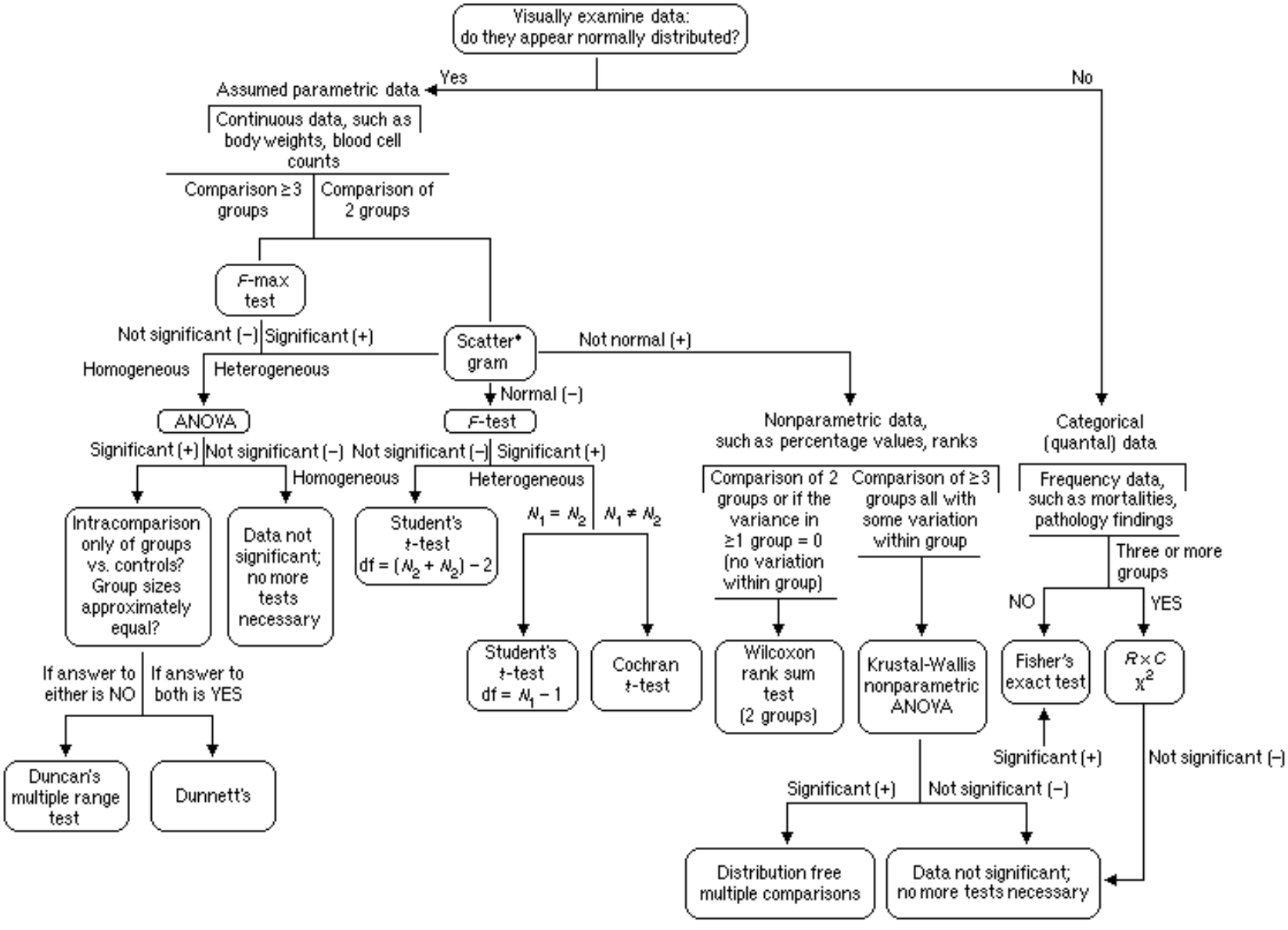
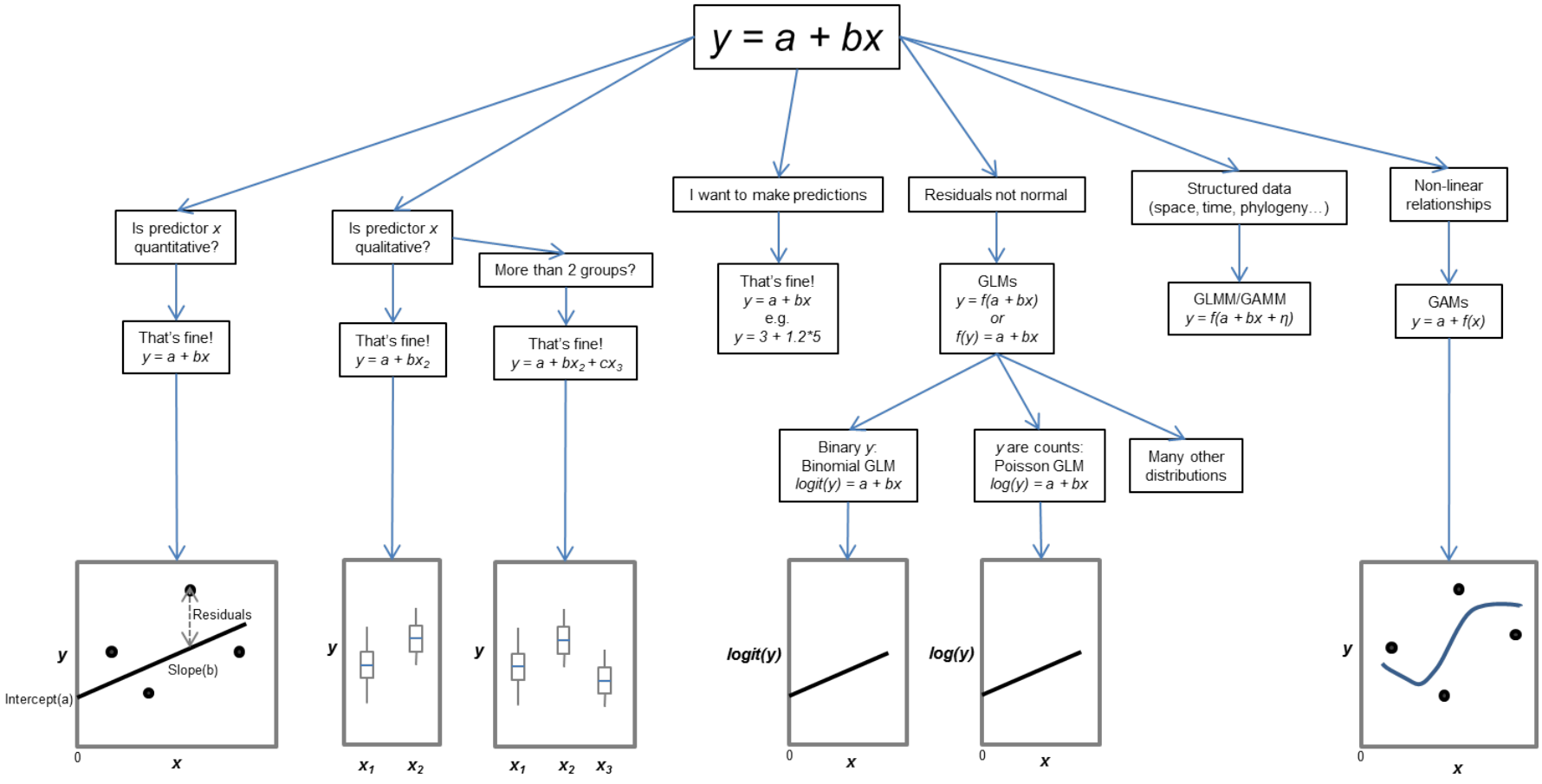
Introduction to linear models

