Bayes factor: A calculus of Evidence

Akshika Rastogi, Vineeta Verma

SVBPUA & T, Meerut, U.P, India

Abstract

1. Abstract

In the Bayesian analysis of the reliability characteristics, the parameter(s) representing reliability characteristics, in a life time distribution, follow random variations represented by prior(s). The prior(s) are updated with experimental data. However, it is felt that even the basic life time distribution can be updated in view of variations in its parameter. This basic distribution can be updated with the concept of compound distribution. Bayesian inference is usually presented as the method of determining how scientific belief should be modified by data. The Bayesian methodology has been one of the most active area of statistical development in the past 30 years. It is little understood that Bayesian methods have a data based core, which can be used as **Calculus of Evidence.** This core is called the Bayes factor.

In this paper we use Bayes factor for model selection. The bayes factor is one of the popular criteria in selecting models. The Bayes factor can be interpreted as the strength of evidence favoring one of the models for the given data. In fact, Bayes factor is strongly related to the posterior probabilities and hence its interpretation is straight forward. Jeffreys (1961) and Kass and Raftrey (1995) have proposed the rules to determine the strength of evidence which can be associated with the observed value of Bayes factor.

KEYWORDS- Bayesian, Bayes factor, evidence, probabilities, model selection

2. Introduction :

In the Bayesian analysis of reliability characteristics, the parameter (s), representing reliability characteristics, in a life time distributions follow random variations represented by prior (s). The priors are then updated with experimental data. However, it is felt that even the basic life time distribution can be updated in view of variations in its parameter. This basic distribution can be updated with the concept of compound distributions. Obviously, then we have two basic distributions representing entirely two different situations. It is in this light, the investigator proposes to develop statistical techniques for analyzing the robust character of tools in reliability/survival analysis when parameter (s) in the life time distributions are considered as random variable.

Bayesian inference is usually presented as a method of determining how scientific belief should be modified by data. The Bayesian methodology has been one of the most active areas of statistical development in the past 20 years. It is little understood that Bayesian methods have a data based core, which can be used as a calculus of evidence. This core is called the ***"BAYES FACTOR"*** which in its simplest form is also called a likehood ratio.

The Bayes factor is a popular criterion in Bayesian model selection. Let x denote the data and let M_1 and M2 be two competing models. Then the Bayes factor, k(x), is one of the most popular criteria in selecting which models M_1 and M2 fit the data better. The Bayes factor can be interpreted as the strength of evidence favoring one of the models for the given data. Infact, the Bayes factor is strongly related to the posterior probabilities and hence its interpretation is straight forward. Jeffreys (1961) and Kass and Raftery (1995) have proposed the rules to determine the strength of evidence which can be associated with the observed value of the Bayes factor.

On the other hand, before the data are taken, the Bayes factor, k(x), is a random variable and its distribution follow that of x. Then the sampling properties of the Bayes factor can be examined under each of the two models under considerations. Once the data X = x are obtained, these properties can be used to measure the agreement between each of the two models and the data. According to this measure a decision can be made about the goodness of fit of each of models. Thus, it is not appropriate to treat the Bayes factor as fixed once the data is observed due to uncertainty in the random sample. In particular, when the amount of uncertainity is large, the data, which are observed at the different time points but from the same distribution model, could lead to very different observed values of Bayes factor. It is well known that the Bayes factor depends on the choice of the prior distribution and, in particular, it is extremely sensitive to imprecise priors.

3. Notation

x	data under consideration
M_1	Model M_i
M2	Model M2
k(x)	Bayes factor
$P_i(x/\theta_i)$	Probability function for i = 1, 2
$(\theta_i\ )$	Prior distribution for i = 1,
m, (x)	Prior predective distribution for i = 1, 2

4. Comparision of two models through Bayes Factor

Assume that we are interested in comparing two models M_1 and M2, say, as convenient statistical representation of some data x. Let $P_i\ (x/\theta_i)$ and $\pi_i\ (\theta_i)$ denote the probability function and the (proper) prior distribution, respectively, under model M_i for i = 1, 2. Also, let M_i (x) denote the prior predictive distribution under model M_i, that is,

$$m_i,(x)= S\int P_i\ (x/\theta_i)\ \pi_i(\theta_i)d\ \theta \qquad = \qquad 1,2$$

In the context of hypothesis testing the selection between these two models can be expressed as

H1 : M_1 is true v/s H2 : M2 is true.

Given *the observed data, x, the Bayes factor for M_1 against M2 is*

$$k(x)=\frac{m_1(x)}{m_2(x)}$$

Due to recent computational advances, sophisticated techniques for computing Bayes factors have been developed. For example, Kass and Vaidyanathan (1992), Chib (1995), Meng and Wong (1996), Chen and Shao (1997), Dicicco, Kass and Raftery and Wasserman (1997), Gelmon and Meng (1998) and Chib and Jeliazkov (2001). If k (x) > 1, then x is best predicated by M1 and consequently, this model is preferred. Ofcourse, using similar arguments, K(x) < 1 gives support to model M2. Several different ways has been proposed to interpret the strength of evidence according to k(x). Jeffreys (1961) proposed the rule given in Table 1 and, more recently, Kass and Raftery (1995) proposed a slight modification of Jeffreys' proposal as shown in Table 2.

