Baysian Statistics by Means of Heuristic Methods

Antonio Nuić

Department of Mathematics, Osijek

september 2018.

Based on Bayes' rule that provides a rational method for updating beliefs in light of new information.
Bayesian methods provide:

- parameter estimates with good statistical properties;
- parsimonious descriptions of observed data;
- predictions for missing data and forecasts of future data;
- a computational framework for model estimation, selection and validation.

1. For each numerical value $\theta \in \Theta$, our prior distribution $p(\theta)$ describes our belief that θ represents the true population characteristics.
2. For each $\theta \in \Theta$ and $y \in \mathcal{Y}$, our sampling model $p(y|\theta)$ describes our belief that y would be the outcome of our study if we knew θ to be true. Once we obtain the data y, the last step is to update our beliefs about θ.
3. For each numerical value of $\theta \in \Theta$, our posterior distribution $p(\theta|y)$ describes our belief that θ is the true value, having observed dataset y.

Definition

Let X be random variable (or random vector) and x its observed value. Let $f(x|\theta)$ be its density function. Function of parameter θ:

$$L(\theta|x) := f(x|\theta)$$

, $\theta \in \Phi$ (1.2) is called likelihood function of statistic X based on observed value x.

Theorem

(Bayes' theorem) Let A and B be events in Ω and assume $P(B) > 0$. Then:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}.$$

This equation is called Bayes' formula.

Example

Assume we have a box that contains n white or black balls. Proportion of white balls is represented with random variable P*. If the first drawn ball is white, what is the probability of* $P = p_0$*. To solve this we assume that all possible outcomes of* P *are equally likely. In other words, we choose uniform distribution on* $\{0, \frac{1}{n}, \frac{2}{n}..., 1\}$ *as prior distribution of* P*. Then we derive posterior distribution of* P *by using Bayes' theorem and we have:*

$$P(P = p_0 | \text{ given data}) = \frac{p_0 \cdot \frac{1}{n+1}}{\frac{1}{2}} = \frac{2p_0}{n+1}.$$

For $n = 3$:
Prior distribution:

$$\begin{pmatrix} 0 & \frac{1}{3} & \frac{2}{3} & 1 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$$

Posterior distribution:

$$\begin{pmatrix} 0 & \frac{1}{3} & \frac{2}{3} & 1 \\ 0 & \frac{1}{6} & \frac{1}{3} & \frac{1}{2} \end{pmatrix}$$

Example

Out of 860 *individuals* $y = 441$ (51%) *agree with supreme court ruling. Let* θ *be the population proportion agreeing with the ruling. Using a binomial sampling model and a uniform prior distribution, the posterior distribution of* θ *is beta*(442, 420)*. Using the Monte Carlo algorithm, we can obtain samples of the log-odds* $\gamma = \log\left[\frac{\theta}{1-\theta}\right]$

R code:

$a = 1$; $b = 1$

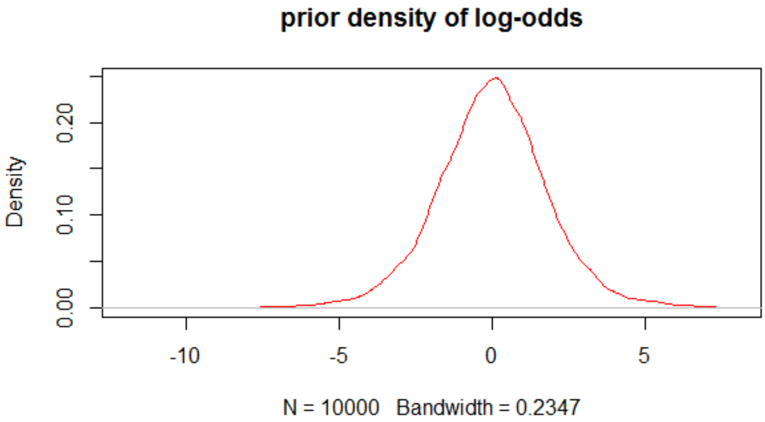
theta.prior.mc=rbeta (10000, a , b)

gamma.prior.mc= log(theta.prior.mc/(1-theta.prior.mc))

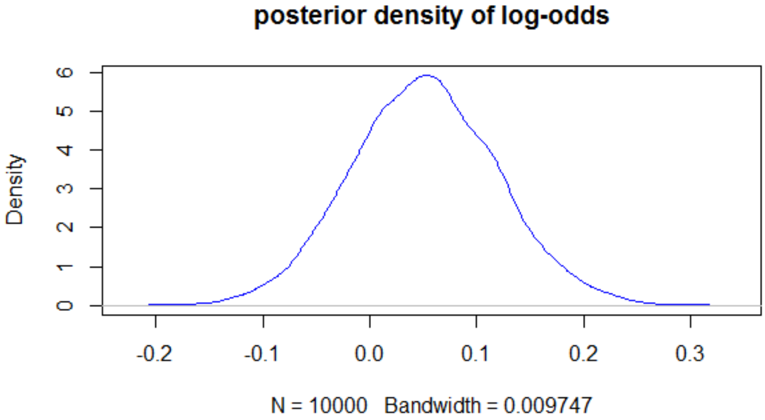
$n0 = 860 - 441$; $n1 = 441$

theta.post.mc=rbeta (10000 , $a + n1$, $b + n0$)

gamma.post.mc= log(theta.post.mc/(1?theta.post.mc))



Slika: Prior density function of log-odds



Slika: Posterior density function of log-odds

P.D.Hoff,*A First Course in Bayesian Statistical Methods*, Department of Statistics University of Washington, Springer 2009.

P. Congdon *Applied Bayesian Modeling*, University of London, UK, John Wiley and Sons, 2003.

MacKay, D. J. C. (1999). *Introduction to Monte Carlo methods*. In Learning in Graphical Models, M. I. Jordan (ed), MIT Press, 1999.