Definition des linearen Modells

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_m x_m + \varepsilon$$

y: Beobachtungen (abhängige Variable)

$x_1, \ldots, x_m$: Regressoren (unabh. Variablen)

$\beta_1, \ldots, \beta_m$: unbekannte Parameter

ε: zufälliger Fehler

Beispiele für lineare Modelle

- $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \varepsilon$
- $y = \beta_0 + \beta_1 \log(x) + \varepsilon$
- $y = \beta_0 + \beta_1 \sin(2x) + \beta_2 \cos(x^2) + \varepsilon$

Linear bedeutet linear in den Parametern!

Kleinste Quadrate Schätzer I

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_m x_{mi} + \varepsilon_i$$

Voraussetzungen an ε_i:

- Erwartungswert $E(\varepsilon_i) = 0$
- Homoskedastizität: $Var(\varepsilon_i) = \sigma^2 < \infty$
- Unkorreliertheit: $Cov(\varepsilon_i, \varepsilon_j) = 0$ $(i \neq j)$
- Falls stochastische Regressoren, dann ε_i und x_{ki} stochastisch unabhängig

Dann: KQ-Schätzer **BLUE**!

Kleinste Quadrate Schätzer II

Zusätzlich:

- ε_i stochastisch unabhängig
- ε_i normalverteilt

Dann: KQ-Schätzer auch ML-Schätzer!

Außerdem: Umfangreiche Testtheorie steht zur Verfügung!

Vorteile linearer Modelle

- Einfache, klare Struktur
- Nichtlineare Zusammenhänge sind lokal durch lineare Funktionen approximierbar
- Optimale Schätzer für große Klasse von Schätzern
- Theorie (nahezu) vollständig bekannt

Nachteile linearer Modelle

- Schlechte Extrapolationseigenschaften (nur lokale Gültigkeit bzw. innerhalb der Daten)
- Große Anfälligkeit gegen Abweichungen von den Modellannahmen bzw. Ausreißern

Wichtig: Validierung des Modells!!!

Alternative: Robuste Schätzer

Beispiele Robuster Schätzer:

- Huber (1981): M-Schätzer
- Hampel et al (1986): M-Schätzer
- Yohai (1987): MM-Schätzer
- Rousseeuw and Leroy (1987): LMS- und LTS-Schätzer
- Rieder (1994): AL-Schätzer

Vorteile Robuster Schätzer

- Robust gegen Modellabweichungen verschiedener Art
  (d.h. Ideales Modell + Umgebung)

- M- bzw. AL-Schätzer **asymptotisch optimal** innerhalb großer Klasse von Schätzern

Nachteile Robuster Schätzer

- Kompliziertere Theorie

  I.d.R. nur asymptotische Aussagen; d.h., asymptotisch optimaler Schätzer, asymptotische Tests, usw.

- Höherer Rechenaufwand