ROBUST BAYESIAN HYPOTHESIS TESTING IN THE PRESENCE OF NUISANCE PARAMETERS

by

James Berger and Julia Mortera
Purdue University Università di Trento

Technical Report #91-69C

Department of Statistics
Purdue University

March 1992

*

ROBUST BAYESIAN HYPOTHESIS TESTING IN THE PRESENCE OF NUISANCE PARAMETERS*

by

James Berger and Julia Mortera
Purdue University   Università di Trento

Abstract. Robust Bayesian testing of point null hypotheses is considered for problems involving the presence of nuisance parameters. The robust Bayesian approach seeks answers that hold for a range of prior distributions. Three techniques for handling the nuisance parameter are studied and compared: (i) utilize a noninformative prior to integrate out the nuisance parameter; (ii) utilize a test statistic whose distribution does not depend on the nuisance parameter; and (iii) use a class of prior distributions for the nuisance parameter. These approaches are studied in two examples, the univariate normal model with unknown mean and variance, and a multivariate normal example.

Key words: Bayes factor; Nuisance parameters; P-values; Point null hypothesis; Robust Bayesian analysis.

AMS subject classification; Primary 62F15, Secondary 62F03.

1. INTRODUCTION

1.1 OVERVIEW

Let X be a random variable having density $f(x|\theta,\tau)$, where $\theta \in \mathbf{R}^{\nu}$ and $\tau \in \mathbf{R}^{k}$. Suppose that it is desired to test the null hypothesis $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$; thus θ is the parameter of interest while τ is a nuisance parameter. The parameters θ and τ will be assumed to be a priori independent.

It is well known that, in testing a precise null hypothesis, there is often a conflict between the conclusions reached using classical and Bayesian measures of evidence (P-values and Bayes factors, respectively). Typical Bayes factors and lower bounds on the Bayes factor over all "plausible" prior distributions are often much larger than the corresponding P-values. See, among others, Lindley (1957), Jeffreys (1961), Edwards, Lindman and

* This research was supported by NSF grant DMS-8923071 and by Italian MURST funds.

ROBUST BAYESIAN HYPOTHESIS TESTING IN THE PRESENCE OF NUISANCE PARAMETERS*

by

James Berger and Julia Mortera

Purdue University Università di Roma

Abstract. Robust Bayesian testing of point null hypotheses is considered for problems involving the presence of nuisance parameters. The robust Bayesian approach seeks answers that hold for a range of prior distributions. Three techniques for handling the nuisance parameter are studied and compared: (i) utilize a noninformative prior to integrate out the nuisance parameter; (ii) utilize a test statistic whose distribution does not depend on the nuisance parameter; and (iii) use a class of prior distributions for the nuisance parameter. These approaches are studied in two examples, the univariate normal model with unknown mean and variance, and a multivariate normal example.

Key words: Bayes factor; Nuisance parameters; P-values; Point null hypothesis; Robust Bayesian analysis.

1. INTRODUCTION

1.1 OVERVIEW

Let X be a random variable having density $f(x|\theta,\tau)$, where $\theta \in \mathbf{R}^{\nu}$ and $\tau \in \mathbf{R}^{k}$. Suppose that it is desired to test the null hypothesis $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$; thus θ is the parameter of interest while τ is a nuisance parameter. The parameters θ and τ will be assumed to be a priori independent.

It is well known that, in testing a precise null hypothesis, there is often a conflict between the conclusions reached using classical and Bayesian measures of evidence (P-values and Bayes factors, respectively). Typical Bayes factors and lower bounds on the Bayes factor over all "plausible" prior distributions are often much larger than the corresponding P-values. See, among others, Lindley (1957), Jeffreys (1961), Edwards, Lindman and

* This research was supported by NSF grant DMS-8923071 and by Italian MURST funds.

Savage (1963), Berger (1985, 1986, 1990), Berger and Delampady (1987), Berger and Sellke (1987), Delampady (1989a, 1989b), Moreno and Cano (1989), Delampady and Berger (1990), and Berger and Mortera (1991). Delampady (1989b) and Delampady and Berger (1990) are of particular relevance to the problem of dealing with nuisance parameters, as will be considered here, and certain results from these papers will be extensively utilized.

The purpose of this paper is to compare, through careful study of two typical examples, the success of three different methods for dealing with nuisance parameters in the development of lower bounds on the Bayes factor. The first method is based on integrating out the nuisance parameter via a "non-informative" prior distribution; the second is based on conditioning on a test statistic; and the third method is to seek lower bounds when both parameters vary over classes of distributions. It will be seen that the first two methods lead to useful lower bounds on the Bayes factor, bounds which are, in general, substantially larger than the corresponding P-values. However, the conflict between classical and Bayesian answers is less marked here than in problems where there are no nuisance parameters. The third approach does not seem to produce useful lower bounds on the Bayes factor, unless the class of prior distributions for the nuisance parameter is sharply constrained. Therefore, we do not in general recommend use of broad classes of priors for the nuisance parameters themselves.

The effectiveness of the second method – developing Bayesian lower bounds based on classical test statistics – was somewhat surprising, especially since we found that useful lower bounds could arise from the class of <u>all</u> prior distributions on the alternative. This can also be thought of as a marginal likelihood approach to the problem; see Bertolino, Piccinato, and Racugno (1990) for development of this approach in an analysis of variance setting. The bounds obtained by this method are elementary to derive and often convey very useful information. Furthermore, in some cases we found these bounds to be equal to more elaborate bounds developed via approach 1 and/or more sophisticated classes of priors. Thus there appears to be a surprising similarity between "standard" robust Bayesian bounds and marginal likelihood ratios in the problems we study.

