The Traditional Behrens-Fisher Problem from the Bayesian Perspective Applied to the Alabama Sturgeon Controversy

Alfred A. Bartolucci[1], Paul D. Blanchard[2],
W. Mike Howell[2] and Karan P. Singh[1]

[1]Department of Biostatistics, University of Alabama at Birmingham
[2]Department of Biology, Samford University
Birmingham, Alabama, U.S.A.

Summary The raw data from a sample of 75 sturgeon (45 listed as shovelnose and 30 described as Alabama sturgeon) were examined to determine if there were two distinct species in this sample. Earlier work by the authors (AAB, PDB, WMH) concluded that the Alabama and shovelnose were indistinguishable based on principal component analyses of the meristic and mensural data. In our present work we perform formal Bayesian sharp hypothesis testing on the distinct principal components from the two groups (Alabama vs. Shovelnose). The question of inference on the difference of two normal means when both the variances are unknown is the celebrated Behrens-Fisher Problem. This methodology applied in the Bayesian context leads to coherent inference on the sharp null hypothesis for the difference of the two component means. Realistic prior parameters are engaged in conjugate normal prior distributions which when combined with the sampling distributions lead to the usual prior to posterior constructs. We further derive the Bayes Factor or Weighted Likelihood Ratio (WLR) for inference on our sharp null hypotheses. The evaluation of the WLR for the prior structure requires at most one dimensional quadratures which are easily computed because the integrands are products of known density functions with familiar properties.

1. Introduction

Inference on two normal means when the variances are unknown is the well known Behrens-Fisher problem. This was first treated in the Bayesian context by Jeffreys (1940) who derived Fisher's "fiducial" distribution for the difference of two means as a posterior distribution. Lindley (1965) followed Jeffrey's treatment with the use of the joint "complete ignorance" prior density for the parameters of the two normal distributions with focus on the tail area of the continuous-type posterior density. Dayal and Dickey (1976) treated the Behrens-Fisher problem in the context of the weighted likelihood ratio or Bayes factor. Their treatise dealt with the global subject of "coherent inference for the Behrens-Fisher problems". It is the methodology of Dayal and Dickey (1976) that we pursue in this paper. Our application is to the Alabama Sturgeon Controversy in which we test the equality of means on a statistically reduced data set of mensural characteristics from two alleged distinct species of sturgeon, i.e., the shovelnose sturgeon and the Alabama sturgeon. Our approach is via the Bayes factor method. We first present the test and results for equality of means assigning integrable prior densities to the parameters of the two normal distributions. Our second approach is to test the complete equality of the two populations, i.e., equal means and variances relying more heavily on non-integrable prior constant densities for the difference of the mean parameters. Our results are applied to the actual sample of mensural data.

1.1 Statistical Methods

The traditional Behrens-Fisher problem is one of inferring from two independent normal samples with $Data = (X_{11}, X_{12}, \ldots, X_{1n_1}; X_{21}, X_{22}, \ldots X_{2n_2})$. The means are denoted by μ_1 and μ_2 with variances σ_1^2 and σ_2^2, respectively. Our goal is to present Bayesian inference as developed by Dayal and Dickey (1976) for the two normal means and realistic tractable forms of the prior density which we denote by $p(\mu_1, \mu_2, \sigma_1, \sigma_2)$.

We briefly review necessary formulae and concepts from the theory of Bayes factors for a sharp hypothesis (single dimension reduction). For the problem of two normal populations, define $\eta = \mu_1 - \mu_2$ and let ζ denote the nuisance parameters $\zeta = (\mu_1, \sigma_1, \sigma_2)$. The transformation to η and ζ from μ_1 and μ_2 is simple with unit Jacobian. Then the Bayes factor $B(H_0)$ for the hypothesis H_0: $\eta = \eta_0$ versus the alternative H_1: $\eta \neq \eta_0$ (prior probabilities $0 < P(H_0) < 1$ and $P(H_1) = 1 - P(H_0)$) is given by

