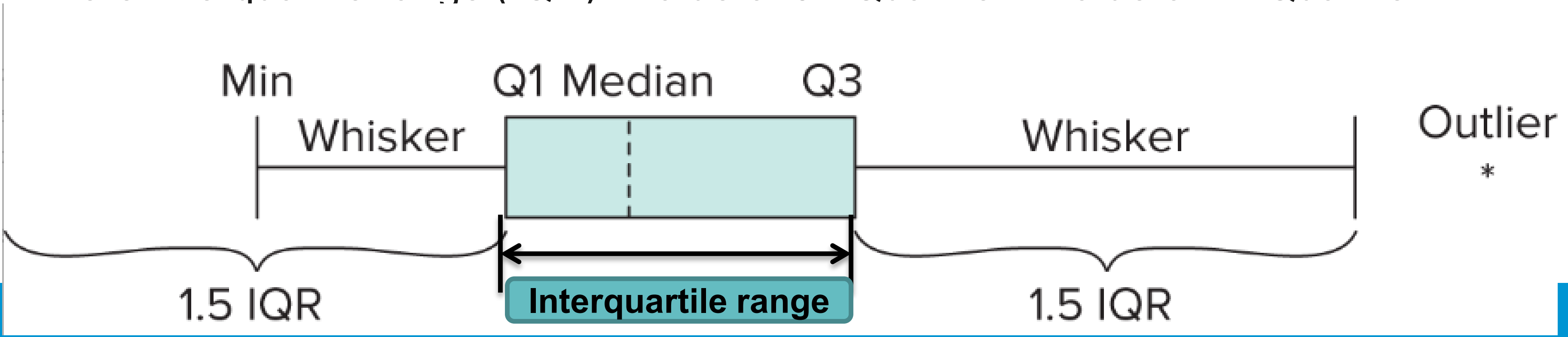
- Minimum
- 1st Quartile
- Median (2nd Quartile)
- 3rd Quartile
- Maximum

To determine the outlier

Smaller than 1.5×Interquartile Range below the 1st Quartile, OR

Larger than 1.5×Interquartile Range above the 3rd Quartile

Where Interquartile range (IQR)= Value of 3rd Quartile – Value of 1st Quartile



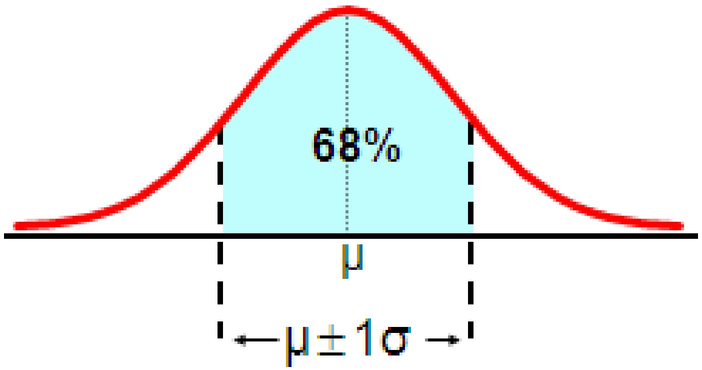
Empirical Rule

Symmetric and bell-shaped distribution:

68% of observations fall in the interval $\mu \pm \sigma$

95% of observations fall in the interval $\mu \pm 2\sigma$

99.7% of observations fall in the interval $\mu \pm 3\sigma$



Example: 280 students in one lecture class with a mean sore 74 and a standard deviation of 8 (assuming distribution is symmetric and bell-shaped).

1. Approximately how many students scored between 58 and 90?
2. Approximately how many students scored more than 90?

Probability Essentials

- Probability – the chance that an uncertain event will occur (always between 0 and 1)
- Experiment – a process that produces outcomes for uncertain events
- Sample Space (or event) – the collection of all possible experimental outcomes
- Concept of probability is quite intuitive; however, the rules of probability are not always intuitive or easy to master.
  - An event with probability zero cannot occur.
  - An event with probability 1 is certain to occur.
  - An event with probability greater than 0 and less than 1 involves uncertainty, but the closer its probability is to 1 the more likely it is to occur.

