math*overflow*

what can be said about the choice of a prior in Bayesian statistics?

Asked 8 years, 9 months ago   Active 8 years, 9 months ago   Viewed 2k times

▲

10

▼

★

4

When reading about the Bayesian approach to statistics, priors are an important component of the whole methodology.

Yet, it seems like priors are chosen without any specific theoretical motivation. There is the theory of conjugate priors, which is motivated mostly computationally, I believe, but still, I haven't seen a result in the line of "the choice of a certain prior will lead to faster convergence rate" or something similar to that.

Is there a good reference that analyzes the choice of a prior somehow, instead of always assuming that it is given, and assuming that it is completely the modeler's choice?

big-picture   st.statistics

asked Feb 17 '11 at 20:58

bayesianlearner
101   3

▲
⚑

(just as a note: I am familiar with the empirical Bayesian approach, but I don't think it gives the kind of information I am looking for. I would be happy to be proven wrong about it.) – bayesianlearner   Feb 17 '11 at 21:04

---

4   ▲
⚑

Semi-rant follows: my experience in the military-industrial complex has involved lots of cases where people just guess discrete (or few-parameter) priors based on nothing more than their gut feeling or its equivalent, and even more often folks don't bother to do a stability analysis of the results. So yes, the choices of priors are important. But alas, priors are very often chosen recklessly. Garbage in, garbage out. Nowadays when I hear "Bayesian" my first impulse is to cringe unless or until I become convinced that the person saying that actually demonstrates some level of care. – Steve Huntsman Feb 17 '11 at 21:09

---

1   ▲
⚑

You could search for literature from "objective Bayesian" community, who focus on principled methods on choosing priors, esp survey papers from people like Wasserman, Berger, Bernardo, Dawid – Yaroslav Bulatov Feb 17 '11 at 23:02

---

5 Answers

▲

13

▼

There are many approaches to this problem. Here are three.

1. The **subjective** Bayes approach says the prior should simply quantify what is known or believed before the experiment takes place. Period. End of discussion.

In practice someone may use any and all of these approaches, even within the same model. For example, they may use a subjective prior on parameters where there is a considerable amount of prior knowledge and use a reference prior on other parameters that are less important or less understood.

Often it simply doesn't matter much what prior you use. For example, you might show that a variety of priors, say an optimistic prior and a pessimistic prior, lead to essentially the same conclusion. This is particularly the case when there's a lot of data: the impact of the prior fades as data accrue. But for other applications, such as hypothesis testing, priors matter more.

edited Feb 18 '11 at 3:08   answered Feb 17 '11 at 23:18

John D. Cook
3,785   36   56

▲

4

▼

You can (at least in spirit, modulo technical details) define a prior by assigning to each hypothesis H the probability $2^{-K(H)}$, where K is Kolmogorov complexity. Obviously this prior is not as good as the prior that assigns to each hypothesis its true probability, but there are multiple senses in which it's (provably) almost as good. There are, for example, theorems saying that when you use this method to predict the values of a binary string, the total expected prediction error over infinitely many predictions is bounded above by a (finite) constant. An important key phrase here is minimum message length.

answered Feb 19 '11 at 4:54

Steven Landsburg
18.5k   3   75   128

1   ▲   That's a nice way to precisely (up to an additive constant depending on the model of computation) state Occam's razor. (Yes, I realize that I have just perpetrated the
⚑   mother of all split infinitives.) – Tracy Hall Feb 19 '11 at 5:10

▲   Another important keyword in this context is "universal Solomonoff prior". – Marcin Kotowski Feb 19 '11 at 22:18
⚑

▲

4

▼

Short answer from someone who doesn't know much about Bayesian statistics:

You should read about "reference priors". Have a look here: http://arxiv.org/pdf/0904.0156

answered Feb 17 '11 at 21:04

passerby51
1,301   6   9

▲

The key point is to show that your analysis does not depend much on the prior in the first place.

You could try

and if you get answers that are reasonably close then that should provide some support that your analysis does not depend on arbitrary choice of prior.

BUGS and similar systems give quite a bit of flexibility regarding the prior but otherwise, from a practical viewpoint, easy computability and the availability of software may restrict your choice of prior.

See:

- http://en.wikipedia.org/wiki/Conjugate_prior
- http://en.wikipedia.org/wiki/Uniform_prior
- http://en.wikipedia.org/wiki/Jeffreys_prior

edited Feb 20 '11 at 17:04   answered Feb 20 '11 at 16:26

G. Grothendieck

31   2

---

1   You are not the first "Grothendieck" around... If it is not your real name, please see tea.mathoverflow.net/discussion/302/... – Mariano Suárez-Álvarez Feb 20 '11 at 16:48

Sometimes you have to pick the prior from a probability distribution, called a hyperprior. That itself can be drawn from a hyperhyperprior and so forth. Eventually you to just pick something: see Sunrise problem.

0

answered Feb 18 '11 at 3:59

11   1