It should be remembered, of course, that lower bounds on Bayes factors are lower bounds. We certainly encourage use of Bayes factors for actual subjective priors, but fear that such will have difficulty replacing the ubiquitous P-value. The lower bounds we discuss are as "automatic" as the P-value, and are considerably more sensible as measures of evidence.

Sections 1.2 and 1.3 give definitions of the classical and Bayesian measures of evidence and the classes of prior distributions that will be considered. An introduction to the three approaches for handling the nuisance parameter will be given in Section 1.4. The different approaches are then applied to a univariate normal model in Section 2 and a multivariate normal model in Section 3.

1.2 MEASURES OF EVIDENCE

The following three measures of evidence will be considered:

<u>P-VALUE:</u>

Classical significance testing is based on a test statistic $T(X)$, large absolute values of which are considered to be evidence against H_0. The observed significance level, the P-value, when x is observed, is defined to be

$$p = P(|T(X)| \geq |T(x)| \; |\theta_0). \tag{1.1}$$

(In our examples, p will not depend upon τ.)

<u>BAYES FACTOR:</u>

It will be assumed that θ and τ are a priori independent, with τ having density $g_2(\tau)$ (w.r.t. Lebesgue measure), and θ, given H_1 is true, having density (w.r.t. Lebesgue measure) $g_1(\theta)$ on $\{\theta : \theta \neq \theta_0\}$. Then the Bayes factor for H_0 versus H_1 is

$$B(x, g_1, g_2) = \frac{\int f(x|\theta_0, \tau) g_2(\tau) d\tau}{\int\int f(x|\theta, \tau) g_1(\theta) g_2(\tau) d\theta d\tau}. \tag{1.2}$$

<u>BOUNDS ON B:</u>

Of interest will be lower bounds on (1.2) as g_1 and g_2 vary over classes $\mathcal{G}_1$ and $\mathcal{G}_2$, respectively. Thus define

$$\underline{B}(x, \mathcal{G}_1, \mathcal{G}_2) = \inf_{g_1 \in \mathcal{G}_1, g_2 \in \mathcal{G}_2} B(x, g_1, g_2), \tag{1.3}$$

where $B(x, g_1, g_2)$ is given in (1.2).

1.3 CLASSES OF PRIOR DISTRIBUTIONS

The classes of prior distributions that will be considered are (letting $\mathcal{U}[a,b]$ stand for the uniform distribution on the interval (a,b)):

$$\mathcal{G}_A = \{\text{all distributions}\},$$
$$\mathcal{G}_S = \{\text{all symmetric distributions about } \theta_0\},$$
$$\mathcal{G}_{US} = \{\text{all symmetric unimodal distributions with mode at } \theta_0\},$$
$$\mathcal{G}_{SU} = \{\text{all } \mathcal{U}[\theta_0 - r, \theta_0 + r] \text{ distributions}, r > 0\},$$
$$\mathcal{G}_D = \{\text{all nonincreasing densities on } [0, \infty)\},$$
$$\mathcal{G}_U = \{\text{all } \mathcal{U}[0, k] \text{ distributions}, k > 0\}.$$

The first four will typically be utilized as classes of g_1, and the last two as classes of g_2. Useful formulas for computing the lower bounds on the Bayes factor over the classes $\mathcal{G}_A$ and $\mathcal{G}_{US}$ are given, for example, in Lemmas 2.1 and 2.2 of Berger and Mortera (BM) (1991).

1.4 HANDLING THE NUISANCE PARAMETER

Three different approaches of dealing with the nuisance parameter τ will be considered.

Approach 1: *Integrate Out Over* τ

The nuisance parameter can be eliminated by integrating $f(x|\theta, \tau)$ w.r.t. a "non-informative" prior distribution $g_2^*(\tau)$. The lower bound on the Bayes factor is then

$$\underline{B}(x, \mathcal{G}_1, g_2^*) = \frac{\int f(x|\theta_0, \tau) g_2^*(\tau) d\tau}{\sup\limits_{g_1 \in \mathcal{G}_1} \int [\int f(x|\theta, \tau) g_2^*(\tau) d\tau] g_1(\theta) d\theta}. \tag{1.4}$$

Approach 2: *Utilize a Test Statistic with a* τ*-Free Test*

A second approach for eliminating the nuisance parameter τ is to base the analysis on a test statistic $T(X)$ which has a density $h(T(x)|\delta)$, where δ is a function of θ and τ such that the problem of testing the null hypothesis $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$ becomes equivalent to testing $H_0 : \delta = \delta_0$ versus $H_1 : \delta \neq \delta_0$. The original test for θ based on X is then replaced by the test for δ based on $T(X)$. In this latter problem, let g denote a prior distribution for δ given H_1 (i.e., a distribution on $\{\delta : \delta \neq \delta_0\}$), and $\mathcal{G}_A$ stand for all such prior distributions. Then we will consider the lower bound on the Bayes factor

$$\underline{B}(T(x), \mathcal{G}_A) \equiv \inf_{g \in \mathcal{G}_A} \frac{h(T(x)|\delta_0)}{\int h(T(x)|\delta) g(\delta) d\delta}$$
$$= h(T(x)|\delta_0) / h(T(x)|\hat{\delta}), \tag{1.5}$$

where $\hat{\delta}$ is a value of δ which maximizes $h(T(x)|\delta)$.

