Statistics by Hand
An Introductory Course for Psychologists

Normality

Version 3.0

Reminder

- Sample and population
- Variance
- New: Standard deviation
  - The square root of variance

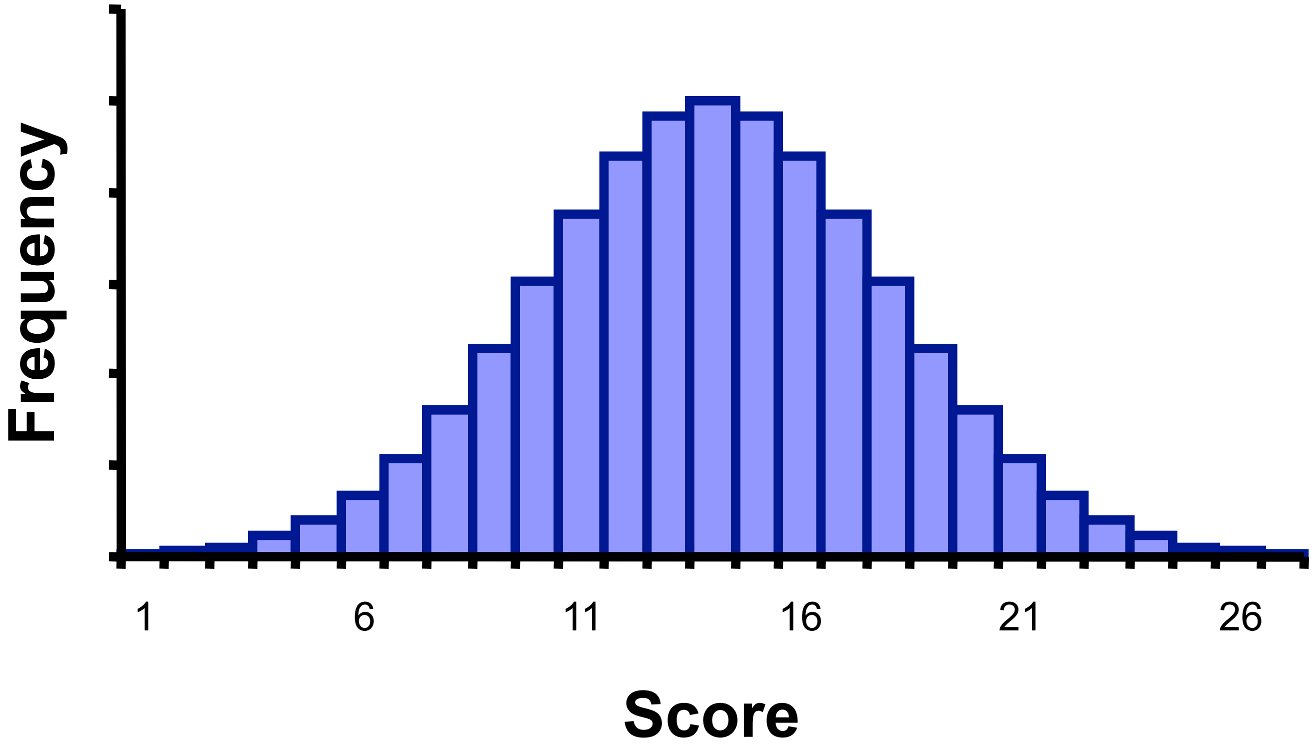
$$s^2 = \frac{\sum \left(X - \overline{X}\right)^2}{N-1}$$

