Statistics by Hand

An Introductory Course for Psychologists

Introduction

Version 3.0

Sally Clark

- Christopher (11 weeks) died in his sleep.
- One year later, Harry(8 weeks) also died.
- Forensic evidence was “slim to non-existent”
- No evidence that Sally had been an uncaring or violent mother.

- Sally was convicted of murdering both children.

Roy Meadows – expert witness

- Probability of *two* cot deaths in the same family are

1 in 73 million

Less than once a century in the UK

Understanding probability

- A *probability* is a number between 0 and 1 that expresses the likelihood of an event.

- P = 1 event is sure to occur
- P = 0 event will never occur

- P(Heads) = 0.5
- P(Tails) = 0.5

Multiplicative Law

The probability of a set of independent events all happening is the product of the individual probabilities.

P(H)=0.5 P(H)=0.5

0.5 x 0.5 = 0.25

Roy Meadows – expert witness

- Chances of a randomly chosen baby dying a cot death are 1 in 1,303 (p = 0.0008)
- If family is affluent, and mother over 26, then 1 in 8,500 (p = 0.0001)
- Probability of *two* cot deaths is

$0.0001 \times 0.0001 = 1.4 \times 10^{-8}$

1 in 73 million

Less than once a century in the UK

Logical flaw

The probability of a set of *independent* events all happening is the product of the individual probabilities.

Cot deaths in the same family are unlikely to be independent events (e.g. shared genes).

Siblings of a child who died of cot death are at least 10x more likely to die of cot death than the average.

Using the 1 in 1,303 figure:

$P = 0.0008 \times (0.0008 \times 10) = 6.4 \times 10^{-6}$ 1 in 150,000

About 4 cases a year in the UK!

A happier example

- What's the probability that two children in this class share a birthday?
- Assume birthdays are independent events.
- Assume 30 children.

P = 0
P = 0.003
P = 0.08
P = 0.25
P = 0.50
P = 0.75
P = 1

How many pairs in a group of 30?

$$C_r^N = \frac{N!}{r!(N-r)!}$$

The combination rule

- General rule for picking *r* things out of *N* things.
- *N = 30* (group size) , *r = 2* (pairs)

How many pairs in a group of 30?

- 4! = 4 x 3 x 2 x 1
- *N* = *30* (group size) , *r* = *2* (pairs)

$$C_r^N = \frac{N!}{r!(N-r)!} = \frac{30!}{2!28!} = \frac{30 \times 29 \times 28 \ldots \times 1}{2 \times 28 \times 27 \times 26 \ldots \times 1} = \frac{30 \times 29}{2} = 435$$

The answer explained

- The probability of a pair of children sharing a birthday is, let's say, 1/365 = 0.003
- The probability of a pair of children *not* sharing a birthday is therefore 1 – 0.003 = 0.997
- There are no shared birthdays *only if* no pairs of children share a birthday.
- Multiplicative law: $0.997^{435} = 0.28$
- Probability that there is at least one shared birthday =

1 – 0.28 = 0.72

Lies, damned lies, and statistics

“It has been said by some wits that there are three degrees of unveracity: lies, d---d lies, and statistics”.

- Mrs. Andrew Crosse (1892)

A modern example

"A recent survey showed that average income in Bristol is higher than in Exeter".

- Bristol Evening post

"A recent survey showed that average income in Exeter is higher than in Bristol".

- Express and Echo

It's the *same* survey, using the *same* sample!

Terminology

Sample: The group of people you are testing.

Population: The total set of people you could theoretically test.

A modern example

"A recent survey showed that average income in Bristol is higher than in Exeter".

- Bristol Evening post

"A recent survey showed that average income in Exeter is higher than in Bristol".

- Express and Echo

It's the *same* survey, using the *same* sample!