Considering the test based on $T(x)$ has several motivations. The first is simply to note that this is the common practice in classical and likelihood statistics, and following this practice will allow us to provide comparative results. More fundamentally, the reduction to $T(X)$ and δ can often follow from Bayesian (and classical) invariance arguments. This was observed and developed in Delampady (1989b). Finally, it sometimes happens that one is presented only with $T(X)$. An example is when one learns only the P-value (which then essentially becomes $T(x)$); see Berger and Mortera (1991).

One of the surprises of this study is that, for analysis based on $T(X)$, it frequently suffices to use the crude $\mathcal{G}_A$ to obtain $\underline{B}$, rather than the more reasonable (but difficult to work with) classes such as $\mathcal{G}_S$ or $\mathcal{G}_{US}$. This is particularly interesting because $\underline{B}(T(x), \mathcal{G}_A)$ corresponds to the natural likelihood ratio measure for comparing hypotheses in the marginal likelihood approach (cf., Bertolino, Piccinato and Racugno, 1990).

Approach 3: *Compute Lower Bounds Over g_1 and g_2.*

A third possible approach to the problem of dealing with nuisance parameters is to determine the lower bound on the Bayes factor when $g_1(\theta)$ and $g_2(\tau)$ vary over $\mathcal{G}_1$ and $\mathcal{G}_2$, respectively. The bound is then simply given by (1.3).

2. THE UNIVARIATE NORMAL PROBLEM

Throughout this section the following example will be considered. Let $X_1, X_2, \ldots, X_n$ be iid $\mathcal{N}(\mu, \sigma^2)$, where both μ and σ^2 are unknown, and suppose it is desired to test $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$. Having observed $\underset{\sim}{x} = (x_1, \ldots, x_n)$, the likelihood is

$$f(\underset{\sim}{x}|\mu, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \exp\left\{-\frac{s^2}{2\sigma^2}\left[\frac{n(\bar{x}-\mu)^2}{s^2} + 1\right]\right\}, \tag{2.1}$$

where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ and $s^2 = \sum_{i=1}^{n}(x_i - \bar{x})^2$. Approaches 1 to 3 will now be applied to this example.

2.1 Approach 1: Integrating Out σ

The common noninformative prior for σ is $g_2^*(\sigma) = \sigma^{-1}$. It is well known that integrating over σ in (2.1) with respect to g_2^* yields a marginal likelihood for μ proportional to a t-density with location $\bar{x}$, scale $s/\sqrt{n}$, and $(n-1)$ degrees of freedom. Thus (1.4) becomes

$$\underline{B}(\underset{\sim}{x}, \mathcal{G}_1, g_2^*) = \frac{[1 + t^2/(n-1)]^{-n/2}}{\sup\limits_{g_1 \in \mathcal{G}_1} \int [1 + \frac{n(\mu - \bar{x})^2}{s^2}]^{-n/2} g_1(\mu) d\mu}, \tag{2.2}$$

where $t = \sqrt{n}(\bar{x} - \mu_0)/(s/\sqrt{n-1})$. Note that t is the classical t-statistic for testing $H_0 : \mu = \mu_0$.

<u>Case 1</u>: $\mathcal{G}_1 = \mathcal{G}_A = \{$ *all distributions on* $\{\mu : \mu \neq \mu_0\}\}$.

With this choice of $\mathcal{G}_1$, the supremum in (2.2) is clearly attained by a point mass at $\bar{x}$, so that

$$\underline{B}_A \equiv \underline{B}(\underset{\sim}{x}, \mathcal{G}_A, g_2^*) = [1 + t^2/(n-1)]^{-n/2}. \tag{2.3}$$

Table 1 gives the values of $\underline{B}_A$ for some standard t, given in terms of their corresponding P-values. Observe that, for small n, roughly $n \leq 4$, $\underline{B}_A$ is smaller than the P-value, in contrast to the well-studied normal case ($n = \infty$).

Table 1. Lower bounds on Bayes factors when σ is integrated out

	$p = .001$			$p = .01$			$p = .05$		
$n-1$	$\underline{B}_A$	$\underline{B}_S$	$\underline{B}_{US}$	$\underline{B}_A$	$\underline{B}_S$	$\underline{B}_{US}$	$\underline{B}_A$	$\underline{B}_S$	$\underline{B}_{US}$
1	2.5×10^{-6}	5.0×10^{-6}	.001	.00025	.00049	.0115	.0062	.0123	.0685
2	9.0×10^{-5}	1.8×10^{-4}	.002	.0028	.0056	.0261	.0304	.0606	.1447
4	.00059	.0012	.005	.010	.0201	.0512	.0683	.1359	.2381
6	.0012	.0023	.007	.0155	.0309	.0674	.0887	.1768	.2847
8	.0016	.0032	.009	.0192	.0384	.0779	.1009	.2013	.3118
10	.0020	.0039	.010	.0218	.0437	.0852	.1090	.2174	.3294
20	.0030	.0059	.013	.0282	.0564	.1022	.1266	.2528	.3674
∞	.0044	.0089	.018	.0362	.0725	.1223	.1465	.2928	.4084

Case 2: $\mathcal{G}_1 = \mathcal{G}_S = \{$ *distributions symmetric about* $\mu_0\}$.

