International Journal of Mathematics and Statistics Invention (IJMSI)
ISSN (Online): 2321 – 4767, ISSN (Print): 2321 – 4759
www.ijmsi.org ||Volume 5 Issue 6 ||June. 2017 || PP.06-09

A Novel Bayes Factor for Inverse Model Selection Problem based on Inverse Reference Distribution

Debashis Chatterjee [1]
[1](Interdisciplinary Statistical Research Unit / Indian Statistical Institute, India)

Abstract : *Statistical model selection problem can be divided into two broad categories based on Forward and Inverse problem. Compared to a wealthy of literature available for Forward model selection, there are very few methods applicable for Inverse model selection context. In this article we propose a novel Bayes factor for model selection in Bayesian Inverse Problem context. The proposed Bayes Factor is specially designed for Inverse problem with the help of Inverse Reference Distribution (IRD). We will discuss our proposal from decision theoretic perspective.*

Keywords: *Bayes Factor, Model selection, Inverse problem, Inverse Reference Distribution*

I. INTRODUCTION

The question of choice of model is one of the central statistical problems. The main purpose of forward model selection problem is to infer about the model that actually generated the data, where as the inverse model selection primarily deals with the task of finding the most suitable model which has highest likelihood to generate the actual unobserved covariate or the unknown data generating function. Many approaches have been proposed in literature for proceeding with this central issue in the forward sense. The main objective of Inverse model selection is to infer about the "best" model so that one can infer about the unknown function which generates the data. Both frequentest and Bayesian schools have plenty of methods for model selection in forward problem. Broadly the methods fall into two categories (refer to [1]). One contains algorithms for picking a useful model (for example : exhaustive search, stepwise selection, backward and forward selection procedures, F tests for nested models etc.), other category contains criteria for judging the quality of a model (for example : AIC, BIC, Bayes Factor, cross validation etc.) There exists several methods in the statistical literature for model assessment and model selection for forward problem which purport themselves specifically as Bayesian predictive methods but the same for Inverse model problem is scarce. Indeed, Inverse model selection procedure available so far is not adequate in general in contrast to the wide range of practical necessity and applicability of statistical inverse problems.

II. INVERSE PROBLEM

Suppose one is interested in inferring about covariate $x \in \mathsf{A}$ from the given noisy observed response $y \in \mathsf{B}$, where A and B are Hilbert space. Let the relationship between x and y is governed by following equation (1) :

$$y = F(x) + \eta \quad (1)$$

where $F : \mathsf{A} \to \mathsf{B}$ is the forward operator from A to B and $\eta \in \mathsf{B}$ is the noise associated with the observation. In general, we allow F to be non-injective so that the standard approximate inversion of F may not serve the purpose. Moreover, except some statistical properties, η is unknown and there is no covenant that noisy observation y should fall in the image set of F. As stated in [2], for finite dimensional version of the above problem there might be situations when the number of equations is smaller than the number of unknowns. Modern statistical research is increasingly coming across the statistical inverse problems termed as "ill-posed" where the maximum likelihood solution or classical least square may not be uniquely defined at all and with very bad perturbation sensitivity of the classical solution.

III. NOTATIONS & PRELIMINARIES

We will use the notation $\pi(\cdot)$ to denote any suitable probability density without loss of generality. From decision theoretic perspective, let the unknown state of the world is given by the parameter $\theta \in \Theta$. Decision to an inference problem will be denoted by $a \in \mathsf{A}$ equipped with suitable utility function $u(a, \theta)$. For given data D and model M, we will denote beliefs about the unknown state of the world are described by the posterior distribution $\pi(\theta \mid D, M)$. Assume μ_0 to be the prior measure with Lebesgue density $\pi_0(x)$ for random variable X corresponding to the outcome $x \in \mathsf{A}$. We will assume joint distribution for (Y, X) and the conditional random variable $X \mid Y$ are well defined. $X \mid Y$ will be denoted by X_Y. Let us assume ε comes from distribution Q_0 where Q_0 is a measure on B such that $\eta \perp X$. Let the Lebesgue density for Q_0 be $\pi(\eta)$. Let also assume F to follow Q_F for some suitable measure Q_F.

IV. BAYES THEOREM FOR GENERAL INVERSE PROBLEM

We can generalize the Bayes theorem for inverse problem, which is much similar to [3] and [2] under the following assumption (1).

