Bayesian Statistics

Luiz Max de Carvalho[lmax.fgv@gmail.com]

PhD-level course
School of Applied Mathematics (EMAp/FGV), Rio de Janeiro.

17 de Março de 2021

Welcome!

- This is a 60-hour, PhD-level course on Bayesian inference.
- We have 11 planned weeks. Reading material is posted at https://github.com/maxbiostat/BayesianStatisticsCourse/
- Assessment will be done via a written exam (70% marks) and an assignment (30%);
- Tenets:
  - Respect the instructor and your classmates;
  - Read before class;
  - Engage in the discussion;
  - Don't be afraid to ask/disagree.
- Books are
  - Robert (2007);
  - Hoff (2009);
  - Bernardo and Smith (2009).

Bayes's Theorem

What do

$$\Pr(A \mid B) = \frac{\Pr(B \mid A)\Pr(A)}{\Pr(B)}, \tag{1}$$

and

$$\Pr(A_i \mid B) = \frac{\Pr(B \mid A)\Pr(A)}{\sum_{i=1}^{n} \Pr(B \mid A_i)\Pr(A_i)}, \tag{2}$$

and

$$p(\theta \mid \boldsymbol{y}) = \frac{l(\boldsymbol{y} \mid \theta)\pi(\theta)}{\int_{\Theta} l(\boldsymbol{y} \mid t)\pi(t)\,dt}, \tag{3}$$

and

$$p(\theta \mid \boldsymbol{y}) = \frac{l(\boldsymbol{y} \mid \theta)\pi(\theta)}{m(\boldsymbol{y})}, \tag{4}$$

all have in common?

Statistical model: informal definition

Stuff you say in the bar:

Definition 1 (Statistical model: informal)

DeGroot, def 7.1.1, pp. 377 A statistical model consists in identifying the random variables of interest (observable and potentially observable), the specification of the joint distribution of these variables and the identification of parameters (θ) that index this joint distribution. Sometimes it is also convenient to assum that the parameters are themselves random variables, but then one needs to specify a joint distribution for θ also.

Statistical model: formal definition

Stuff you say in a Lecture:

Definition 2 (Statistical model: formal)

McCullagh, 2002. Let $\mathcal{X}$ be an arbitrary sample space, Θ a non-empty set and $\mathcal{P}(\mathcal{X})$ the set of all probability distributions on $\mathcal{X}$, i.e. $P : \Theta \to [0, \infty)$, $P \in \mathcal{P}$. A parametric statistical model is a function $P : \Theta \to \mathcal{P}(\mathcal{X})$, that associates each point $\theta \in \Theta$ to a probability distribution P_θ over $\mathcal{X}$.

Examples:

- Put $\mathcal{X} = \mathbb{R}$ and $\Theta = (-\infty, \infty) \times (0, \infty)$. We say P is a *normal* (or *Gaussian*) statistical model[1] if for every $\theta = \{\mu, \sigma^2\} \in \Theta$,

$$P_\theta(x) \equiv \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \; x \in \mathbb{R}.$$

- Put $\mathcal{X} = \mathbb{N} \cup \{0\}$ and $\Theta = (0, \infty)$. P is a Poisson statistical model if, for $\lambda \in \Theta$,

$$P_\lambda(k) \equiv \frac{e^{-\lambda}\lambda^k}{k!}, \; k = 0, 1, \ldots$$

[1] Note the abuse of notation: striclty speaking, P_θ is a probability **measure** and not a *density* as we have presented it here.

References

Bernardo, J. M. and Smith, A. F. (2009). *Bayesian theory*, volume 405. John Wiley & Sons.

DeGroot, M. H. and Schervish, M. J. (2012). *Probability and Statistics.* Pearson Education.

Hoff, P. D. (2009). *A first course in Bayesian statistical methods*, volume 580. Springer.

Robert, C. (2007). *The Bayesian choice: from decision-theoretic foundations to computational implementation.* Springer Science & Business Media.