AARHUS UNIVERSITY

Methods 2 - 2

Chris Mathys

BSc Programme in Cognitive Science

Spring 2022

Data and measurement

- Data are the basis of modelling

- Aspects of data quality – discuss with your neighbour for 5 mins and come up with a list
- List:
  - **Validity**
    - Construct validity
    - Face validity
    - Ecological validity
  - **Reliability**
  - **Sample selection**
    - Selection bias – why a problem? What to do about it? Discuss for 5 mins and write down an answer

Graphs

- All graphs are comparisons
- Examples of graphs?
- Advantages, disadvantages?

Vectors and matrices (this and more: video 2a)

At the simplest level:

- *Vector:* list of numbers

$$\begin{pmatrix} 43.3 \\ 46.3 \\ 55.3 \end{pmatrix}$$

- *Matrix:* rectangular array of numbers

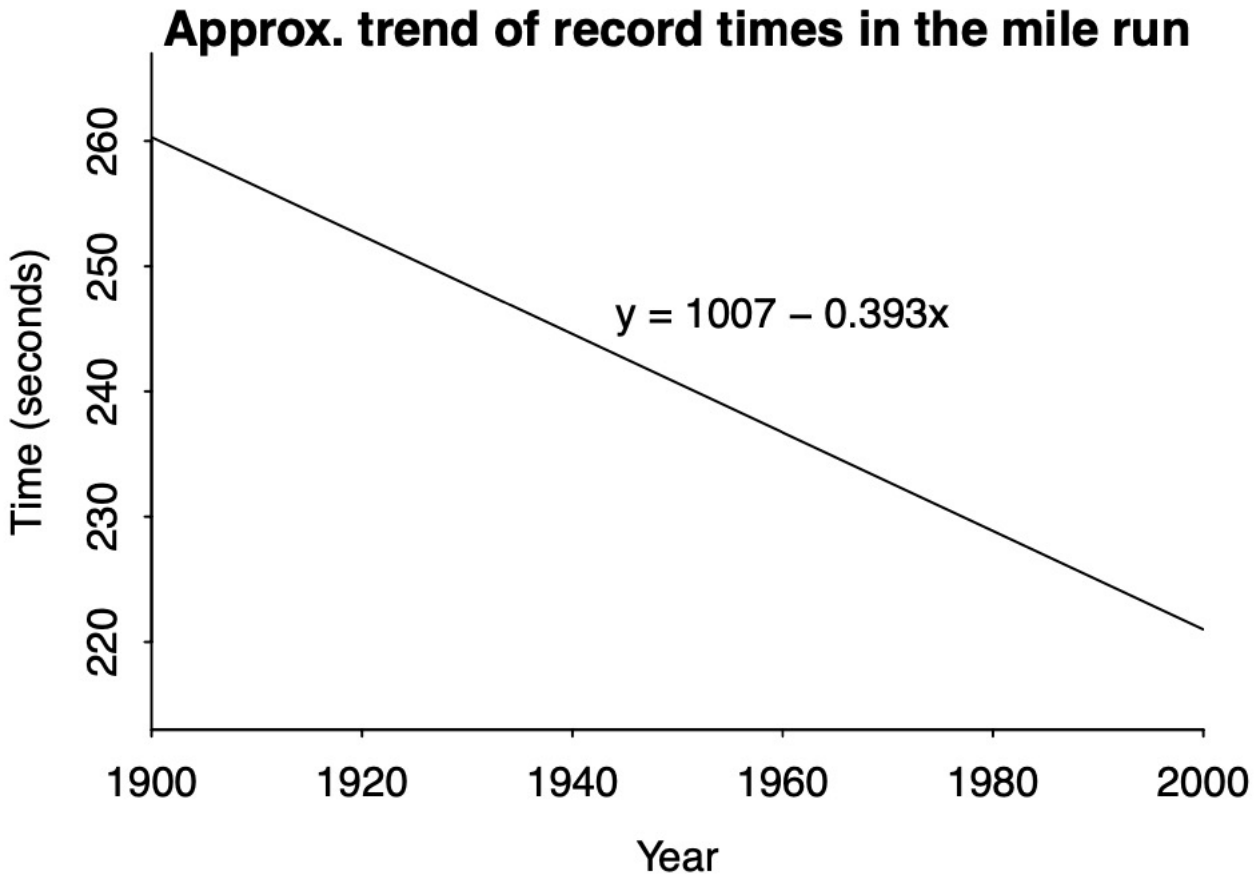
$$\begin{pmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 3 \end{pmatrix}$$