Interpretations of Probability

Objective and Subjective

- **Classical** - Propensity – i.e. the physical design of an object (die)
- **Empirical** - Relative frequency in the long run (mortality tables for calculating insurance premiums)
- **Subjective** - Reasonable degree of subjective belief (reliability of witness testimony)

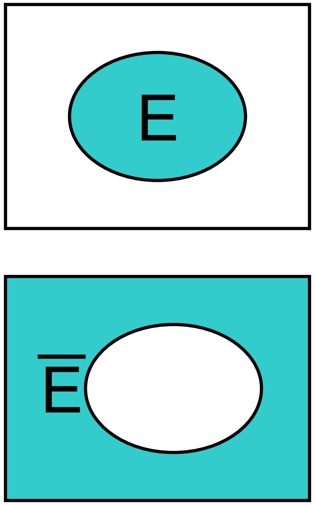
Rules of Probability

Complement Rule (simplest)

- The complement of an event E is the collection of all possible elementary events **NOT** contained in event E ($\overline{E}$ or E^C).

- Complement Rule:

e.g., If there is a 40% chance it will rain, there is a 60% chance it won't rain!



$$P(\overline{E}) = 1 - P(E)$$

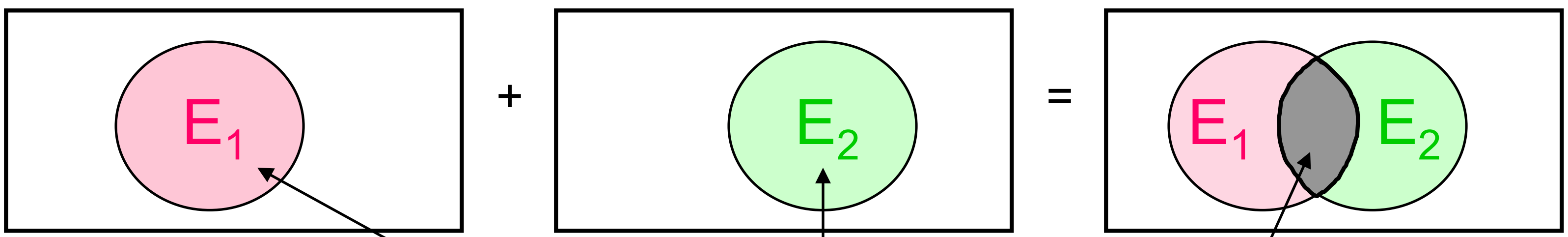
or

$$P(E) + P(\overline{E}) = 1$$

Addition Rule for Two Events

Addition Rule (when events are NOT mutually exclusive):

$$P(E_1 \text{ or } E_2) = P(E_1) + P(E_2) - P(E_1 \text{ and } E_2)$$



$$P(E_1 \text{ or } E_2) = P(E_1) + P(E_2) - P(E_1 \text{ and } E_2)$$

Don’t count common elements twice!

NOTE: **“OR” indicates addition of events**

***joint* probability “AND”**

Conditional Probability

$$P(\mathrm{A}\,|\,\mathrm{B}) = \frac{P(A \text{ and } \mathrm{B})}{P(B)}$$

- Probabilities are always assessed relative to the information currently available. As new information becomes available, probabilities often change.
- A formal way to revise probabilities on the basis of new information is to use *conditional probabilities*.
- Let *A* and *B* be any events with probabilities *P*(*A*) and *P*(*B*). Typically the probability *P*(*A*) is assessed without knowledge of whether *B* does or does not occur. However if we are told B has occurred, the probability of A might change.

Multiplication Rule

- In the conditional probability rule the numerator is the probability that both A and B occur.
- However, in some applications $P(A|B)$ and $P(B)$ are known; in these cases we can multiply both side of the conditional probability formula by $P(B)$ to obtain the **multiplication rule**.

$$P(A \text{ and } B) = P(A|B)P(B)$$

- The conditional probability formula and the multiplication rule are both valid; in fact, they are equivalent.