Variance

$$s = \sqrt{\frac{\sum \left(X - \overline{X}\right)^2}{N-1}}$$

Standard deviation

“Normal” distribution

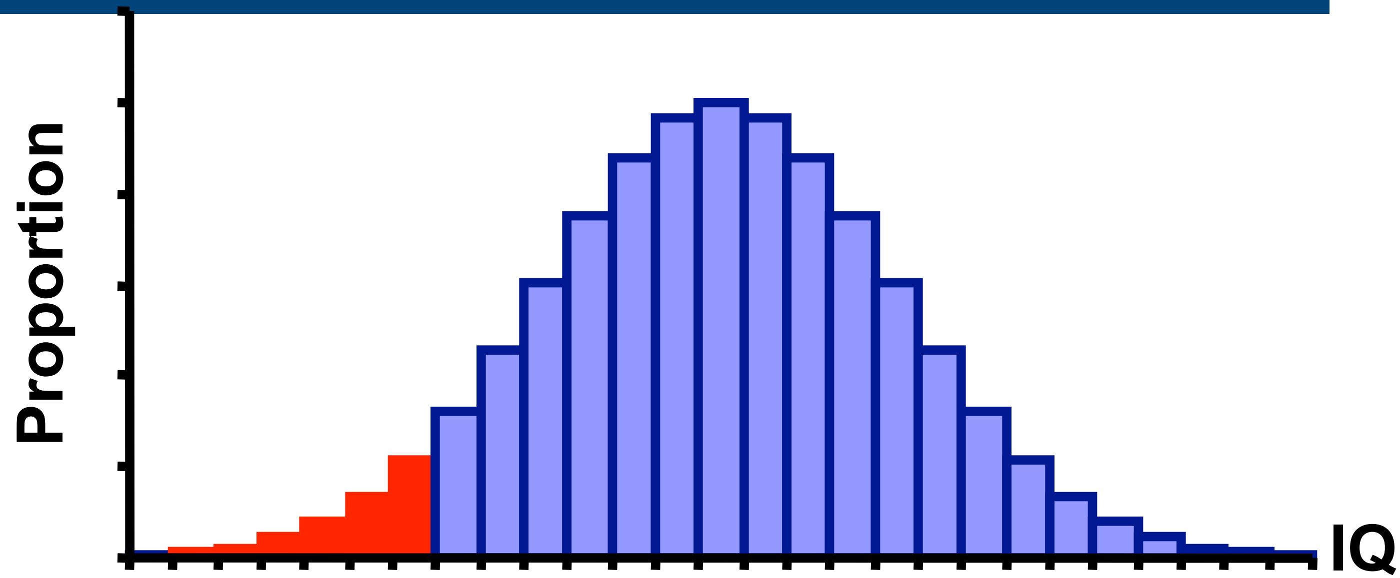


- Many types of data are *normally distributed.*
- Height, Astronomical observations, IQ

Normal distribution

- Example application
  - A patient incurs brain damage following a car crash.
  - His IQ on the WAIS is now 85.
  - He was not tested prior to the crash.
  - What's the likelihood that his IQ has been adversely affected by the crash?
- IQ scores (WAIS)
  - Normally distributed for population of non-brain damaged people.
  - $\mu = 100$, $\sigma = 15$

Example application



- We know the population is normal, mean 100, s.d. 15. For a mathematician, that's enough to draw this plot.
- From the plot, you can work out the proportion of non-brain damaged people who have a score of 85 or lower.

Z-tests

- That proportion is the probability that our patient comes from the non-brain-damaged population.
- All this represents an awful lot of work.
- Fortunately, there's a short-cut: Z-tests

Our patient

$$z = \frac{X - \mu}{\sigma} = \frac{85 - 100}{15} = -1$$

- P = 0.16
- Doesn't reach conventional levels of significance.

Z-test

- In psychology, use a Z-test where:
  - There is just one participant
  - That participant's data is just one number
  - The population the participant comes from is known to be **normally distributed**
  - The standard deviation of the population is known.
- All fairly rare, but if true then few other tests would work.

Normality of a sample

- A **population** either is, or is not, *normally distributed*.
- However, most of the tests covered in the rest of the course assume that the **sample** is <u>sufficiently</u> normal (i.e. more or less normal).
- How do you tell whether a sample is “sufficiently” normal?
- Here’s a rough-and-ready way to tell (the only way we’ll cover in this course).

Normality of a sample

Scores out of 60 (whole numbers) on a behavioural problems index

1,6,12,16,20,13,8,3,6,14,7,15,12,9,13

1-5	||
6-10	|||||
11-15	||||||
16-20	||

- Create roughly N/4 equal sized "bins"
- Make a mark for each number in the data set
- It'll never look great with small samples. This data set is roughly normal.

- Main things to look for: bimodality and asymmetry (skew)
- If N<10 then there's not really enough data to do this.

Bimodality

- Two peaks:

1-5	|||
6-10	|||||||||
11-15	|||
16-20	||||||||||

Skew

- Severe lack of symmetry:

<table>
<tr><td>1-5</td><td>||||||||||</td><td>1-5</td><td>|</td></tr>
<tr><td>6-10</td><td>||||||</td><td>6-10</td><td>|||||</td></tr>
<tr><td>11-15</td><td>|||</td><td>11-15</td><td>||||||</td></tr>
<tr><td>16-20</td><td>||</td><td>16-20</td><td>||||||||</td></tr>
</table>

Normality of distribution

Scores out of 60 on a behavioural problems index (ADHD children)
50, 49, 25, 27, 29, 45, 52, 51, 48, 26, 27, 30, 43, 51

20 |||||
30 |
40 ||||
50 ||||

- This data set does not look particularly normal: some evidence of bimodality and/ or skew.

Practice

- Problem Z-1
- Problem E-1