Mean and median

- Mean:

  $\bar{X} = \sum X / N$ $\quad\quad$ $X = [1, 2, 3, 4, 8]$

  $N = 5$ $\quad\quad$ $\bar{X} = 3.6$

- Median: The middle number when the data is placed in rank order

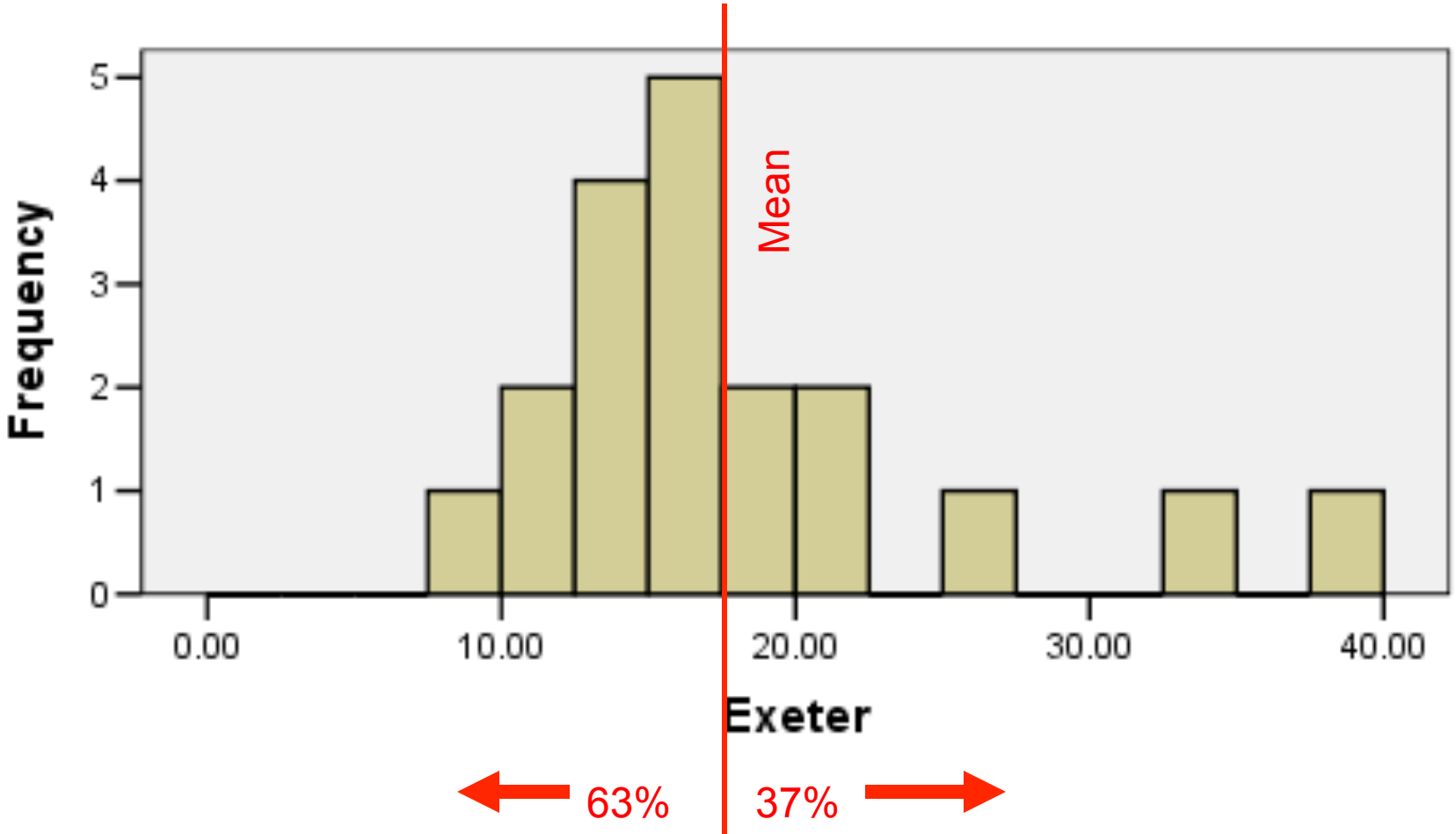
  1, 2, 3, 4, 8 $\quad\quad$ median = 3

Mean and median salaries (£k)

	Mean	Median
Bristol	16.0	16.0
Exeter	17.6	15.0

- Both headlines are supported by statistics.
- Relativism: Both *summary statistics* (mean and median) are equally true.
- Realism: One *summary statistic* is a better description of the truth than the other.

Exeter's distribution



Over to you…

- Attempt problems PCB-1 and PCB-2
- Remember:
  - The probability of a set of independent events all happening is the product of the individual probabilities.

$$C_r^N = \frac{N!}{r!(N-r)!}$$