Introduction to linear models

Modern statistics are easier than this

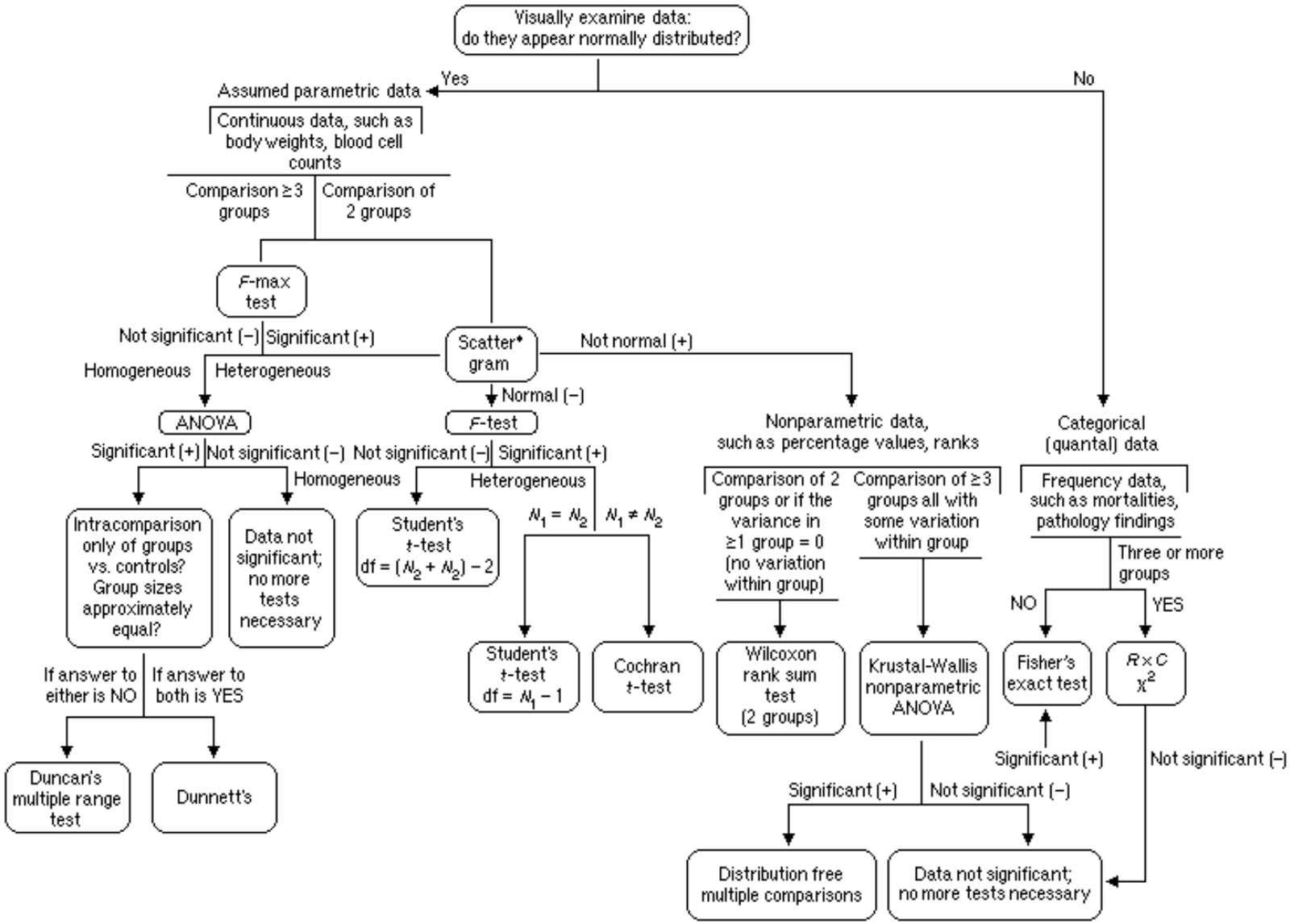
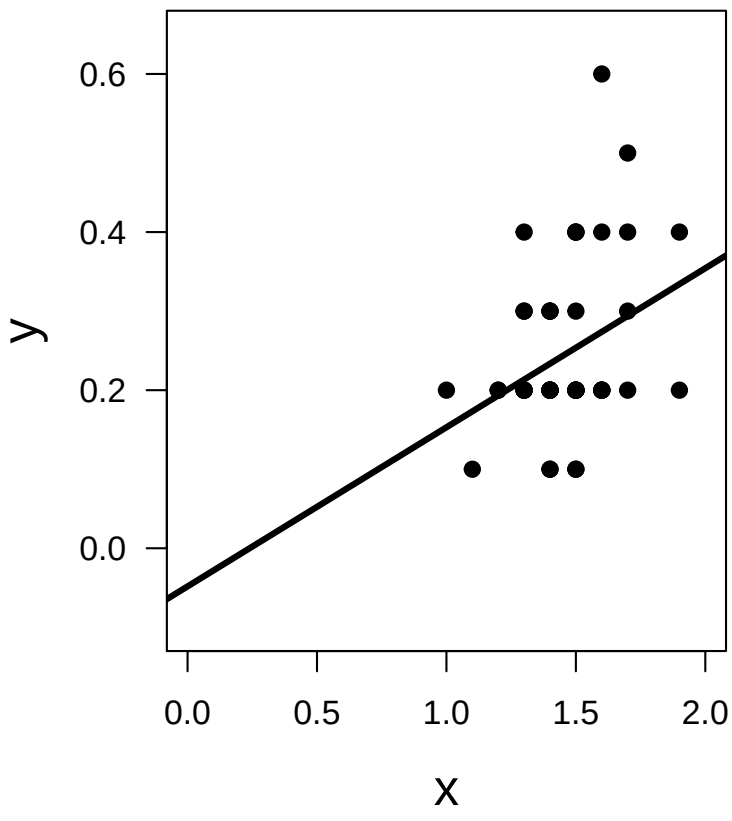