For this choice of $\mathcal{G}_1$, (2.2) and Theorem 3 of Berger and Sellke (1987) yield

$$\underline{B}_S \equiv \underline{B}(\underset{\sim}{x}, \mathcal{G}_S, g_2^*) = \frac{f_{n-1}(t)}{\sup_{r>0}\{\frac{1}{2}f_{n-1}(t+r) + \frac{1}{2}f_{n-1}(t-r)\}}, \tag{2.4}$$

where f_{n-1} is the t density with $(n-1)$ degrees of freedom. Table 1 presents values of $\underline{B}_S$ for common P-values.

Case 3: $\mathcal{G}_1 = \mathcal{G}_{US} = \{$ *symmetric, unimodal distributions about* $\mu_0\}$.

For this choice of $\mathcal{G}_1$, (2.2) and Lemma 2.2 of BM (1991) yield

$$\underline{B}_{US} \equiv \underline{B}(\underset{\sim}{x}, \mathcal{G}_{US}, g_2^*) = \frac{f_{n-1}(t)}{\sup_{r>0}\{\frac{1}{2r}[F_{n-1}(r-t) - F_{n-1}(-r-t)]\}}, \tag{2.5}$$

where F_{n-1} is the cdf of the t-density with $(n-1)$ degrees of freedom. Values of $\underline{B}_{US}$ are also given in Table 1.

Because $\mathcal{G}_{US} \subset \mathcal{G}_S \subset \mathcal{G}_A$, the ordering $\underline{B}_A \leq \underline{B}_S \leq \underline{B}_{US}$, observed in Table 1, must always hold. The difference between $\underline{B}_A$ and $\underline{B}_S$ is only a factor of about 2, so that the assumption of symmetry in g_1 is of only minor benefit in improving the lower bound. The additional assumption of unimodality has a more significant effect, increasing the lower bound by an additional factor of as much as 200. We feel that this assumption is warranted if an effort is being made to be "objective", and so would advocate use of $\underline{B}_{US}$ as the more sensible lower bound.

Interestingly, $\underline{B}_{US}$ appears to always be larger than the corresponding P-value, even for $n-1=1$. Indeed, the following lemma establishes this in an asymptotic (large t) sense. This lemma complements Theorem 7 of Berger and Sellke (1987), which gives the asymptotic behavior of P-value/$\underline{B}_{US}$ for the normal case.

Lemma 2.1

$$\lim_{t\to\infty} \frac{P\text{-value}}{\underline{B}_{US}} = \frac{1}{(n-1)}. \tag{2.6}$$

Proof: See Appendix I. □

2.2 Approach 2: Utilizing a Test Statistic

Consider the test statistic

$$T(X) = \frac{\sqrt{n}\,|\bar{X} - \mu_0|}{S/\sqrt{n-1}}, \tag{2.7}$$

where $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ and $S^2 = \sum_{i=1}^{n}(X_i - \bar{X})^2$. The density of T is $\frac{1}{2}[h_{n-1}(t|\delta) + h_{n-1}(-t|\delta)]$, where $h_{n-1}(\cdot|\delta)$ is the non-central t-distribution with $(n-1)$ degrees of freedom and noncentrality parameter $\delta = \sqrt{n}(\mu - \mu_0)/\sigma$. Testing $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$ is equivalent to testing $H_0 : \delta = 0$ versus $H_1 : \delta \neq 0$. From (1.5), the lower bound on the Bayes factor, when $g \in \mathcal{G}_A$, is

$$\underline{B}(T(x), \mathcal{G}_A) = \frac{2h_{n-1}(t|0)}{\sup_{\delta}[h_{n-1}(t|\delta) + h_{n-1}(-t|\delta)]}. \tag{2.8}$$

Simple algebra yields

$$\underline{B}(T(x), \mathcal{G}_A) = \frac{2^{\frac{1}{2}(\nu+1)}\Gamma(\frac{\nu+1}{2})}{\sup_{\delta}\{e^{-\frac{1}{2}\delta^2}[\text{ int }(t,\nu,\delta) + \text{ int }(-t,\nu,\delta)]\}}, \tag{2.9}$$

where $\nu = n - 1$ and

$$\text{int }(t,\nu,\delta) = \int_0^\infty z^\nu \exp\Big[-\frac{1}{2}z^2 + \frac{z\delta t}{\sqrt{\nu + t^2}}\Big]dz.$$

For $\nu > 2$, a useful approximation for computing (2.9) is

$$\underline{B}(T(x), \mathcal{G}_A) \cong \frac{2\Gamma(\frac{\nu+1}{2})}{\sup_{\delta}\{e^{-\frac{1}{2}\delta^2}[\sum_{j=0}^{200}[D_j(t,\nu,\delta) + D_j(-t,\nu,\delta)]]\}}, \tag{2.10}$$

where

$$D_j(t,\nu,\delta) = \Gamma\Big(\frac{\nu+j+1}{2}\Big)(j!)^{-1}\Big[\frac{t\delta\sqrt{2}}{\sqrt{\nu+t^2}}\Big]^j.$$

Results are shown in Table 2 where also the value, $\hat{\delta}$, that maximizes the denominator is given. Note that $\underline{B}(T(x), \mathcal{G}_A)$ ranges from twice the P-value at $\nu = 1$ to about 6 times the P-value at $\nu = 20$.