$$B(H_0) = l(\eta_0) \Big/ \int_{-\infty}^{\infty} l(\eta)\, p(\eta|H_1)\, d\eta = \frac{p(Data|H_0)}{p(Data|H_1)} \qquad (1)$$

where $l(\eta)$ is the marginal (prior integrated) likelihood of η and $p(\eta|H_1)$ is the prior density of η on H_1 induced by the prior of (η, ζ) on the whole parameter space. Note that the Bayes factor does not depend on $P(H_0)$ or $P(H_1)$. Also, note the posterior odds for H_0 vs. H_1 can be obtained by multiplying $B(H_0)$ by the prior odds. It has been shown (Dickey and Lientz, 1970) that $B(H_0)$ can also be obtained as Savage's density ratio

$$B(H_0) = \lim_{\eta \to \eta_0} p(\eta|H_1; Data) \,/\, p(\eta|H_1) \qquad (2)$$

provided $l(\eta)$ is continuous at η_0. We note that the Bayes factors discussed here provide an approximation to a neighborhood hypothesis Bayes factor by a sharp hypothesis Bayes factor. The neighborhood null hypothesis takes the form

$$\eta^c = (\mu_1^c - \mu_2^c)/c. \tag{3}$$

for small δ.

Recall that $\eta = \mu_1 - \mu_2$ or the difference of two normal means. One could also consider η to be equal to the ratio μ_1/μ_2. As it stands the difference and ratio are special cases of a major general discrepancy measure.

One can show that the expression (3) becomes $\mu_1 - \mu_2$ as $c \to 1$ and that (3) is equal to μ_1/μ_2 as $c \to 0$. The Bayes factors for η^c as $c \to 1$ and $c \to 0$ are not equal. As an example of inference on ratio parameters, see Bartolucci and Dickey (1977) or Bartolucci and Singh (1993).

The joint likelihood of the four parameters for our two independently sampled normal distributions is:

$$l(\mu_1, \mu_2, \sigma_1, \sigma_2) \propto l(\mu_1, \sigma_1) l(\mu_2, \sigma_2)$$

where in general,

$$l(\mu, \sigma) \propto (\sigma^2)^{\frac{n}{2}} \exp\left[-\frac{1}{2}\sigma^{-2}(n(\mu - m)^2 + \nu s^2)\right], \tag{4}$$

$nm = x_1 + x_2 + ... + x_n$, $\nu s^2 = (x_1 - m)^2 + ... + (x_n - m)^2$, $\nu = n - 1$. Specifically, for each sample the n, m, ν, and s^2 are subscripted by 1 or 2 respectively.

The familiar traditional distribution of the sufficient statistics m and s^2 depends on the sample size n and thus holds for samples of fixed size. The likelihood (4) applies for any non-informative stopping rule to yield a single Bayesian inference. For the purposes of Bayesian inference from real data, we may assume that m and s^2 have their traditional fixed size sampling distribution.

We further define $\phi_\nu(\cdot\,;\,\cdot)$ and $\phi_{\nu_1,\nu_2}(\cdot\,,\,\cdot\,,\,\cdot)$ the t-density and Behrens-Fisher density as follows:

$$\phi_\nu(x; s^2) = s^{-1}\left[\nu^{\frac{1}{2}} Beta\left(\frac{\nu}{2}, \frac{1}{2}\right)\right]^{-1} \left(1 + \nu^{-1}(x/s)^2\right)^{-\frac{1}{2}(\nu+1)}$$

$$\phi_{\nu_1,\nu_2}\left(x, s_1^2, s_2^2\right) = \int_{-\infty}^{\infty} \phi_{\nu_1}\left(x - z, s_1^2\right) \phi_{\nu_2}(z, s_2^2)\, dz \tag{5}$$

in which Beta(a, b) denotes the complete beta integral $\int_0^1 \mu^{a-1}(1-\mu)^{b-1}\, d\mu$.