Modern statistics are easier than this

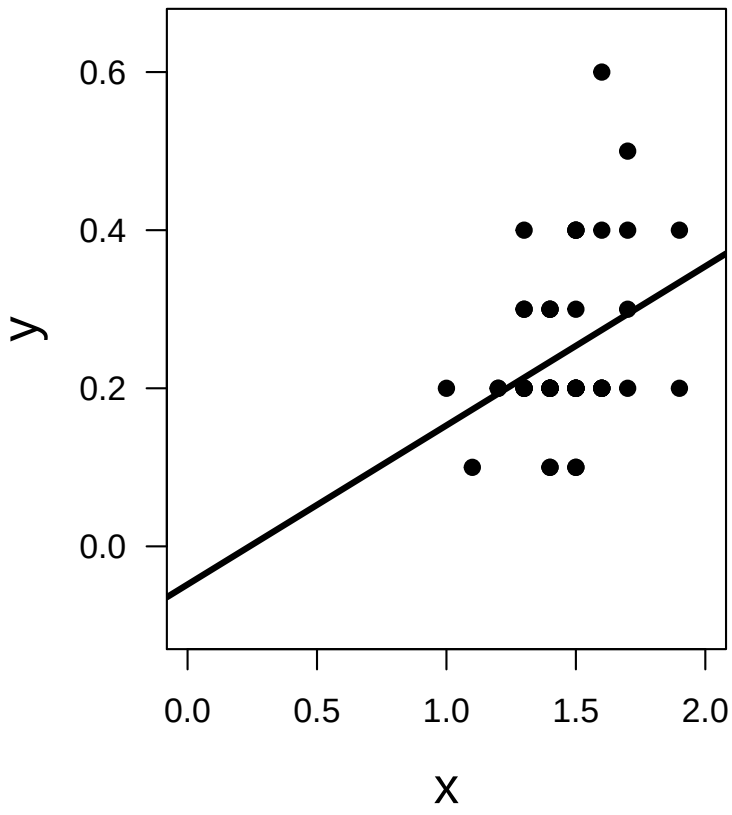


A unified framework



Our unified regression framework

$$y_i = a + bx_i + \varepsilon_i$$
$$\varepsilon_i \sim N\left(0, \sigma^2\right)$$