Figure 1:

Our overarching 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

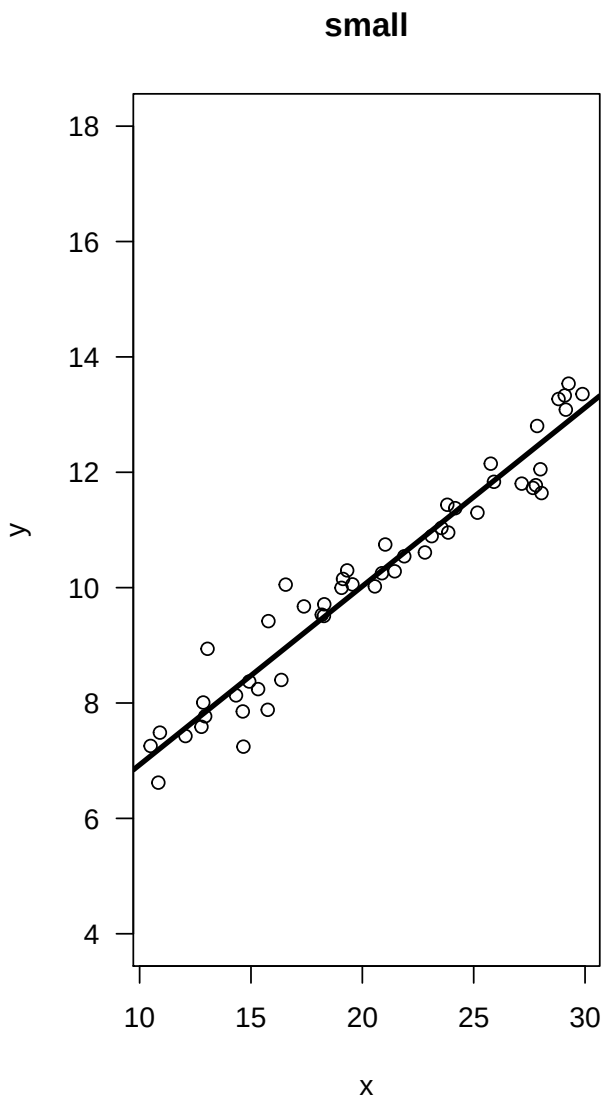
$a =$ intercept

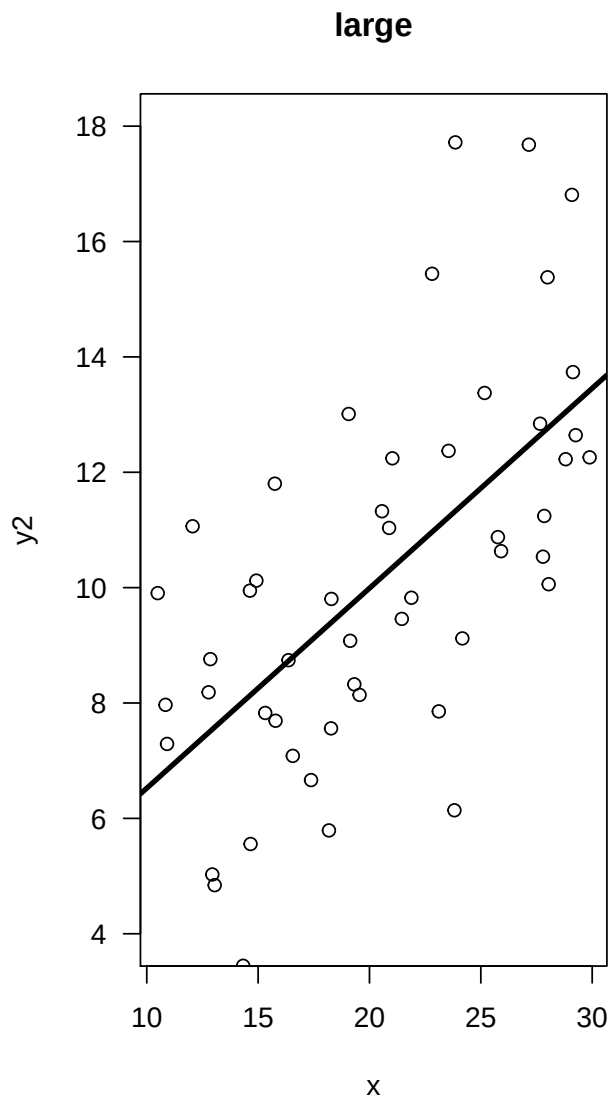