Table 2. Lower bounds on Bayes factors when utilizing a test statistic

	$p = .001$		$p = .01$		$p = .05$	
$\nu = n - 1$	$\underline{B}(T(x), \mathcal{G}_A)$	$\hat{\delta}$	$\underline{B}(T(x), \mathcal{G}_A)$	$\hat{\delta}$	$\underline{B}(T(x), \mathcal{G}_A)$	$\hat{\delta}$
1	.0021	636.215	.0267	63.670	.1035	12.745
2	.0027	31.604	.0273	9.975	.1391	4.405
3	.0033	12.954	.0329	5.917	.1676	3.287
4	.0037	8.664	.0377	4.686	.1887	2.869
6	.0045	6.024	.0449	3.782	.2158	2.517
8	.0051	5.104	.0498	3.418	.2321	2.361
10	.0056	4.644	.0533	3.223	.2428	2.277
20	.0069	3.886	.0618	2.875	.2663	2.108
∞	.0089	3.291	.0725	2.576	.2928	1.958

Observe in the $\nu = \infty$ rows of Tables 1 and 2 that $\underline{B}_S = \underline{B}(T(x), \mathcal{G}_A)$. This is indeed generally true for $\nu = \infty$, as can be seen from Lemmas 2.3 and 2.4 of BM (1991).

It is interesting to compare the lower bounds in Table 2 with those in Table 2 of Delampady (1989b). The Table 2 entries here are roughly twice the analogous bounds of Delampady. This difference is due to the fact that we utilize the test statistic $\sqrt{n(n-1)}|\bar{X} - \mu_0|/S$, while Delampady effectively utilizes $\sqrt{n(n-1)}(\bar{X} - \mu_0)/S$.

2.3 Approach 3: Lower Bounds Over $g_1(\mu)$ and $g_2(\sigma)$.

First, note that if $g_1(\mu) \in \mathcal{G}_A$, the lower bound on the Bayes factor as a function of σ^2 is (slightly abusing notation)

$$\begin{aligned}\underline{B}(\underset{\sim}{x}, \mathcal{G}_A, \sigma^2) &= \frac{f(\underset{\sim}{x}|\mu_0, \sigma^2)}{f(\underset{\sim}{x}|\hat{\mu}, \sigma^2)} \\ &= \exp\{-\frac{s^2}{2\sigma^2} q^2\},\end{aligned}$$

where $q = \sqrt{n}(\bar{x} - \mu_0)/s$ and $\hat{\mu} = \bar{x}$. Since $\lim_{\sigma^2 \to 0} \underline{B}(\underset{\sim}{x}, \mathcal{G}_A, \sigma^2) = 0$ and $\lim_{\sigma^2 \to \infty} \underline{B}(\underset{\sim}{x}, \mathcal{G}_A, \sigma^2) = 1$, one cannot find useful bounds when also using $\mathcal{G}_A$ for the class of prior distributions on σ^2.

One might hope that use of "smoother" classes of prior distributions would give more useful lower bounds. Natural candidates are $\mathcal{G}_{US}$ for $g_1(\mu)$ and $\mathcal{G}_D$ for $g_2(\sigma)$. Unfortunately, these classes are still too large to give useful bounds, as the following Lemma shows.

Lemma 2.2. Defining $q = \sqrt{n}(\bar{x} - \mu_0)/s$ and $s^2 = \sum(x_i - \bar{x})^2$,

$$\underline{B}(\underset{\sim}{x}, \mathcal{G}_{US}, \mathcal{G}_D) = \inf_k \left\{ \frac{\int_0^k \sigma^{-n} \exp\{-\frac{s^2}{2\sigma^2}[q^2+1]\} d\sigma}{\sup_r \int_{\mu_0 - r}^{\mu_0 + r} \int_0^k \frac{1}{2r} \sigma^{-n} \exp\{-\frac{1}{2\sigma^2}[n(\bar{x}-\mu)^2 + s^2]\} d\mu d\sigma} \right\} \quad (2.11)$$
$$= 0.$$

Proof: See Appendix I. □

Although $\underline{B}(\underset{\sim}{x}, \mathcal{G}_{US}, \mathcal{G}_D)$ is thus useless, the representation in (2.11) (see also the Proof of Lemma 2.2 in Appendix I) suggests that useful information about the Bayes factor can be conveyed by considering $B(\underset{\sim}{x}, g_{1,r}, g_{2,k})$, where $g_{1,r}(\mu)$ is $\mathcal{U}[\mu_0 - r, \mu_0 + r]$ and $g_{2,k}(\sigma)$ is $\mathcal{U}[0, k]$, the "extreme points" of $\mathcal{G}_{US}$ and $\mathcal{G}_D$, respectively. Graphing $B(\underset{\sim}{x}, g_{1,r}, g_{2,k})$ over the range of r and k deemed to be relevant can indicate the degree of robustness of B. In the special case $n = 4$, the following can be used to do the computation.

Lemma 2.3.

$$B(\underset{\sim}{x}, g_{1,r}, g_{2,k}) = \frac{r^* A^{-3/2} \Gamma(\frac{3}{2}, Ak^{*2})}{2\sqrt{2} D(k^*, q, r^*)}, \quad (2.12)$$

where $q = 2(\bar{x} - \mu_0)/s$, $A = \frac{1}{2}(1 + q^2)$, $k^* = s/k$, $r^* = r/2s$,

$$\Gamma(a, x) = \frac{1}{\Gamma(a)} \int_x^\infty e^{-t} t^{a-1} dt \quad \text{for } a > 0, \quad (2.13)$$

and

$$\begin{aligned} D(k^*, q, r^*) = {} & e^{-\frac{1}{2}k^{*2}} \{\Phi[k^*(q + r^*)] - \Phi[k^*(q - r^*)]\} \\ & + \frac{(q + r^*)}{\sqrt{(q + r^*)^2 + 1}} \Phi(-k^* \sqrt{1 + (q + r^*)^2}) \\ & - \frac{(q - r^*)}{\sqrt{(q - r^*)^2 + 1}} \Phi(-k^* \sqrt{1 + (q - r^*)^2}). \end{aligned} \quad (2.14)$$

