University of Texas, MD Anderson Cancer Center

UT MD Anderson Cancer Center Department of Biostatistics Working Paper Series

<i>Year</i> 2005	Paper 6
•	-

A Note on the Consistency and Interpretation of Bayes Factors Based

Valen Johnson*

*vejohnson@mdanderson.org

This working paper is hosted by The Berkeley Electronic Press (bepress) and may not be commercially reproduced without the permission of the copyright holder.

http://biostats.bepress.com/mdandersonbiostat/paper6

Copyright ©2005 by the author.

A Note on the Consistency and Interpretation of Bayes Factors Based on Test Statistics

Valen E. Johnson

A method for defining Bayes factors based on the sampling distribution of test statistics was proposed in Johnson (2005). To implement this method, it is necessary to specify the sampling distribution of test statistics under alternative hypotheses. For Bayes factors based on χ^2 and F statistics, these distributions can be naturally defined as noncentral versions of the null distribution. In this note, I describe criteria for setting hyperparameters that determine these noncentral distributions so that the resulting Bayes factors achieve consistency.

Let BF(1|2) denote the Bayes factors between models 1 and 2, i.e. the ratio of the marginal density of the data under model 1 to the marginal density of the data under model 2. For present purposes, BF(1|2) will be called *consistent* if (a) $BF(1|2) \xrightarrow{p} \infty$ as the sample size $n \to \infty$ when model 1 is true, and (b) $BF(1|2) \xrightarrow{p} 0$ as $n \to \infty$ when model 2 is true.

For the remainder of this note, 'J5" refers to Johnson (2005) and, unless otherwise stated, notation and regularity conditions stated in J5 apply here also.

 χ^2 tests for multinomial data. Let **p** denote a multinomial probability vector which satisfies a given null hypothesis, and suppose that under the alternative hypothesis the multinomial probability vector **q** is drawn from a Dirichlet distribution with parameter c**p**. Letting K - s - 1 denote the degrees of freedom of the χ^2 statistic x_n (as defined in Sec. 2 of J5), the logarithm of the Bayes factor between the alternative hypothesis (model 2) and null hypothesis (model 1) may be written

$$\left[\frac{nx_n}{2(1+c+n)}\right] + \left(\frac{K-s-1}{2}\right)\log\left(\frac{1+c}{1+c+n}\right).$$
(1)

Under the null model, the distribution of x_n is χ^2_{K-s-1} . In this case, the first term of (1) is bounded in probability, while the second tends to $-\infty$ as $n \to \infty$. Thus, $BF(2|1) \xrightarrow{p} -\infty$ under the null model.

Under the alternative hypothesis, the distribution of x_n is noncentral $\chi^2_{K-s-1}(\lambda)$ with noncentrality parameter $\lambda = n\kappa'\kappa$, $\kappa = \{(q_i - p_i)/\sqrt{p_i}\}$. If this hypothesis pertains, the first term in (1) dominates and $BF(2|1) \xrightarrow{p} \infty$ as $n \to \infty$. It follows that the Bayes factor is consistent for constant c. Note that if marginal maximum likelihood estimation (MMLE) is used to estimate $\alpha = (c+1)/n$, the resulting Bayes factor tends to ∞ under the alternative hypothesis, but remains bounded under the null hypothesis and so is not consistent.

It is worth noting that Lemmas 1–4 of J5 assume that $(\mathbf{q} - \mathbf{p}) = O_p(1\sqrt{n})$. Under the alternative hypothesis, this assumption is required for x_n to converge to a noncentral χ^2 distribution as $n \to \infty$. For $(\mathbf{q} - \mathbf{p}) = O_p(1)$ (the case of finite samples), the

Collection of Biostatistics Research Archive distribution of x_n under the alternative model falls between $\min_i \{q_i/p_i\}$ and $\max_i \{q_i/p_i\}$ times the noncentral χ^2 distribution that would be obtained by replacing $\sqrt{np_i}$ by $\sqrt{nq_i}$ in the denominator of the normal deviates used in the definition of x_n . Thus, provided that the values of q_i/p_i are not too different from 1, the non-central χ^2 approximation to the distribution of the test statistic under the alternative model will often be adequate.