Although these rules are given in terms of evidence against M2, the same rule (at least in principle) can be used to interpret the evidence against M_1, by inverting the value of Bayes factor.

Table 1. Jeffreys' Scale of evidence

K	Evidence against M_2
1 : 3.2	Not worth more than a bare mention
3.2 : 10	Substanital
10 : 100	Strong
100 : ∞	Decisive

Table 2. Scale of evidence proposed by Kass and Raftery (1995)

K	Evidence against M_2
1 : 3	Not worth more than a bare mention
3 : 20	**Positive**
20 : 150	Strong
150 : ∞	Very strong

4 Statistical Background

5. **Bayes factor through prior and posterior distribution in the analysis:**

Suppose we have a random variable which produces either a success or a failure, we want to compare a model M_i where the probability of success is **q =** 1/2 and another model l's/1) where q is completely unknown and we take a prior distribution for q which is uniform on [0, 1] with p.d.f.

$$f(x; a, b) = 1 \quad ; \quad 0 < X < 1 \quad ; \quad a, b \geq 0$$

We take a sample of size 25 and find 15 success and 10 failures. Then, the likelihood function or probability of observing 15success is:

$$\binom{25}{15} q^{15}(1-q)^{10} \qquad (1)$$

So, for q=1/2, we have

$$P[x = 15/M_1] = \binom{25}{15}\left(\frac{1}{2}\right)^{25}$$

$$=0.0974166 \qquad (2)$$

Let us all this situations as case I.

Case II : Bayes factor when prior is used in the analysis

Now we use the prior distribution in case I to define the likelihood function in the case of model M_2.

$$P[x = 15/M_2] \qquad = \int_{q=0}^{1}\binom{25}{15} q^{15}(1-q)^{10}dq$$

$$=\frac{1.5512\times 10^{25}}{4.033\times 10^{26}}$$

$$= 0.03846 \qquad (3)$$

In view of (2) and (3), the ratio or bayes factor is then

K=2.5329

Which is 'barely worth mentioning' even if its point very slightly toward M_1.

Case III : Bayes factor when posterior is used in analysis along with case I

After observing the experimental data x = $\{x_1 \dots \dots \dots x_n\}$

Now,

The posterior distribution of q is

$$\pi(q/\underset{\sim}{x}) = \frac{L(x_1 \dots \dots \dots x_n)g(q)dq}{\int_{\Omega} L(x_1 \dots \dots \dots x_n)g(q)dq}$$

$$= \frac{\binom{25}{15}q^{15}(1-q)^{10}\cdot 1}{\int_0^1 \binom{25}{15}q^{15}(1-q)^{10}\cdot 1\, dq}$$

$$= \frac{1}{\beta(16,11)} q^{16-1}(1-q)^{11-1} \qquad (4)$$

Then, we use the posterior distribution in (4) to define the likelihood function in the case of model M_2.

$$P[x = 15\big/M_2] = \int_{q=0}^{1}\binom{25}{15} q^{15}(1-q)^{10}\frac{1}{\beta(16,11)}\, dq$$

$$= \frac{8509236.92}{6665292.22}\, q^{16-1}(1-q)^{11-1}$$

$$= 0.115579 \qquad (5)$$

The view of (2) and (5), thye ratio orbayes factor is then,

K= 0.8427

4.2 Other Stattistical Measure for the model:

(i) For Model M_1:

$$Mean = np \Rightarrow 12.5$$

$$Variance = npq \Rightarrow 6.25$$

And $C.V. = \frac{\sigma}{\bar{x}} \times 100 \Rightarrow 20$

(ii) For Model M_2:

$$Mean = \frac{b+a}{2} \Rightarrow 0.50$$

$$Variance = \frac{1}{12}(b-a)^2 \Rightarrow 0.0833$$

$$C.V. = \frac{\sigma}{\bar{x}} \times 100 \Rightarrow 57.72$$

6. Bayes factor through prior and posterior distribution in the analysis: Binomal with Beta distribution of first kind on [1,1]

$$\mathrm{f}\,(\mathrm{x};\mathrm{a},\mathrm{b}) = \frac{1}{\beta(\mathrm{a},\mathrm{b})}\mathrm{x}^{\mathrm{a}-1}(1-\mathrm{x})^{\mathrm{b}-1}; 0 < x < 1; a, b \geq 0$$

We take a sample of size 25 and find 15 success and 10 failures. Then, the likelihood function or probability of observing 15success is:

$$\binom{25}{15} q^{15}(1-q)^{10} \qquad (1)$$

So, for q=1/2, we have

$$P[x = 15/M_1] = \binom{25}{15}\left(\frac{1}{2}\right)^{25}$$

$$=0.0974166 \qquad (2)$$

Let us call this situations as case I.