Example of quiz

1. An analyst collects data on the weekly closing price of gold throughout a year. The scale of data measurement is ________________.

$\boxed{\text{A. ratio scale}}$

B. ordinal scale

C. interval scale

D. nominal scale

A

Example of quiz

2. In the accompanying stem-and-leaf diagram, the values in the stem and leaf portions represent 10s and 1s digits, respectively.

Stem	Leaf
1	3 5 6 8 8 9
2	0 1 2 2 3 5 6 6 8 8 8 9
3	0 1 2 2 8
4	2 2

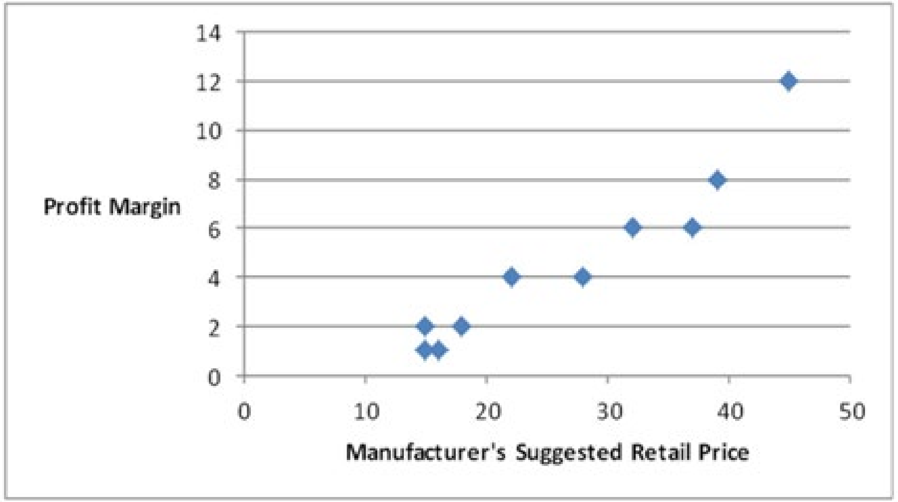
What is the frequency of the class 35 up to 45, that is {x; 35 ≤x< 45}?

A. 0
B. 1
C. 2
D. 3

D

Example of quiz

3. A car dealership created a scatterplot showing the manufacturer's suggested retail price (MSRP) and profit margin for the cars on its lot.



As the MSRP increases, the profit margin tends to:

A. increase

B. decrease

C. stay the same

D. None of the answers

A

Example of quiz

4. Consider a population with data values of
   12, 8, 28, 22, 12, 30, 14

The mean and median are:

A. Mean=12, Median=14

B. Mean=18, Median=14.

C. Mean=18, Median=12

D. Mean=22, Median=12

B

Example of quiz

5. An analyst gathered the following information about the net profit margins of companies in two industries:

Net Profit Margin	Industry A	Industry B
Mean	15.0%	5.0%
Standard deviation	2.0%	0.8%
Range	10.0%	15.0%

Compared with the other industry, the relative dispersion of net profit margins is smaller for Industry:

A. B because it has a smaller mean deviation.

B. B because it has a smaller range of variation.

C. A because it has a smaller standard deviation.

D. A because it has a smaller coefficient of variation.

D

Example of quiz

6. Two hundred people were asked if they had read a book in the past month. The accompanying contingency table, cross-classified by age, was produced.

	Under 30	30 +
Yes	76	65
No	24	35

Given a respondent read a book in the past month, the probability he or she was at least 30 years old is *closest* to __________.

A. 0.33

B. 0.46

C. 0.65

D. 0.88

B

Example of quiz

7. The following discrete probability distribution represents the anticipated profit/loss ($1,000s) for a small business in the upcoming financial year.

x	-10	0	10	20
$P(X=x)$	0.35	0.10	0.15	0.40

What is the probability the small business makes a profit?

A. 0.10

B. 0.35

C. 0.55

D. 0.65

C

Examples:

8. A hedge fund returns on average 26% per year with a standard deviation of 12%. Using the empirical rule, approximate the probability the fund returns over 50% next year.

A. 0.5%

B. 1%

C. 2.5%

D. 5%

C