Proof: See Appendix I. □

A graph and a contour plot of $B(\underset{\sim}{x}, g_{1,r}, g_{2,k})$, for $n = 4, s = 1$, and $\underset{\sim}{x}$ such that the P-value is 0.05, are given in Figures 1a and 1b, respectively. It was visually superior to use $k^* = 1/k$ as an argument of B instead of k. It is immediately apparent that B, in this example, is large (near 1) only if r and k^* are small, while B goes rapidly to zero as k^* grows large ($k \to 0$). Note that B is typically larger than 0.05, the P-value.

As an example of the use of such figures to ascertain robustness, suppose the user's subjective beliefs are reflected by choices $1.25 \le r \le 2.5$ and $2.5 \le k \le 100$ (i.e., any unimodal symmetric prior for μ between the "extreme points" $\mathcal{U}[\mu_0 - 1.25, \mu_0 + 1.25]$ and $\mathcal{U}[\mu_0 - 2.5, \mu_0 + 2.5]$ is deemed possible, as are nonincreasing priors for σ between the "extreme points" $\mathcal{U}[0, 2.5]$ and $\mathcal{U}[0, 100]$). From Figure 1b, the corresponding range of B can be seen to be $0.14 \le B \le 0.27$, which would be a reasonably robust conclusion.

3. THE SYMMETRIC MULTIVARIATE NORMAL PROBLEM

All computations in this section refer to the following example. Let $\underset{\sim}{X} = (X_1, \ldots, X_k) \sim N(\underset{\sim}{\mu}, I)$ and suppose it is desired to test $H_0 : \underset{\sim}{\mu} = \underset{\sim}{\mu}_0$ versus $H_1 : \underset{\sim}{\mu} \ne \underset{\sim}{\mu}_0$. We do not consider Approach 3 in this example, because the "nuisance parameter" is derived, rather than natural.

3.1 Approach 1: Integrating Out the Nuisance Parameter

Berger and Delampady (1987) found the lower bounds for this example when $g(\underset{\sim}{\mu}) \in \mathcal{G}_{US}$. This is equivalent to adopting approach 1, with $\theta = |\underset{\sim}{\mu} - \underset{\sim}{\mu}_0|$ being the parameter of interest, $\underset{\sim}{\tau} = (\underset{\sim}{\mu} - \underset{\sim}{\mu}_0)/|\underset{\sim}{\mu} - \underset{\sim}{\mu}_0|$ being the nuisance parameter, $g_2(\underset{\sim}{\tau})$ being the uniform distribution on the unit ball, and $g_1(\theta) \in \mathcal{G}^* = \{$ densities on $(0, \infty)$, proportional to $\theta^{k-1} w(\theta)$, with w nonincreasing $\}$. Also the null hypothesis becomes $H_0 : \theta = 0$ and the alternative becomes $H_1 : \theta \ne 0$. The lower bound on the Bayes factor, $\underline{B}^* = \underline{B}(\underset{\sim}{x}, \mathcal{G}^*, g_2)$, is given in Table 3 for some standard P-values (see Berger and Delampady (1987), although the entries computed there for the $k = \infty$ case were incorrect). Note that $\underline{B}^*$ remains almost constant as the dimensionality k increases.

Table 3. Lower bounds on Bayes factors for the multivariate normal problem

Dimension	$p = .001$			$p = .01$			$p = .05$		
k	$\underline{B}^*$	$\underline{B}(T(\underset{\sim}{x}), \mathcal{G}_A)$	$\hat{\delta}$	$\underline{B}^*$	$\underline{B}(T(\underset{\sim}{x}), \mathcal{G}_A)$	$\hat{\delta}$	$\underline{B}^*$	$\underline{B}(T(\underset{\sim}{x}), \mathcal{G}_A)$	$\hat{\delta}$
1	.0182	.0089	10.83	.1227	.0725	6.64	.4092	.2929	3.83
2	.0143	.00914	12.78	.0978	.0739	8.14	.3481	.2923	4.87
4	.0114	.00930	15.37	.0850	.0745	10.14	.3141	.2907	6.27
6	.0097	.00936	17.32	.0807	.0747	11.62	.3023	.2891	7.31
8	.0095	.00939	18.96	.0789	.0746	12.86	.2963	.2879	8.18
10	.0094	.00940	20.40	.0777	.0745	13.95	.2927	.2867	8.95
20	.0093	.00939	26.04	.0743	.0739	18.21	.2844	.2828	11.95
∞	.0085	.00845		.0669	.0669		.2585	.2585	