Case II : Bayes factor when prior is used in the analysis

Now we use the prior distribution in case I to define the likelihood function in the case of model M_2.

$$P[x = 15/M_2] = \int_{\mathrm{q}=0}^{1}\binom{25}{15}\,\mathrm{q}^{15}\,\frac{1}{\beta(1,1)}\,\mathrm{q}^{1-1}(1-\mathrm{q})^{1-1}\mathrm{dq}$$

$$= 0.03846 \qquad (3)$$

In view of (2) and (3), the ratio or bayes factor is then

K=2.5329

Which is 'barely worth mentioning' even if its point very slightly toward M_1.

Case III : Bayes factor when posterior is used in analysis along with case I

After observing the experimental data x = $\{x_1 \ldots \ldots \ldots x_n\}$

Now,

The posterior distribution of q is

$$\pi(q/\underset{\sim}{x}) = \frac{\mathrm{L}(\mathrm{x}_1\mathrm{x}_2 \ldots \ldots \ldots \mathrm{x}_n).\mathrm{g}(\mathrm{q})\mathrm{dq}}{\int_{\Omega} \mathrm{L}(\mathrm{x}_1\mathrm{x}_2 \ldots \ldots \ldots \mathrm{x}_n).\mathrm{g}(\mathrm{q})\mathrm{dq}}$$

$$= \frac{\binom{25}{15} q^{15}(1-q)^{10} \frac{1}{\beta(1,1)} q^{1-1}(1-q)^{1-1}}{\int_0^1 \binom{25}{15} q^{15}(1-q)^{10} \cdot \frac{1}{\beta(1,1)} q^{1-1}(1-q)^{1-1} 1\, dq}$$

$$= \frac{1}{\beta(16,11)} q^{16-1}(1-q)^{11-1} \qquad (4)$$

Then, we use the posterior distribution in (4) to define the likelihood function in the case of model M_2.

$$P[x = 15 / M_2] = \int_{q=0}^{1} \binom{25}{15} q^{15} (1-q)^{10} \frac{1}{\beta(16,11)} dq$$

$$= \frac{8509236.92}{6665292.22} q^{16-1}(1-q)^{11-1}$$

$$= 0.115579 \qquad (5)$$

The view of (2) and (5), the ratio or bayes factor is then,

K= 0.8427

7. Conclusion-

The Bayes factor is a comparison of how well two hypothesis predict the data. The hypothesis that predicts the observed data better is the one that is said to have more evidence supporting it. So it is with bayes factor that modifies prior probabilities , and after seeinghow much Bayes factor of certain size change various prior probabilities, we begin to understand what represent strong evidence and weak evidence.

Bayes factor is used to Study the impact of experimental data on model. In other words, this also reflect the robust character of evidence towards the two models in respect of data information.

References-

- https://en.wikipedia.org/wiki/Bayes_factor
- Berger, J. O. and Sellke, T. (1987). Testing a point null hypothesis: Irreconcilability of P values and evidence (with discussion), Journal of the American Statistical Association 82: 112–139.

- Edwards, W., Lindman, H. and Savage, L. J. (1963). Bayesian statistical inference in psychological research, Psychological Review 70: 193–242.
- Held, L. (2010). A nomogram for P values, BMC Med Res Methodol 10: 21. Madigan, D. and York, J. (1995). Bayesian graphical models for discrete data, International Statistical Review 63(2): 215–232.
- Saban´es Bov´e, D. and Held, L. (2011b). Hyper-g priors for generalized linear models, Bayesian Analysis. Forthcoming article as of 18/2/2011. Available from: http://ba.stat.cmu.edu/abstracts/Sabanes.php.
- Sauerbrei, W. and Royston, P. (1999). Building multivariable prognostic and diagnostic models: transformation of the predictors by using fractional polynomials, Journal of the Royal Statistical Society. Series A (Statistics in Society) 162(1): 71–94. Scott, J. G. and Berger, J. O. (2010). Bayes and empirical-Bayes multiplicity adjustment in the variable-selection problem, Annals of Statistics 38(5): 2587–2619.
- Sellke, T., Bayarri, M. J. and Berger, J. O. (2001). Calibration of p values for testing precise null hypotheses, The American Statistician 55: 62–71. West, M. (1985).
- Carlin C,Louis T. Bayes and Empirical Bayes Methods for Data Analysis. London: Chapman and Hall ;1996.
- Howard J. The 232 table :a discussion from a Bayesian viewpoint. Statistical Science. 1999;13:351-67.
- Kass R, Raftrey A. Bayes Factors. Journal of the American Statistical Association. 1995;90:773-95.