

# Bayes'scher Ansatz

Peter von Rohr

26 März 2018

# 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

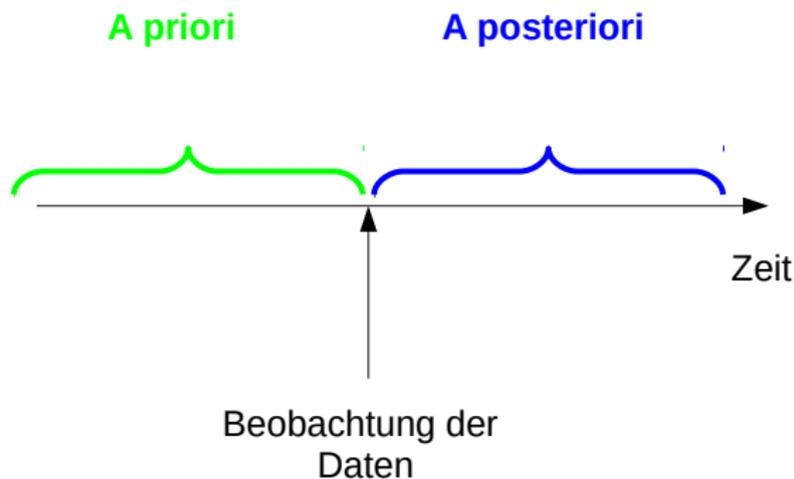
Angenommen: einfaches lineares Regressionsmodell

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

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 unbekanntenen Grössen gegeben die bekannten Grössen



# A Posteriori Verteilung

- ▶ Für unser Regressionsmodell
  - ▶ Annahme:  $\sigma^2$  sei bekannt  $\rightarrow$  weggelassen in Herleitung
  - ▶ a posteriori Verteilung für  $\beta$ :  $f(\beta|y)$
- ▶ Berechnung durch **Satz von Bayes**, dieser basiert auf der Definition der bedingten Wahrscheinlichkeit

$$\begin{aligned} f(\beta|y) &= \frac{f(\beta, y)}{f(y)} \\ &= \frac{f(y|\beta)f(\beta)}{f(y)} \\ &\propto f(y|\beta)f(\beta) \end{aligned} \tag{2}$$

# Komponenten der A Posteriori Verteilung

- ▶  $f(y|\beta)$ : Likelihood
- ▶  $f(\beta)$ : a priori Verteilung von  $\beta$ , oft als konstant angenommen, d.h. uninformative a priori Verteilung
- ▶  $f(y)$ : Normalisierungskonstante

→ uninformative a priori Verteilungen führen

$$f(\beta|y) \propto f(y|\beta)$$

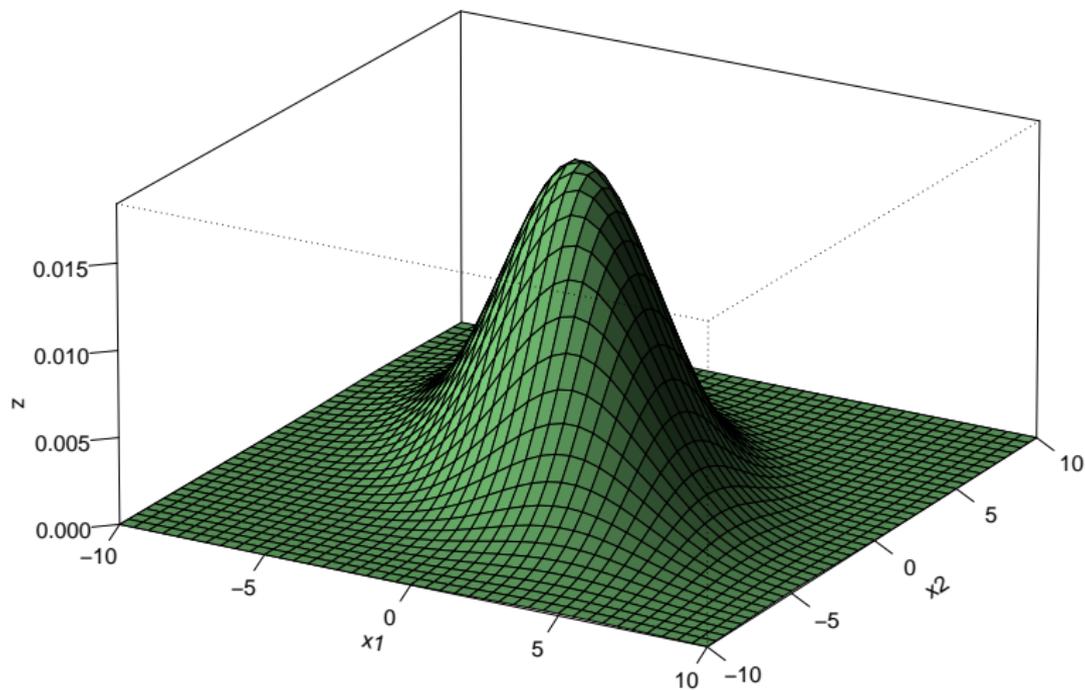
→ Annahme, dass  $y$  normalverteilt:

$$f(y|\beta) = (2\pi\sigma^2)^{-n/2} \exp \left\{ -\frac{1}{2} \frac{(y - X\beta)^T (y - X\beta)}{\sigma^2} \right\}$$

# Zwei-dimensionale Normalverteilung

## Two dimensional Normal Distribution

$$\mu_1 = 0, \mu_2 = 0, \sigma_{11} = 10, \sigma_{22} = 10, \sigma_{12} = 15, \rho = 0.5$$



# 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:  $f(\beta|y)$
- ▶ Vollbedingte Verteilungen für Regression:
  - ▶  $f(\beta_0|\beta_1, y)$
  - ▶  $f(\beta_1|\beta_0, y)$

wobei  $\beta$  der Vektor bestehen aus  $\beta^T = \begin{bmatrix} \beta_0 & \beta_1 \end{bmatrix}$

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

## R-Programm

```
beta = c(0, 0); meanBeta = c(0, 0)
niter = 10000 # number of samples
for (iter in 1:niter) {# loop for Gibbs sampler
  # sampling intercept
  w = y - X[, 2] * beta[2]; x = X[, 1]
  xpxi = 1/(t(x) %*% x); betaHat = t(x) %*% w * xpxi
  # using residual var = 1
  beta[1] = rnorm(1, betaHat, sqrt(xpxi))
  # sampling slope
  w = y - X[, 1] * beta[1]; x = X[, 2]
  xpxi = 1/(t(x) %*% x)
  betaHat = t(x) %*% w * xpxi
  # using residual var = 1
  beta[2] = rnorm(1, betaHat, sqrt(xpxi))
  meanBeta = meanBeta + beta
}
```

# Fragen und Dank

- ▶ Fragen?
- ▶ Vielen Dank!!