$b =$ slope

$\sigma =$ residual variation

$\varepsilon =$ residuals

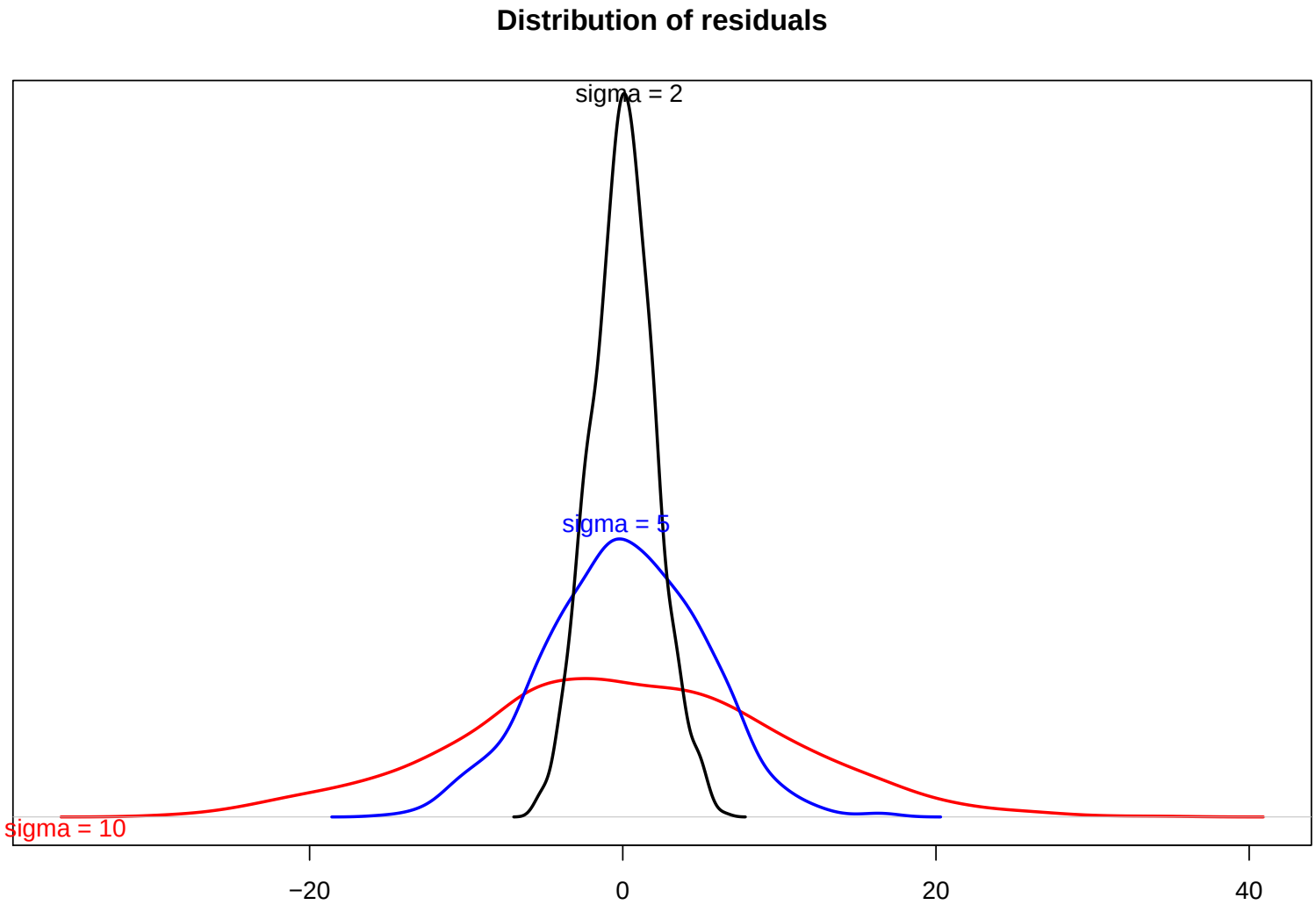
Residual variation (error)