Let us now assign realistic integrable prior densities to $(\mu_1, \mu_2, \sigma_1, \sigma_2)$ all four independent given H_1 each of σ_1^2 and σ_2^2 distributed as $\tau g^2/\chi_\tau^2$ and each of μ_1 and μ_2 having density $p(\mu|H_1) = \phi_{\nu_0}(\mu - m_0, s_0^2)$.

Note that $p(\mu|H_1)$ is normal for $\nu_0 = \infty$. Let

$$\begin{aligned} \beta &= \nu_1 + \tau \\ \gamma &= \nu_2 + \tau \\ \varepsilon &= (\nu_1 s_1^2 + \tau g^2)/\beta \\ \theta &= (\nu_2 s_2^2 + \tau g^2)/\gamma, \end{aligned}$$

then $B(H_0)$ for $H_0 = \mu_1 - \mu_2 = 0$ is

$$\begin{aligned} B(H_0) = {} & \int_{-\infty}^{\infty} \phi_\beta(y_1 - m_1;\, \varepsilon/n_1)\, \phi_\gamma(y - m_2;\, \theta/n_2)\, \phi_{\nu_0}(y - m_0;\, s_0^2)\, dy \\ & /\phi_{\beta,\nu_0}\left(m_1 - m_0;\, \varepsilon/n_1;\, s_0^2\right) \phi_{\gamma,\nu_0}\left(m_2 - m_0;\, \theta/n_2;\, s_0^2\right) \\ & \cdot \phi_{\nu_0,\nu_0}\left(0;\, s_0^2,\, s_0^2\right). \end{aligned} \tag{6}$$

Since the integrand consists of known densities with familiar properties it is easily computed numerically. The numerator is a convolution of t-densities and the denominator is a product of Behrens-Fisher densities.

Dayal and Dickey (1976) also considered the question of both parameters shared versus non-shared. That is to say, we can consider two normal distributions of being nearly the same distribution. Conceivably, we can infer that they share the same mean and variance, i.e. define the hypothesis

$$\begin{aligned} K_0&: \ \eta = 0 \ \ and \ \ \rho = 1 \\ where&\ \ \rho = \sigma_1/\sigma_2 \ \ and \\ K_1&: \ \eta \neq 0 \ \ and \ \ \rho \neq 1. \end{aligned}$$

We consider the μ_1 and μ_2 as having approximate prior constant density which we denote as

$$\begin{aligned} p(\mu_1, \mu_2|H_1) &= const_\eta\, const_\zeta \\ &= p(\eta|H_1) p(\zeta|H_1), \end{aligned}$$

and σ_1 and σ_2 with prior density as before, i.e. $\tau g^2/\chi_\tau^2$.

Define

$$\begin{aligned} n_b &= n_1 + n_2 \\ \nu_b &= n_1 + n_2 - 1 \\ D &= \beta\varepsilon + n_1 n_2 n_b^{-1}(m_1 - m_2)^2 \end{aligned}$$

Then the Bayes factor for testing K_0 versus K_1 is

$$\beta(K_0) = \frac{1}{const_\eta} \frac{k(2\tau,\, 2\tau g^2)}{k(\nu_b + 2\tau,\, D + 2\tau g^2)} \Big/ \frac{k(\tau,\, \tau g^2)}{k(\beta,\, \beta\varepsilon) k(\gamma,\, \gamma\theta)} \tag{7}$$

where k(a, c) is the normalizing constant of the density

a/χ_c^2

or

$$k(a,\ c) = \left(\frac{c}{2}\right)^{\frac{a}{2}} \Gamma\left(\frac{a}{2}\right)$$

1.2 Results

Of the 75 sturgeon in the sample, 45 were listed as Shovelnose (S) and 30 were labeled as Alabama sturgeon (A). There were 35 mensural variables in the data set. Four principal components explained approximately 60% of the variation in the data. The analysis was geared to compare the mean of the reduced data in A to that of S. Table 1 summarizes the data. The overall classical significance level is around 0.08 indicating no statistical difference between the two samples. We wish to conduct our Bayesian analysis which shows even less extreme coherent conclusions.