Data

y = response variable

x = predictor

Parameters

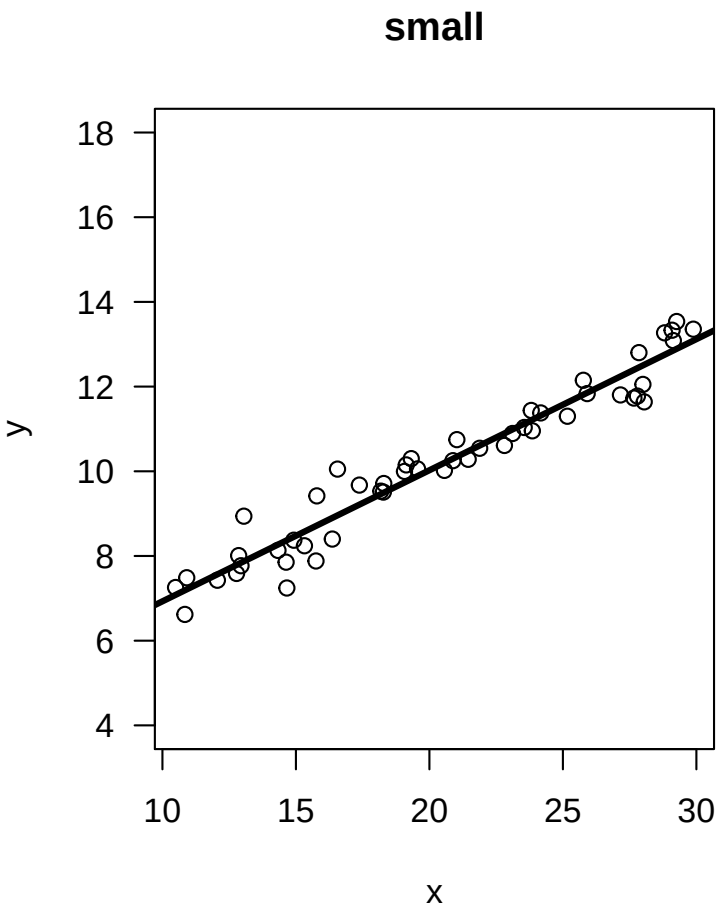
a = intercept

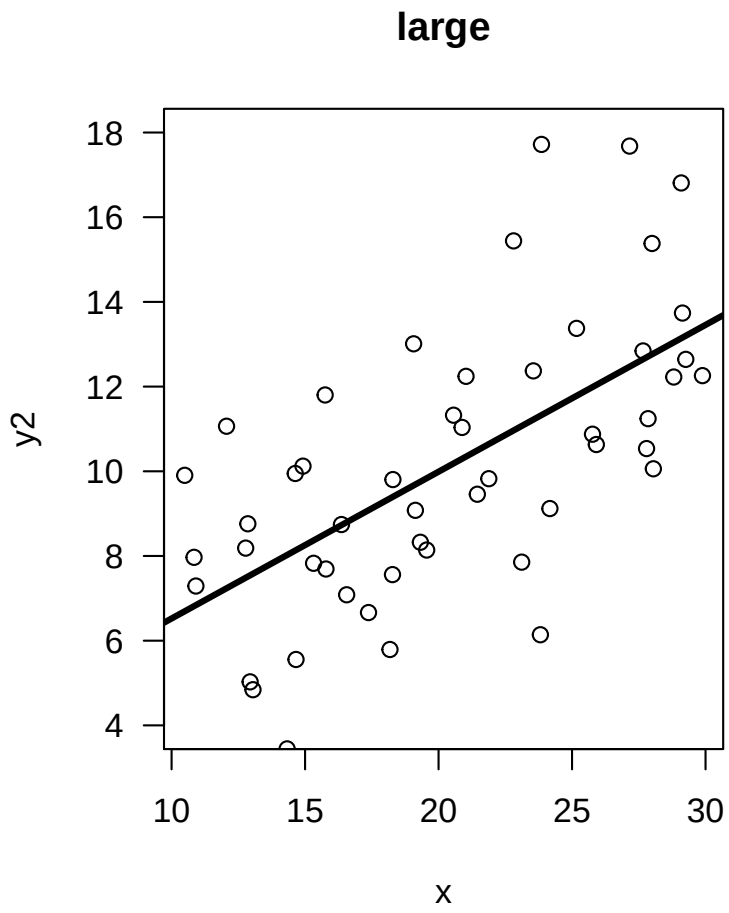