F tests for linear models. Suppose that

$$\mathbf{y} \mid \boldsymbol{\beta}, \sigma^2 \sim N_n(\mathbf{X}\boldsymbol{\beta}, \sigma^2 \mathbf{I}),$$

where \mathbf{y} is an $n \times 1$ observation vector, $\boldsymbol{\beta}$ is an $r \times 1$ regression parameter, \mathbf{X} is an $n \times r$ matrix of rank r, and σ^2 is a scalar variance parameter. Assume further that under the null hypothesis, $\mathbf{H}'\boldsymbol{\beta} = \boldsymbol{\xi}$, where \mathbf{H} is an $r \times k$ matrix of rank k whose range space is contained in the range space of \mathbf{X}' , and let f_n denote the standard F statistic for testing the null hypothesis against the alternative that $\boldsymbol{\beta}$ does not satisfy this constraint. Under the alternative hypothesis, if $\boldsymbol{\beta}$ is drawn from an r-variate normal distribution centered on a value that does satisfy the null constraint and having covariance matrix $\tau \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$, then the logarithm of the Bayes factor in favor of the alternative, say $\log(BF(2|1))$, can be written

$$-\frac{K}{2}\log(1+n\tau^*) + \frac{k+m}{2}\log\left(1+\frac{kf_n}{m}\right) - \frac{k+m}{2}\log\left(1+\frac{kf_n}{m(1+n\tau^*)}\right), \quad (2)$$

where m = n - r and $n\tau^* = \tau$.

Under the null hypothesis, $f_n = O_p(1)$. This implies that the first term in (2) dominates the sum, so that $BF(2|1) \xrightarrow{p} -\infty$ as $n \to \infty$.

Under the alternative hypothesis, $f_n/(1 + n\tau^*)$ has a $F_{k,m}$ distribution. Consequently, $f/m = O_p(1)$ and f/m > 0 with probability 1. The second term in (2) is thus linear in m; the remaining terms are $O_p(\log(n))$ or less. It follows that $BF(2|1) \xrightarrow{p} \infty$ under the alternative hypothesis. Therefore, the Bayes factor based on the F statistic is consistent for fixed values of τ^* (but not for fixed values of τ).

Stomach cancer data revisited. White and Eisenberg (1959) provided a cross-classification of stomach cancer site with blood type for 707 cancer patients (Table 1). The purpose of their study was to determine whether there was an association between blood type and cancer site. The χ^2 test for independence for these data is 12.65 on 6 degrees of freedom.

Because White and Eisenberg did not specify an alternative model for these data, it is not clear what value of c should be used to define the distribution of the χ^2 test statistic under the alternative hypothesis. Without an explicit alternative hypothesis, this difficulty can be circumvented by reporting test results for a range of alternative models. This strategy is particularly appealing if the "weight of evidence" criteria suggested in Kass and Raftery (1995) are used. According to their scheme (which represents a variation on criteria proposed by Jeffreys (1961)), relative evidence in favor of one of the tested

Collection of Biostatistics Research Archive

 $\mathbf{2}$

	Blood Group		
Site	0	А	B or AB
Pylorus and antrum	104	140	52
Body and fundus	116	117	52
Cardia	28	39	11
Extensive	28	12	8

Table 1: White and Eisenberg's classification of cancer patients

hypotheses is classified according to the value of twice the natural logarithm of Bayes factors. Based on this value, experimental evidence can be classified as "not worth more than a bare mention" (0-2), "positive" (2-6), "strong" (6-10) or "very strong" (> 10).

Using these classifications, the Bayes factor based on the χ^2 statistic suggests that White and Eisenberg's data provide (a) very strong evidence against alternative hypotheses generated from values of c in (0,16.5); (b) strong evidence against alternatives generated from values of c in (16.5,35.9); (c) positive evidence against alternatives generated from values of c in (35.9,86.0); and (d) evidence not worth mentioning for alternative models generated from values of c in (86,412) or values of c > 1050. There is positive evidence for alternative models generated with c in the range (412,1050), and the maximum evidence in favor of the alternative hypothesis occurs when c=636. These domains of support are illustrated in Fig. 1.

When c = 636, the Bayes factor in favor of the alternative hypothesis is slightly less than 3 (twice the log of the Bayes factor is 2.17). Standard deviations of cell probabilities generated from this alternative hypothesis are $\sqrt{p_i(1-p_i)}/25$, where $\{p_i\}$ denotes a probability vector satisfying the independence assumption. Such deviations ($\approx 4\%$) may or may not be regarded as substantively important.

References

- [1] Jeffreys, H. (1961), Theory of Probability, London: Oxford University Press.
- [2] Johnson, V.E. (2005), "Bayes Factors Based on Test Statistics", to appear in the Journal of the Royal Statistical Society, Series B.
- [3] Kass, R.E. and Raftery, A.E. (1995), "Bayes Factors," Journal of the American Statistical Association, 90, 773–795.
- [4] White, C. and Eisenberg, H. (1959) "ABO blood groups and cancer of the stomach," Yale Journal of Biology and Medicine, 32, 58-61.



3



Figure 1: Domains of relative support for null (independence) and alternative hypotheses for White and Eisenberg data.

COBRA A BEPRESS REPOSITORY Collection of Biostatistics Research Archive