Multiplying a vector by a matrix

$$\hat{y} = X\hat{\beta}$$

$$\hat{y} = \begin{pmatrix} 43.3 \\ 46.3 \\ 55.3 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 46.3 \\ 3.0 \end{pmatrix}$$

Graphing a line



- Invent your own example (2 points in a plane) and calculate the corresponding equation!

Exponential and power-law relations

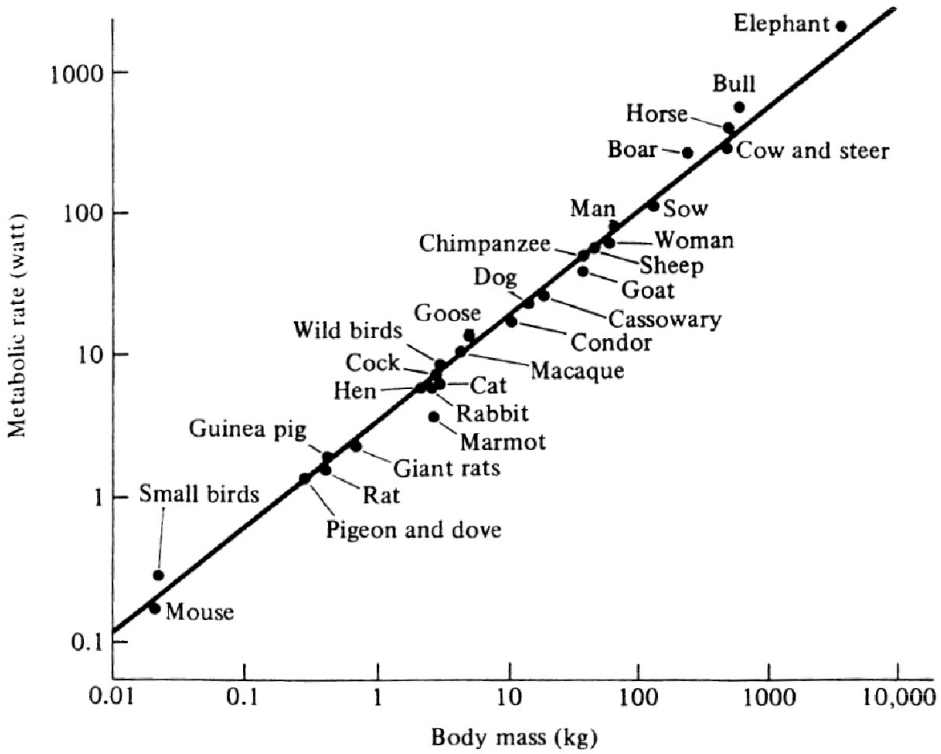


Figure 3.4 *Log metabolic rate vs. log body mass of animals, from Schmidt-Nielsen (1984). These data illustrate the log-log transformation. The fitted line has a slope of 0.74. See also Figure 3.5.*

- Invent your own examples for exponential and power law relations and calculate the corresponding equation!