b = slope

σ = residual variation

ε = residuals

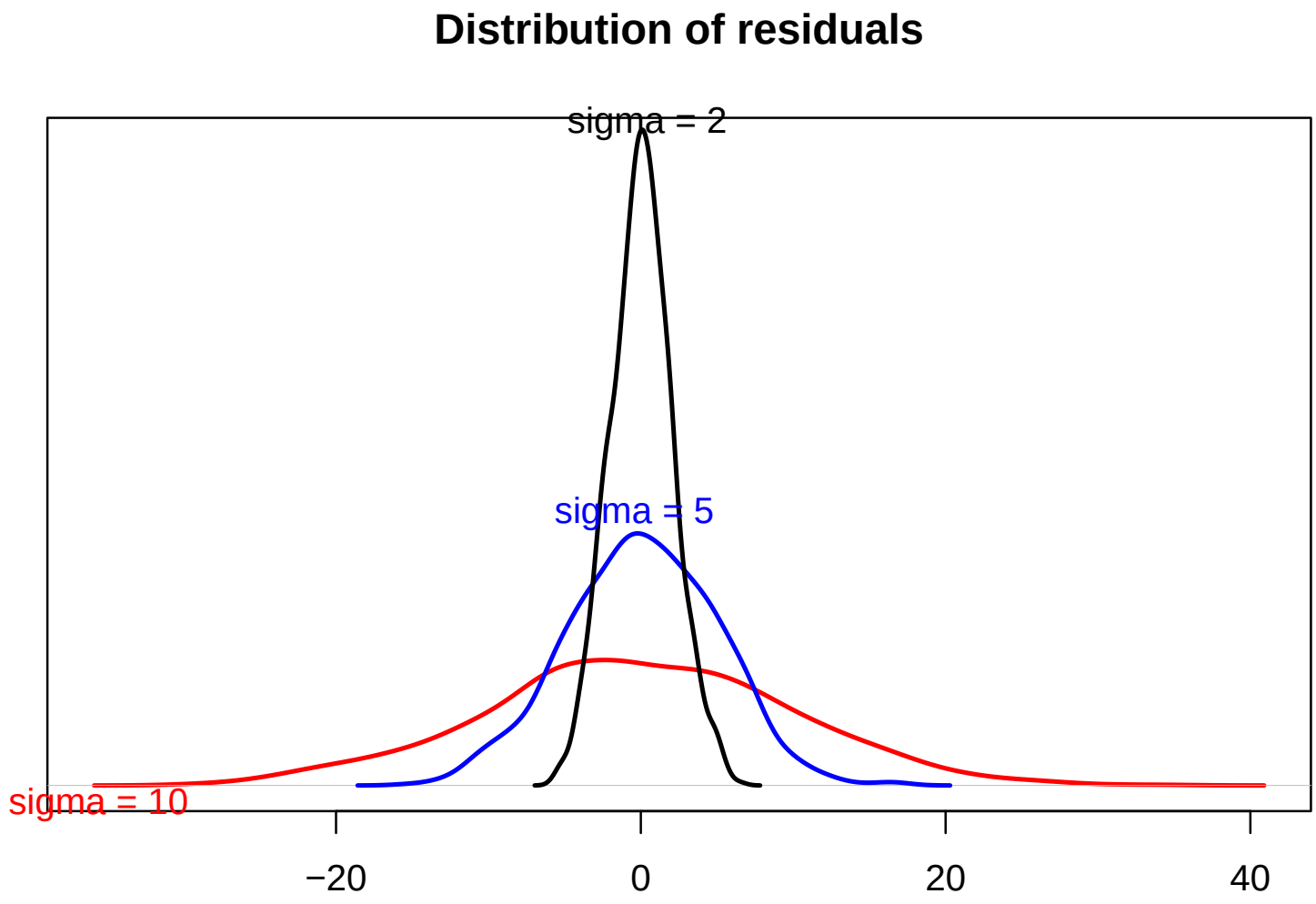
Residual variation (error)