Assumption 1 *Let* $Z := \int\int \pi\left(y - F(x)\right) dQ_F \; d\mu_0(x) > 0$

Theorem 1 (Bayes Theorem for General Inverse Problem) *If $Q_X \ll Q_0$ for X a.s μ_0, then under assumption (1) the random variable $X \mid Y := X_Y$ is well defined with Lebesgue density $\pi_y(x)$ given by :*

$$\pi_y(x) = \frac{1}{Z}\pi_x(y)\cdot\pi_0(x) = \frac{1}{Z}\left(\int \pi\left(y - F(x)\right)dQ_F\right)\cdot\pi_0(x) \qquad (2)$$

Let us define Potential $\Phi(x, y)$ as follows:

$$\Phi(x, y) := -\log \pi_x(y) = -\log\left(\int \pi\left(y - F(x)\right)dQ_F\right). \qquad (3)$$

We can then rewrite Theorem (1) similar to [2] as follows:

$$\frac{dQ_Y}{dQ_0}(x) \propto \exp\left(-\Phi(x, y)\right) \qquad (4)$$

V. INVERSE REFERENCE DISTRIBUTION

Given a model M, a Bayesian model usually consists of a statistical model equipped with suitable probability density $\pi\left(y, x \mid \theta, M\right)$ for the observations with a prior distribution $\pi(\theta \mid M)$ for the model parameters (refer to [1]).

We can obtain the following conditional density:

$$\pi(y, \theta \mid x, M) = \pi(y \mid x, \theta, M)\cdot\pi(\theta \mid x, M). \qquad (5)$$

Conditioning on observed data D in the posterior distribution $\pi(\theta \mid D, M)$ for the model parameters, we can determine the posterior predictive distribution

$$\pi(\tilde{y} \mid \tilde{x}, D, M) = \int \pi(\tilde{y} \mid \tilde{x}, \theta, D, M)\cdot\pi(\theta \mid \tilde{x}, D, M)\, d\theta \qquad (6)$$

Suppose we have data set $D = \{x_i, y_i\}_{i=1}^{n}$ with a probability distribution $\pi(y_i \mid x_i, \theta)$ with θ as model parameter. It is of interest to assess a given model M based on the corresponding unknown covariate x. In [4] the author has proposed a bayesian model adequacy for such scenario. The method is based on Inverse Reference Distribution (IRD). Given a model with dataset $\{(y_i, x_i);\; i = 1,2, \ldots, n\}$ where y_i is the observed response and $X = \{x_1, x_2, \ldots, x_n\}$ is a set of (non-random) covariates, the approach is to obtain reference distributions based on the posterior $\pi(\tilde{X} \mid Y)$, where $\tilde{X}$ is interpreted as the unobserved random vector corresponding to the observed covariates X (refer to [4]). Let $L(Y, \tilde{X}, \theta)$ denote the Likelihood treating $(\tilde{X}, \theta)$ as unknown with the model parameter θ. Then the posterior $\pi(\tilde{X} \mid Y)$ can be obtained from equation (7).

$$\pi(\tilde{X} \mid Y) \propto \int \pi(\tilde{X}, \theta)\, L(Y, \tilde{X}, \theta)\, d\theta. \qquad (7)$$

The Inverse Reference Distribution given by equation (7). For model M it is given by equation (8).

$$\pi(\tilde{X} \mid Y, M) \propto \int \pi(\tilde{X}, \theta, M)\cdot L(Y, \tilde{X}, \theta, M)\, d\theta. \qquad (8)$$

VI. BAYES FACTOR FOR INVERSE MODEL SELECTION

For rest of the paper we will assume M closed view for inverse model selection, where by assumption it is possible to enumerate all possible model structures $\{M_i\}_{i=1}^{k}$ and the truth model as M_*. We can assign prior distribution $\pi(M_i)$ over the model set or specify a non-parametric model with a prior distribution on a suitable function space (refer to [1]). For a model M_i, we will denote the parameter space as Θ_i with the parameter $\theta_i \in \Theta_i$. The posterior probability of model M_i given the coordinate X is $\pi(M_i \mid X)$ given by

$$\pi(M_i \mid X) = \int \pi(M_i \mid X, Y) \cdot \pi(Y \mid X)\, dy \qquad (9)$$

We can define Bayes Factor for Inverse Model Selection as follows:

