



# Bayes'scher Ansatz

Peter von Rohr

# Frequentisten und Bayesianer

Unterschiede zwischen Frequentisten und Bayesianern bestehen hauptsächlich in

- deren Verständnis von Wahrscheinlichkeiten
- deren Unterteilung von Modell- und Datenkomponenten
- deren Techniken zur Schätzung von Parametern

# Bekannte und Unbekannte Größen

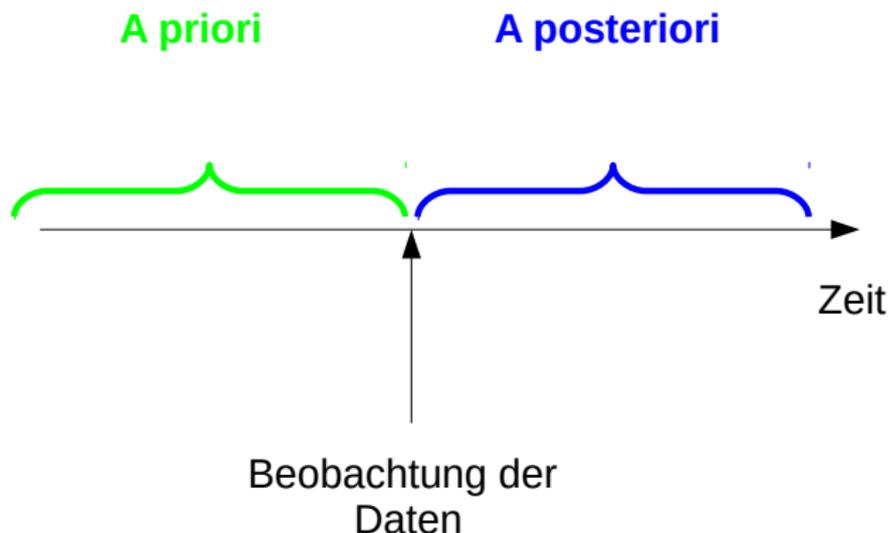
Angenommen: einfaches lineares Regressionsmodell

$$y_i = \beta_0 + \beta_1 x_{i1} + \epsilon_i \quad (1)$$

Was	bekannt	unbekannt
$y_i$	X	
$x_{i1}$	X	
$\beta_0$		X
$\beta_1$		X
$\sigma^2$		X

# Schätzung Unbekannter Grössen

- Parameterschätzung
- a posteriori Verteilung der unbekannt Grössen



# A Posteriori Verteilung

- Für Beispiel unseres Regressionsmodells:  $f(\beta, \sigma^2 | \mathbf{y})$
- Berechnung durch **Satz von Bayes**, basiert auf Definition der bedingten Wahrscheinlichkeit

$$\begin{aligned} f(\beta, \sigma^2 | \mathbf{y}) &= \frac{f(\beta, \sigma^2, \mathbf{y})}{f(\mathbf{y})} \\ &= \frac{f(\mathbf{y} | \beta, \sigma^2) f(\beta) f(\sigma^2)}{f(\mathbf{y})} \end{aligned} \quad (2)$$

# Komponenten der A Posteriori Verteilung

- $f(\mathbf{y}|\beta, \sigma^2)$ : Likelihood
- $f(\beta), f(\sigma^2)$ : a priori Verteilungen
- $f(\mathbf{y})$ : Normalisierungskonstante

# Problem

- A Posteriori Verteilung häufig nicht explizit als Verteilung darstellbar
- Lösung durch
  - 1 Julian Besag 1974: A Posteriori Verteilung ist bestimmt durch vollbedingte Verteilungen
  - 2 Gute Pseudozufallszahlen-Generatoren in Software
- A Posteriori Verteilung für Regression:  $f(\beta, \sigma^2 | \mathbf{y})$
- Vollbedingte Verteilungen für Regression:
  - $f(\beta_0 | \beta_1, \sigma^2, \mathbf{y})$
  - $f(\beta_1 | \beta_0, \sigma^2, \mathbf{y})$
  - $f(\sigma^2 | \beta_0, \beta_1, \mathbf{y})$

# Ablauf einer Analyse: Vorbereitung

- Schritt 1: Festlegung der a priori Verteilungen
- Schritt 2: Bestimmung der Likelihood aufgrund von Daten und Modell
- Schritt 3: Berechnung der a posteriori Verteilung
- Schritt 4: Bestimmung der vollbedingten Verteilungen

# Ablauf einer Analyse: Umsetzung

## Beispiel der Regression

- Schritt 5: Initialisierung aller unbekanntenen Größen ( $\beta_0, \beta_1, \sigma^2$ ) auf einen Startwert
- Schritt 6: Bestimme neuen Wert für  $\beta_0$  durch Ziehen einer Zufallszahl aus  $f(\beta_0|\beta_1, \sigma^2, \mathbf{y})$
- Schritt 7: Bestimme neuen Wert für  $\beta_1$  durch Ziehen einer Zufallszahl aus  $f(\beta_1|\beta_0, \sigma^2, \mathbf{y})$
- Schritt 8: Bestimme neuen Wert für  $\sigma^2$  durch Ziehen einer Zufallszahl aus  $f(\sigma^2|\beta_0, \beta_1, \mathbf{y})$
- Schritt 9: Loop viele Wiederholungen über Schritte 6-8 und speichere alle gezogenen Zahlen
- Schritt 10: Parameterschätzungen als Mittelwerte der gespeicherten Zufallszahlen

# Fragen und Dank

- Fragen?
- Vielen Dank!!