Linear model assumptions

Three of five

1. Linear relation between the response y and predictor x:

$$y_i = \alpha + \beta x_i + u_i$$

- 2. The sample $[y_i, x_i]$ is a random sample of the population
- 3. The errors (u_i) have zero mean when conditioned on x

$$E(u \mid x) = 0$$

OLS estimates $\hat{\alpha}$ and $\hat{\beta}$ are unbiased

Linear model assumptions

Two more

- 1. Linear relation between y and x:
- 2. The sample $[y_i, x_i]$ is a random
- 3. $E(y | x) = \alpha + \beta x$
- 4. Homoskedasticity, residual variance is constant: $Var(y \mid x) = \sigma^2$
- 5. The response y has a normal distribution given the predictor: $y \sim N(\alpha + \beta x, \sigma)$

OLS estimates $\hat{\alpha}$ and $\hat{\beta}$ are unbiased

OLS estimate of σ^2 is unbiased

 $\hat{\alpha}$ and $\hat{\beta}$ have known distributions and we can calculate p-values