Definition 5.1 (Bayes Factor for Inverse Model Selection) *Bayes Factor of Model M_i to M_j is defined as:*

$$B_{i,j} = \frac{\pi(X \mid Y, M_i)}{\pi(X \mid Y, M_j)}. \qquad (10)$$

From Inverse Reference Distribution (IRD) given by equation (8) we can compute the Inverse Bayes Factor given by equation (10). Hence we can write equation (10) as follows:

$$B_{i,j} = \frac{\pi(X \mid Y, M_i)}{\pi(X \mid Y, M_j)} = \frac{\int_{\Theta_i} \pi(\tilde{X}, \theta_i, M_i) \cdot L(Y, \tilde{X}, \theta_i, M_i)\, d\theta_i}{\int_{\Theta_j} \pi(\tilde{X}, \theta_j, M_j) \cdot L(Y, \tilde{X}, \theta_j, M_j)\, d\theta_j} \qquad (11)$$

VII. DECISION THEORETIC PERSPECTIVE

We will now discuss the applicability of our proposed Bayes Factor for Inverse Model Selection context under two decision theoretic viewpoints : M - closed view and M - open view.

1) M - closed view

In the M-closed view it is possible to enumerate all possible model structures $\{M_i\}_{i=1}^{k}$. We can assign prior distribution $\pi(M_i)$ over the model set or specify a non-parametric model with a prior distribution on a suitable function space (refer to [1]). Here we assume that there exists a 'true' model from which the data are coming. For predictive Inverse model selection we suggest to find the utility $u(a,\theta)$ as

$$u(a,\theta) = \int u(a, \tilde{x}) \cdot \pi(\tilde{x} \mid \theta, M)\, d\tilde{x}. \qquad (12)$$

The task is to select the model which maximizes the utility.

In our context, let us define $u_i = \pi(X \mid Y, M_i)$ to be the utility of Model M_i. Then

$$B_{i,j} = \frac{\pi(X \mid Y, M_i)}{\pi(X \mid Y, M_j)} = \frac{u_i}{u_j}. \qquad (13)$$

As mentioned earlier, the main objective of forward model selection problem is to the model M that actually generated the data, where as the inverse model selection primarily deals with the task of finding the most suitable model M which has highest likelihood to generate the actual unobserved covariate X_* or function responsible for generating the data (outcome). Now if we can find the truth model M_* under M - closed view then it serves the both purpose then that should be most desirable answer to both problem, known as oracle solution. But In practical often there is no direct method to find the oracle solution. Notice that, the most suitable forward model may not be the most suitable inverse model unless we have found a oracle (true) model.

2) M - open view

In the M-open view we can not explicitly construct the actual belief model, as there is a strong background information about the belief that that any such model would not reflect well the properties of future data ([1]). In other words, we cannot simply infer about a single 'true' model governing the observed data. As state in [1], in this case, we assess the predictive performance of the candidate models under minimal modelling assumptions. Nevertheless, we can obtain pseudo Monte Carlo samples from it which leads to sample re-use methods like cross-validation (refer to [5]).

VIII. CONCLUSION

In this paper we have proposed some novel Bayes Factor specially designed for Inverse Model Selection. It is worth mentioning that Assumption (1) may not be always true. In fact it heavily depends on choice of prior. Cross validation based Inverse model adequacy is one of the very few alternative approach when the assumption (1) is not valid, or in other words when posterior is likely to become improper. Future research work of the author should address those issues.

ACKNOWLEDGEMENTS

This research work is funded by Indian Statistical Institute. The author would like to express gratitude to Prof. Sourabh Bhattacharya, Interdisciplinary Statistical Research Unit, Indian Statistical Institute for sharing his insight to make this research fruitful.

REFERENCES

[1]. Vehtari, Aki; Ojanen, Janne. A survey of Bayesian predictive methods for model assessment, selection and comparison. Statist. Surv. 6 (2012), 142–228. doi:10.1214/12-SS102. http://projecteuclid.org/euclid.ssu/1356628931.

[2]. Dashti, Masoumeh; Stuart, Andrew M. The Bayesian Approach To Inverse Problems. eprint: arXiv:1302.6989, 2015.

[3]. Vollmer, S. Posterior Consistency for Bayesian Inverse Problems Through Stability and Regression Results (2013), Inverse Problems, 29. Article number 125011.

[4]. Bhattacharya, S. . A Fully Bayesian Approach to Assessment of Model Adequacy in Inverse Problems (2013). Statistical Methodology, 12, 71-83.

[5]. Bernardo, J. M. and Smith, A. F. M. (1994). Bayesian Theory. John Wiley & Sons. MR1274699.