Probability distributions

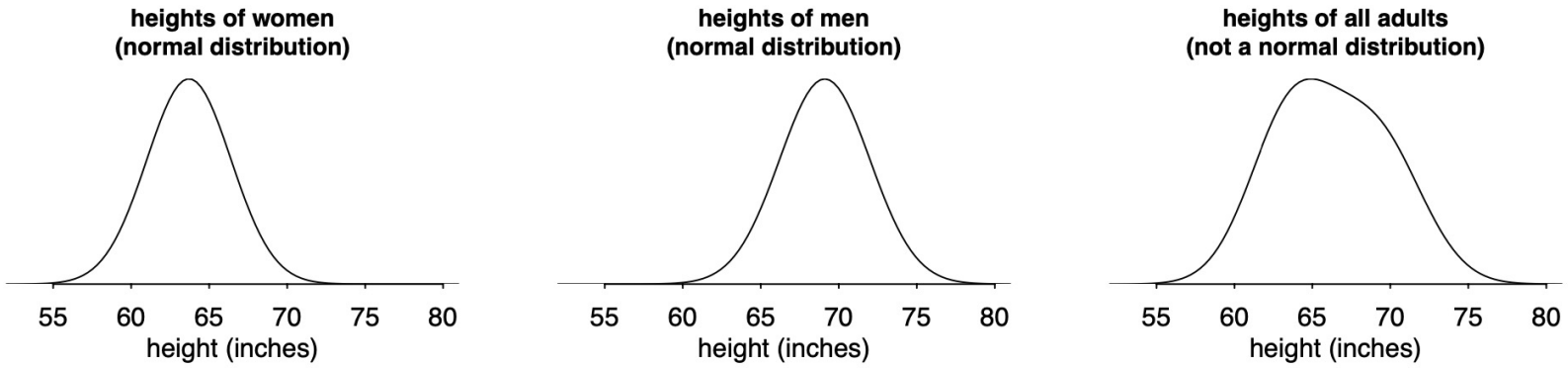


Figure 3.6 *(a) Heights of women, which approximately follow a normal distribution, as predicted from the Central Limit Theorem. The distribution has mean 63.7 and standard deviation 2.7, so about 68% of women have heights in the range* 63.7 ± 2.7. *(b) Heights of men, approximately following a normal distribution with mean 69.1 and standard deviation 2.9. (c) Heights of all adults in the United States, which have the form of a mixture of two normal distributions, one for each sex.*

The Gaussian ("normal") distribution

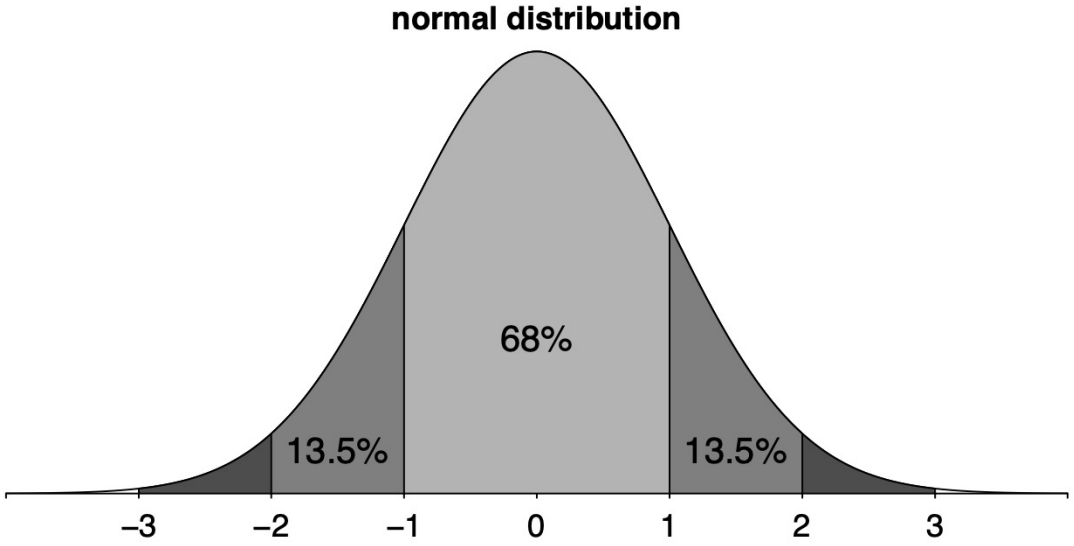


Figure 3.7 *Approximately 50% of the mass of the normal distribution falls within 0.67 standard deviations from the mean, 68% of the mass falls within 1 standard deviation from the mean, 95% within 2 standard deviations of the mean, and 99.7% within 3 standard deviations.*