Table 2 through 4 give the Bayes Factor for testing the equality of means using equation (6). Realistic values of τ and g vary from 5 to 15 and 1 to 2, respectively. These are given in the left margin with a fixed value of m_0 equal to 1.0. The value of s_0 and v_0 vary over a realistic range for subjective prior assessment of these parameters. See Birch and Bartolucci (1983). The values of the Bayes Factors are fairly robust within τ and g categories for varying s_0 and v_0. Note that if the prior odds were even, the posterior odds or Bayes Factor would still show evidence for H_0. This is further evidenced by Figure 1 which is the standardized posterior region over the parameter space for $\eta = \mu_1 - \mu_2$.

In Table 5 are given the WLR for making inference concerning the complete equality of the two populations (equal means and variances). Values are given her for τ ranging from 1 to 5 and g ranging from 1 to 2. The constant in equation (7) is given a wide range of about 5000. this is truly a non-informative prior. In any case, the Bayes Factor shows very strong evidence for K_0 over K, confirming further the equality of the two sturgeon populations with respect to the reduced mensural data.

1.3 Conclusion

The results outlined herein reinforce the conclusion of Howell, Blanchard and Bartolucci (1994). They established that the sturgeon samples, A and S, were indistinguishable species using the principal component analyses. We had the opportunity here to apply the Behrens-Fisher WLR to establish the same. This data reduction technique combined with the Bayesian analysis allows one to avoid inference based solely on single tail area probabilities. One has the ability to examine inferences based on reasonable functions of the parameters in the prior parameter space.

1.4 Acknowledgements

The authors wish to thank Ms. Rita Reid for her time and patience in the typing of this manuscript.

1.5 References

Bartolucci, A.A. and Dickey, J.M. Comparative Bayesian and traditional inference for gamma modeled survival data. *Biometrics*, 33(2), 343-354, 1977.

Bartolucci, A.A. and Singh, K.P. A Bayesian approach to bioequivalence in environmental studies. International Congress on Modelling and Simulation. *Proceedings*, 4, 1821-1830, 1993.

Birch, R. and Bartolucci, A.A. Determination of the hyperparameters of a prior probability model in survival analysis. *Computer Programs in Biomedicine*, 17, 89-94, 1983.

Dayal, H.H. and Dickey, J.M. Bayes factors for Behrens-Fisher problems. *Sankhya, The Indian Journal of Statistics*, 38, Series B, 315-328, 1976.

Dickey, J.M. and Lientz, B.P. The weighted likelihood ratio, sharp hypotheses about chances, the order of a Markov Chain. *Ann. Math. Stat.*, 41, 214-226.

Howell, W.M., Blanchard, P.D. and Bartolucci, A.A. A critique of the written comments submitted by R.L. Mayden and B.R. Kuhajdu to the U.S. Fish and Wildlife Service entitled "Re-evaluation of the taxonomic systematic and conservation status of the Alabama sturgeon, Scaphirhynchus Suttkussi, Williams and Clenmer (Actinopterygii, Acipenseridae)" unpublished report, 1994.

Jeffreys, Harold. Note on the Behrens-Fisher formula. *Ann. Eugen.*,10, 48-51, 1940.

Lindley, D.V. Introduction to Probability and Statistics from a Bayesian Viewpoint: Part Two Inference. Cambridge University Press, Cambridge, 1965.

1.6 Tables

Sturgeon	Mean	Standard Deviation	P-Value
A	-0.4278 (m_1)	1.0841(s_1)	.0771
S	0.2781 (m_2)	1.0836 (s_2)	

Table 1
Traditional Test of the Means
(Average of the First Four Principal Components) from the Sturgeon Sample.