3.2 Approach 2: Utilizing a Test Statistic

Consider the test statistic $T(\underset{\sim}{X}) = |\underset{\sim}{X} - \underset{\sim}{\mu}_0|^2$ which has a non-central chi-squared distribution with k degrees of freedom and non-centrality parameter $\delta = |\underset{\sim}{\mu} - \underset{\sim}{\mu}_0|^2 = \theta^2$, to be denoted by $h_k(t|\delta)$. The test $H_0 : \underset{\sim}{\mu} = \underset{\sim}{\mu}_0$ versus $H_1 : \underset{\sim}{\mu} \neq \underset{\sim}{\mu}_0$ becomes equivalent to the test $H_0 : \delta = 0$ versus $H_1 : \delta > 0$. For $g_1(\delta) \in \mathcal{G}_A$, the lower bound on the Bayes factor is given as in (1.5) by

$$\underline{B}(T(\underset{\sim}{x}), \mathcal{G}_A) = \frac{h_k(t|\delta = 0)}{h_k(t|\hat{\delta})}, \tag{3.1}$$

where $h_k(t|\delta = 0)$ is a central chi-squared distribution with k degrees of freedom and $\hat{\delta}$ is that value of δ which maximizes $h_k(t|\delta)$. Results are shown in Table 3 (see also Delampady (1989b)). Note that, as the dimensionality k increases, $\underline{B}^*$ and $\underline{B}(T(\underset{\sim}{x}), \mathcal{G}_A)$ grow closer together. In fact, as $k \to \infty$ with p remaining fixed, they approach the same limit

$$\lim_{k\to\infty} \underline{B}^* = \lim_{k\to\infty} \underline{B}(T(\underset{\sim}{x}), \mathcal{G}_A) = \exp\{-\frac{1}{2}z_p^2\}, \tag{3.2}$$

where z_p is the p-th quantile of the standard normal distribution (see Appendix II for explanation). This limit is given in the $k = \infty$ row of Table 3. Note that the convergence to this limit is very slow.

The first equality in (3.2) is surprising, since $\underline{B}^*$ is computed using the quite restrictive class of nonincreasing priors in $\delta = |\underset{\sim}{\mu} - \underset{\sim}{\mu}_0|^2$, while $\underline{B}(T(\underset{\sim}{x}), \mathcal{G}_A)$ uses <u>all</u> priors on δ. Note, from Table 3, that $\underline{B}^*$ and $\underline{B}(T(\underset{\sim}{x}), \mathcal{G}_A)$ tend to also agree even for smaller values of k.

APPENDIX I

Proof of Lemma 2.1

Define

$$\psi(r,t) \equiv \frac{f_{n-1}(t)}{\frac{1}{2r}\left[F_{n-1}(r-t) - F_{n-1}(-r-t)\right]}.$$

Write $r = t + g(t)$, and consider any choice of $g(t)$ which satisfies, as $t \to \infty$,

$$g(t) \to \infty \quad \text{and} \quad g(t) = o(t). \tag{A.1}$$

Then, clearly,

$$\begin{aligned}\psi(t+g(t),t) &= \frac{2(t+g(t))f_{n-1}(t)}{F_{n-1}(g(t)) - F_{n-1}(-2t-g(t))} \\ &= 2tf_{n-1}(t)(1+o(1)).\end{aligned}$$

We establish that

$$\underline{B}_{US} = 2tf_{n-1}(t)(1+o(1)) \tag{A.2}$$

through a proof by contradiction, assuming that the minimizing $r^* = t + g^*(t)$ violates one or both of the conditions in (A.1) along some increasing sequence $\{t_i\}$ with $t_i \to \infty$.

Suppose, first, that $g^*(t_i) \le K$, for some $K < \infty$. Then

$$\begin{aligned}\psi(t_i+g^*(t_i),t_i) &= \frac{2(t_i+g^*(t_i))f_{n-1}(t_i)}{F_{n-1}(g^*(t_i))}(1+o(1)) \\ &\ge 2(1+\varepsilon)t_i f_{n-1}(t_i)(1+o(1)),\end{aligned}$$

since $F_{n-1}(g^*(t_i)) \le 1-\varepsilon$ for some $\varepsilon > 0$. But this is larger than (A.2), contradicting the assumption that $(t_i + g^*(t_i))$ is the minimizer.

Suppose, next, that the second condition in (A.1) is violated, i.e., that $g^*(t_i) \ge \varepsilon t_i$ for some $\varepsilon > 0$. Then

$$\begin{aligned}\psi(t_i+g^*(t_i),t_i) &= \frac{2(t_i+g^*(t_i))f_{n-1}(t_i)}{F_{n-1}(g^*(t_i))}(1+o(1)) \\ &\ge 2(1+\varepsilon)t_i f_{n-1}(t_i)(1+o(1)),\end{aligned}$$

since $F_{n-1}(g^*(t_i)) \le 1$. But, again, this is larger than (A.2), which contradicts $t_i + g^*(t_i)$ being the minimizer. Thus we have established (A.2).

It is straightforward to show that the P-value satisfies

$$\begin{aligned}P(|T| > t) &= 2(1 - F_{n-1}(t)) \\ &= 2tf_{n-1}(t)/(n-1),\end{aligned}$$

which together with (A.2) yields the conclusion of the Lemma. □

Proof of Lemma 2.2

Standard representations of $\mathcal{G}_{US}$ and $\mathcal{G}_D$ are

$$\mathcal{G}_{US} = \{g_1(\mu) = \int_0^\infty \frac{1}{2r} 1_{(\mu_0 - r, \mu_0 + r)}(\mu) dF_1(r), \text{ where } F_1 \text{ is any distribution function on } (0,\infty)\},$$

$$\mathcal{G}_D = \{g_2(\sigma) = \int_0^\infty \frac{1}{k} 1_{(0,k)}(\sigma) dF_2(k), \text{ where } F_2 \text{ is any distribution function on } (0,\infty)\}.$$

Defining

$$\psi(\mu, k) = \frac{1}{k}\int_0^k \sigma^{-n} \exp\{-\frac{1}{2\sigma^2}[n(\bar{x}-\mu)^2 + s^2]\} d\sigma,$$

$$\psi^*(r,k) = \frac{1}{2r}\int_{\mu_0 - r}^{\mu_0 + r} \psi(\mu, k) d\mu,$$

it follows that

$$\underline{B}(\underset{\sim}{x}, \mathcal{G}_{US}, \mathcal{G}_D) = \inf_{F_1, F_2} \frac{\int_0^\infty \psi(\mu_0, k) dF_2(k)}{\int_0^\infty \int_0^\infty \psi^*(r,k) dF_2(k) dF_1(r)}.$$

Treating the ratio of integrals above as a ratio of linear functionals w.r.t. the measure $F_1 \times F_2$, Lemma A.1 in Sivaganesan and Berger, 1989, shows that the lower bound is attained at a point mass measure, establishing the first equality in (2.11).

