Fixed Linear Effect Model
using least squares
=> did not work

Bayesian Approaches

GBLUP
using MME

Peter von Rohr

Fixed linear models
use Bayesian
parameter estimation

23.03.2020

Fixed linear model
using LASSO

Statistics

what we have discussed so far
> least squares
> BLUP

The world of statistics is divided into

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

Frequentists
Ratio between successful outcomes and total number of outcomes.

Bayesian: Degree of belief or degree of information

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

Data: y
Model: y = Xb + e
Parameter (unknown): b

Knowns: observable data
Unknows: parameter of the model, missing data

Linear Model

Example: Regression Data-Analysis
where beta_0: intercept;
beta_1: regression slope
x_1: breast circumference
y: body weight

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

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	

Assumption: known error variance

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

Bayesians do not differentiate between fixed effects and random effects but all knowns and unknowns are random variables with a certain distribution (or density). For distribution , we use f()

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

Estimates of unknows (beta) are based on the posterior distribution of the unknowns given the knowns

constant

"proportional to"

Prior

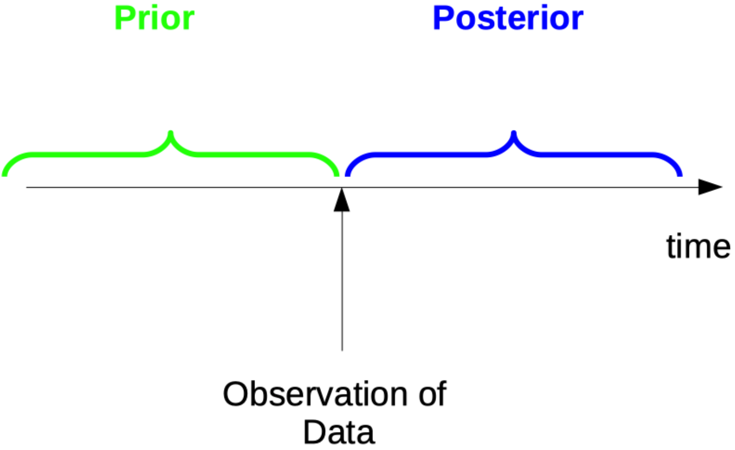
$$
\begin{aligned}
f(\beta|y) &= \frac{f(\beta, y)}{f(y)} \\
&= \frac{f(y|\beta)f(\beta)}{f(y)} \\
&\propto \boxed{f(y|\beta)}\boxed{f(\beta)}
\end{aligned}
$$

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

Prior and Posterior

Any information about unknows beta that originates prior to the event of data observation is encoded in the prior distribution f(beta) of the unknown beta

Any information about the unknowns after having observed the data is encoded in the conditional posterior distribution f(beta I y)
of the unknowns given the data



Posterior Distribution

intercept

two dimensional distribution

slope

- 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

Implied in the solution: Approximation of distributions (densities) by random samples. This was only possible because efficient random number generators are available on the computer.
Example: In R use the command rnorm(1000) gives you 1000 random numbers which are from a normal distribution with mean = 0 and sd = 1

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

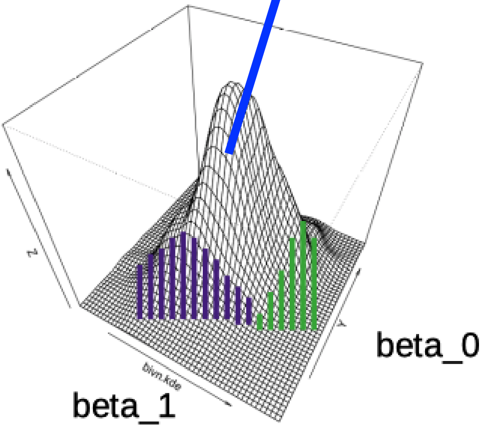
Regression

Solution to get to f(beta l y) is obtained by specifying the full conditional distributions

- 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

two-dimensional normal distribution f(beta l y)



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