Residual variation

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



In a Normal distribution

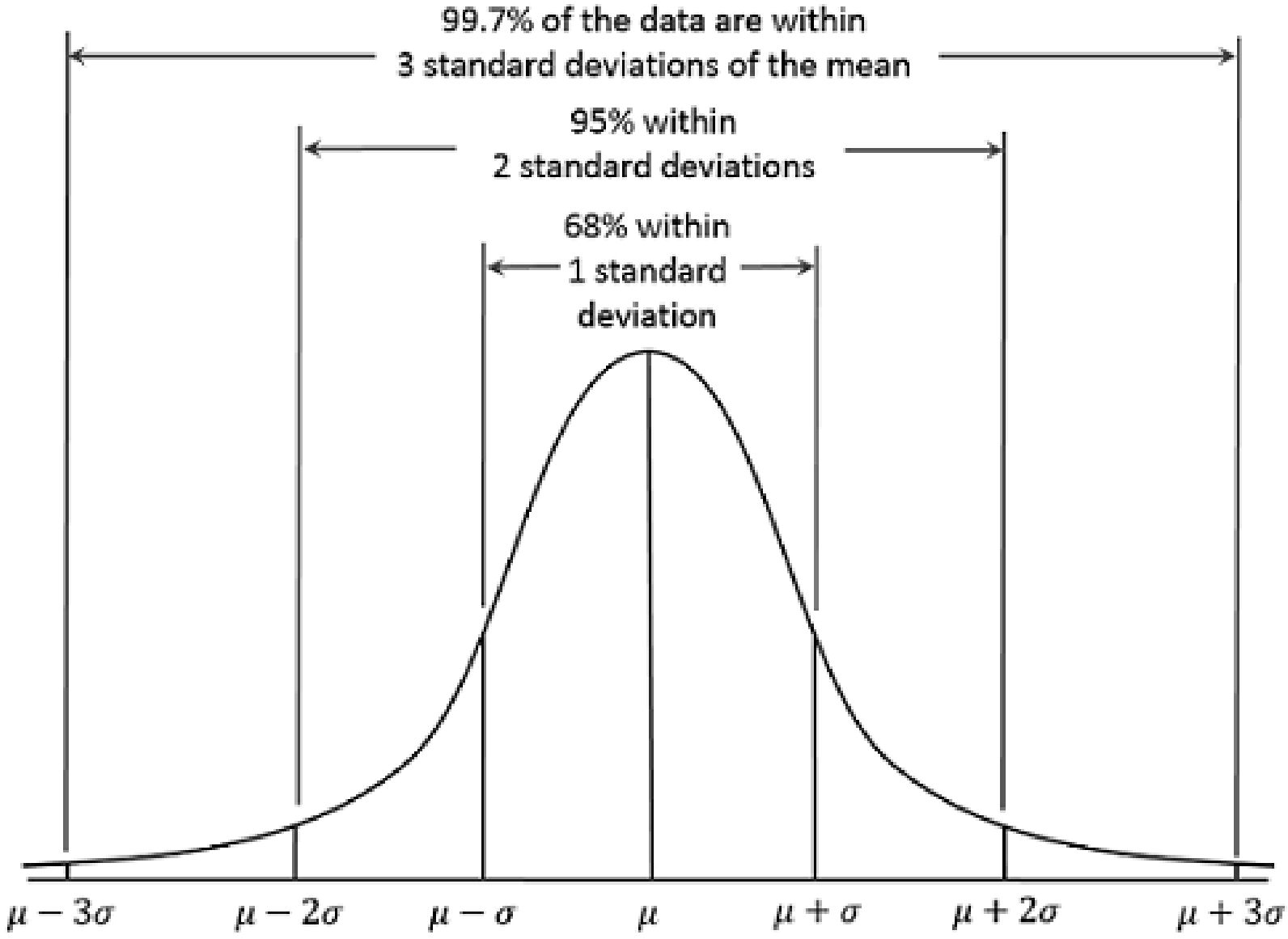


Figure 2:

Different ways to write same model

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

.

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