Bayesian Approaches

Peter von Rohr

2021-03-29

Statistics

The world of statistics is divided into

- **Frequentists** and
- **Bayesians**

Divergence in

- understanding of probability
- differentiation between components of a model and the data
- techniques to estimate parameters

F vs B

Topic	Frequentists	Bayesians
Probability	Ratio between cardinalities of sets	Measure of uncertainty
Model and Data	Parameter are unknown, data are known	Differentiation between knowns and unknowns
Parameter Estimation	ML or REML are used for parameter estimation	MCMC techniques to approximate posterior distributions

Linear Model

$$y_i = \beta_0 + \beta_1 x_{i1} + \epsilon_i$$

Table 1: Separation Into Knowns And Unknowns

Term	Known	Unknown
y_i	X	
x_1	X	
β_0		X
β_1		X
σ^2	X	

Example Dataset

Table 2: Dataset for Regression of Body Weight on Breast Circumference for ten Animals

Animal	Breast Circumference	Body Weight
1	176	471
2	177	463
3	178	481
4	179	470
5	179	496
6	180	491
7	181	518
8	182	511
9	183	510
10	184	541

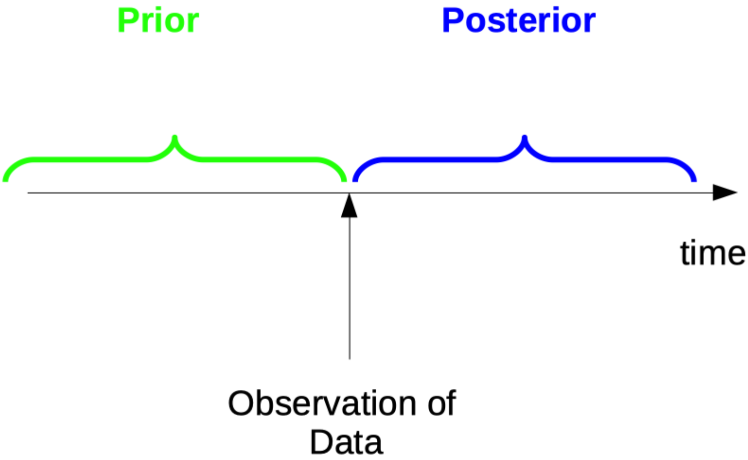
Estimation Of Unknowns

- Estimates of unknowns $\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$
- Using Bayes Theorem:

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

where $f(\beta)$: prior distribution and $f(y|\beta)$: likelihood

Prior and Posterior



Posterior Distribution

- How to get to posterior distribution $f(\beta|y)$
- Use regression as example
- β is a vector with two components, $\beta^{T} = \left[\begin{array}{cc} \beta_0 & \beta_1 \end{array} \right]$
- **Solution**: accumulation of samples from full conditional posterior distributions leads to samples from posterior distribution

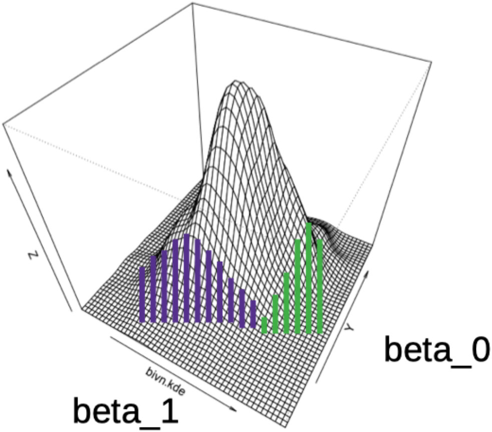
Prior and Likelihood

- What are the distributional assumptions (for regression example and in general)
- Prior: $f(\beta)$ usually assumed to be uniform
- Likelihood: $f(y|\beta)$ assumed to be multivariate normal

Regression

- Full conditional distributions
  - intercept: $f(\beta_0|\beta_1, y)$ is a normal distribution
  - slope: $f(\beta_1|\beta_0, y)$ is normal distribution
- Draw random numbers from full conditional distributions in turn
- Result will be samples from posterior distribution

Full Conditional Distributions



Estimates from Samples

- Given Samples from posterior distribution $f(\beta|y)$
- Estimates are computed as empirical means and standard deviation based on the samples

$$\beta_{Bayes} = \frac{1}{N} \sum_{t=1}^{N} \beta^{(t)}$$

with N samples drawn from full conditional distributions

Gibbs Sampler

- Implementation using full conditional distributions
- Use Gibbs Sampler for regression example
- Step 1: Start with initial values $\beta_0 = \beta_1 = 0$
- Step 2: Compute mean and standard deviation for full conditional distribution of β_0
- Step 3: Draw random sample for β_0
- Step 4 and 5: same for β_1
- Step 6: Repeat 2-5 N times
- Step 7: Compute mean from samples