The Gaussian ("normal") distribution

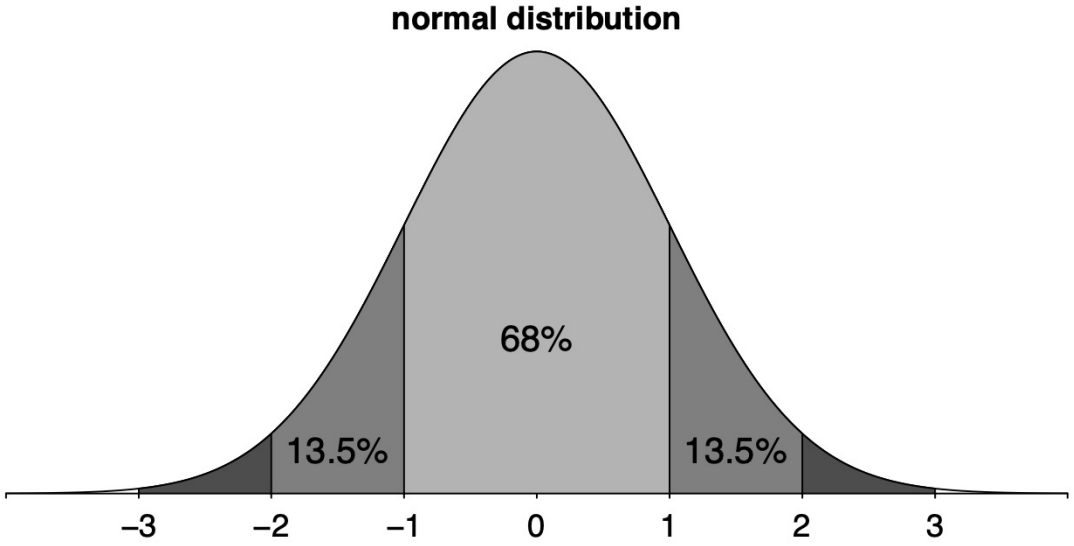


Figure 3.7 *Approximately 50% of the mass of the normal distribution falls within 0.67 standard deviations from the mean, 68% of the mass falls within 1 standard deviation from the mean, 95% within 2 standard deviations of the mean, and 99.7% within 3 standard deviations.*

The log-Gaussian distribution

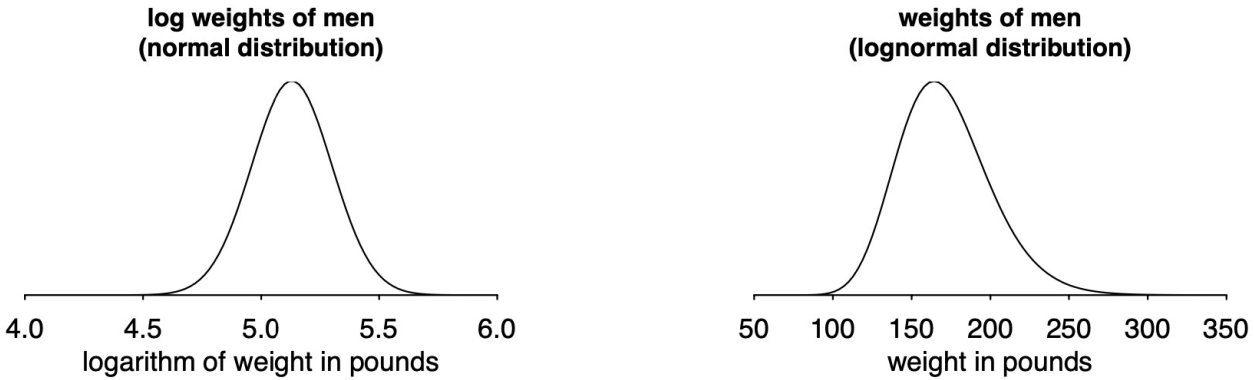


Figure 3.8 *Weights of men (which approximately follow a lognormal distribution, as predicted from the Central Limit Theorem from combining many small multiplicative factors), plotted on the logarithmic and original scales.*