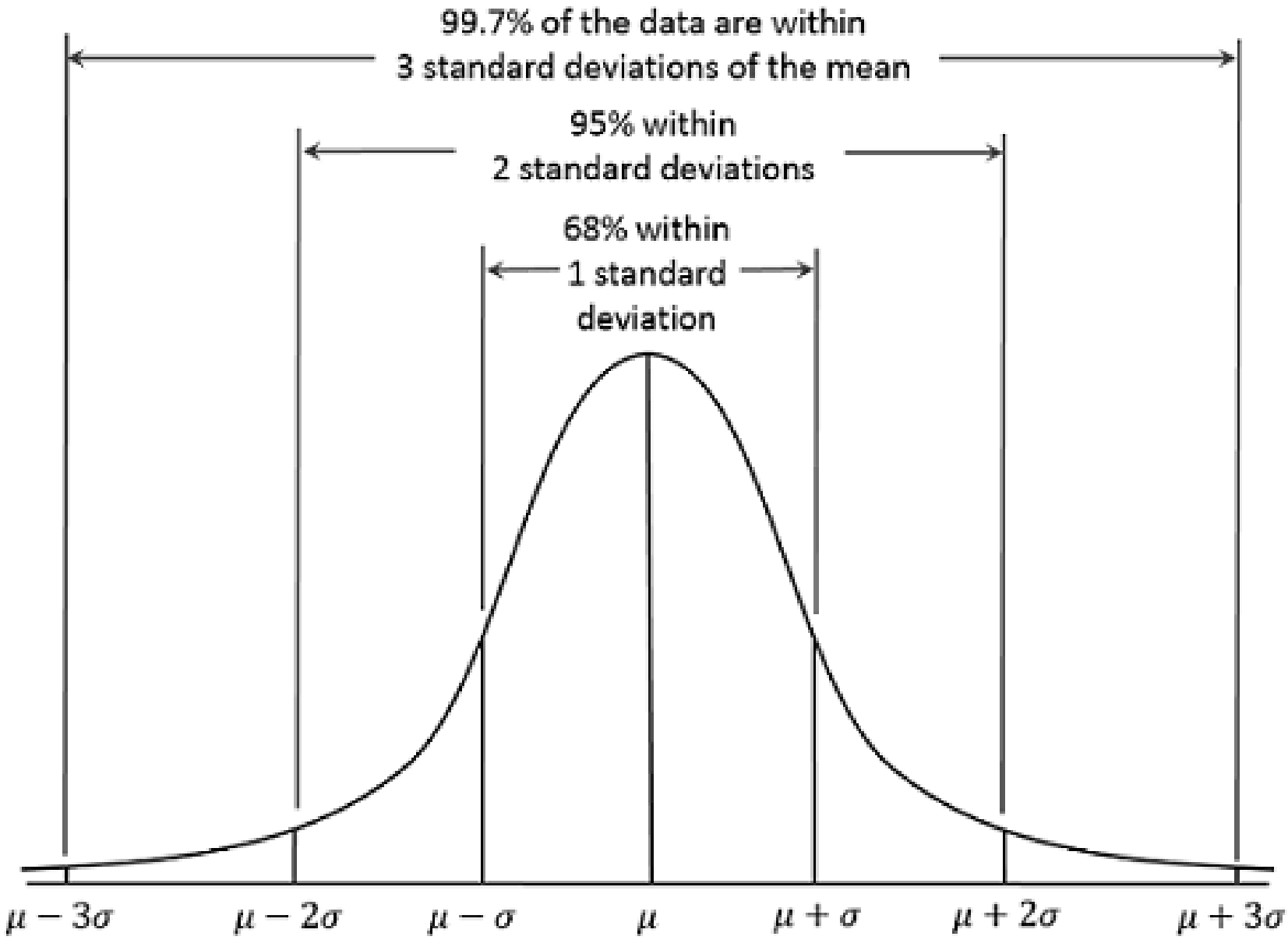


Residual variation

$$\varepsilon_i \sim N\left(0, \sigma^2\right)$$



In a Normal distribution



Different ways to write same model

$$y_i = a + bx_i + \varepsilon_i$$
$$\varepsilon_i \sim N\left(0, \sigma^2\right)$$

.

$$y_i \sim N\left(\mu_i, \sigma^2\right)$$
$$\mu_i = a + bx_i$$
$$\varepsilon_i \sim N\left(0, \sigma^2\right)$$