$\tau = 5.0$ $g = 1.0$	s_0	υ_0	Bayes Factor
$m_0 = 1.0$	0.5	10	0.4374
		15	0.4318
		20	0.4291
		30	0.4213
	1.0	10	0.1171
		15	0.1156
		20	0.1173
		30	0.1214
$\tau = 5.0$ $g = 2.0$	0.5	10	0.5058
		15	0.5038
		20	0.5005
		30	0.5025
	1.0	10	0.1119
		15	0.1131
		20	0.1122
		30	0.1136

Table 2
Bayes Factor to Test H_0: $\eta = 0$
$\tau = 5$, $g = (1, 2)$, Equation 6.

$\tau = 10.0$ $g = 1.0$	s_0	η_0	Bayes Factor
	0.5	10	0.4166
		15	0.4113
		20	0.4086
		30	0.4010
	1.0	10	0.1145
		15	0.1156
		20	0.1147
		30	0.1162
$\tau = 10.0$ $g = 2.0$	0.5	10	0.4790
		15	0.4781
		20	0.4801
		30	0.4822
	1.0	10	0.1015
		15	0.1002
		20	0.1020
		30	0.1033

Table 3
Bayes Factor to Test H_0: $\eta = 0$.
$\tau = 10$, $g = (1, 2)$, Equation 6.

$\tau = 15.0$ $g = 1.0$	s_0	υ_0	Bayes Factor
	0.5	10	0.4009
		15	0.3959
		20	0.3882
		30	0.3858
	1.0	10	0.1119
		15	0.1131
		20	0.1148
		30	0.1162
$\tau = 15.0$ $g = 2.0$	0.5	10	0.4582
		15	0.4627
		20	0.4597
		30	0.4619
	1.0	10	0.0937
		15	0.0925
		20	0.0816
		30	0.0955

Table 4
Bayes Factor to Test H_0: $\eta = 0$.
$\tau = 15$, $g = (1, 2)$, Equation 6.

τ	g	Bayes Factor
1	1	0.92
2	1	34.45
3-5	1	> 1000.00
1	2	23.02
2-5	2	> 1000.00

Table 5
Bayes Factor to Test K_0: $\eta = 0$,
$\rho = 1$ vs. $\eta \neq 0$, $\rho \neq 1$, Equation 7.

1.6 Figure

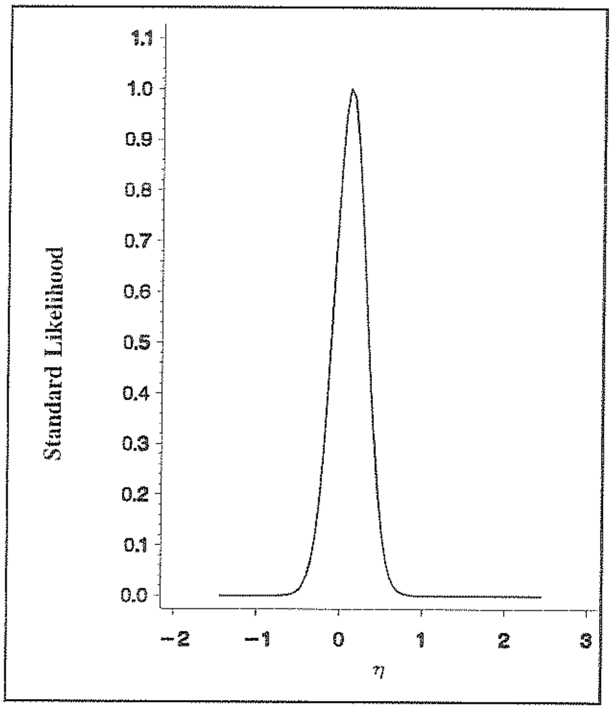

Figure 1
Standard Likelihood Plot of the Difference Between Two Means (η)