Fixing r and applying L'Hospital's rule to $\psi(\mu_0, k)/\psi^*(r,k)$, one obtains

$$\begin{aligned}
\underline{B}(\underset{\sim}{x}, \mathcal{G}_{US}, \mathcal{G}_D) &\le \lim_{k\to 0} \frac{k^{-n}\exp\left\{-\frac{s^2}{2k^2}[q^2+1]\right\}}{\int_{\mu_0 - r}^{\mu_0 + r} \frac{1}{2r} k^{-n} \exp\left\{-\frac{1}{2k^2}[n(\bar{x}-\mu)^2 + s^2]\right\} d\mu} \\
&= \lim_{k\to 0}\left[\frac{1}{2r}\int_{\mu_0 - r}^{\mu_0 + r} \exp\left\{-\frac{n}{2k^2}[(\bar{x}-\mu)^2 - (\bar{x}-\mu_0)^2]\right\} d\mu\right]^{-1} \\
&= \lim_{k\to 0}\left[\frac{1}{2r}\int_{-r}^{r} \exp\left\{\frac{\eta}{2k^2}[2(\bar{x}-\mu_0) - \eta]\right\} d\eta\right]^{-1}, \qquad (A.3)
\end{aligned}$$

where $\eta = \mu - \mu_0$. Without loss of generality, assume $\bar{x} > \mu_0$ and choose $r = \bar{x} - \mu_0$.

Using the monotone convergence theorem, one has

$$\lim_{k\to 0}\int_{-r}^{r} \exp\left\{\frac{n\eta}{2k^2}[2(\bar{x}-\mu_0) - \eta]\right\} d\eta \ge \int_0^r \lim_{k\to 0} \exp\left\{\frac{n\eta}{2k^2}[2r - \eta]\right\} d\eta = \infty.$$

Thus, from (A.3),

$$\lim_{k\to 0} \underline{B}(\underset{\sim}{x}, \mathcal{G}_{US}, \mathcal{G}_D) = 0,$$

which establishes the result.

Proof of Lemma 2.3

$$B(\underset{\sim}{x}, g_{1,r}, g_{2,k}) = \frac{\int_0^k \sigma^{-n} \exp\left\{-\frac{s}{2\sigma^2}[q^2+1]\right\} d\sigma}{\int_0^k \frac{1}{r}\sqrt{\frac{\pi}{2n}}\sigma^{(1-n)} \exp\left\{-\frac{s^2}{2\sigma^2}\right\}[\Phi(-q\frac{s}{\sigma}+\frac{r\sqrt{n}}{\sigma}) - \Phi(-q\frac{s}{\sigma}-\frac{r\sqrt{n}}{\sigma})]d\sigma}.$$

Setting $w = s/\sigma$ and $k = s/k^*$, the numerator is

$$s^{(1-n)} \int_{k^*}^{\infty} w^{n-2} \exp\{-\frac{1}{2}(1+q^2)w^2\}dw = \frac{\Gamma((n-1)/2)}{2} A^{-\frac{n-1}{2}} \Gamma(\frac{n-1}{2}, Ak^{*^2}),$$

where $A = \frac{1}{2}(1+q^2)$ and $\Gamma(a,x)$ is the incomplete Gamma function given in (2.13). Setting $r = r^* s/\sqrt{n}$, the denominator is

$$s^{(1-n)} \frac{\sqrt{2\pi}}{2r^*} \int_{k^*}^{\infty} w^{n-3} \exp\left\{-\frac{1}{2}w^2\right\} \{\Phi[w(-q+r^*)] - \Phi[w(-q-r^*)]\}dw.$$

When $n = 4$, this last integral can be evaluated by integration by parts (using $\int w \exp\{-w^2/2\}dw = -\exp\{-w^2/2\}$), leading to (2.12). □

APPENDIX II

Explanation of (3.2)

Observe first that, as $k \to \infty$, the critical value t_p, for which $Pr(T(\underset{\sim}{x}) \geq t_p | \delta = 0) = p$, is given by

$$t_p = k + \sqrt{2k}\ z_p + O(1),$$

where z_p is the standard normal critical value. Using the fact that $(T-k-\delta)/[2(k+2\delta)]^{1/2}$ tends to a standard normal random variable as $k \to \infty$, it can be shown that

$$\frac{h_k(t_p|0)}{h_k(t_p|\delta)} \to \frac{(2k)^{-1/2}\exp\{-\frac{1}{4k}(\sqrt{2k}\ z_p + O(1))^2\}}{[2(k+2\delta)]^{-1/2}\exp\{-\frac{1}{4(k+2\delta)}(\sqrt{2k}\ z_p - \delta + O(1))^2\}}.$$

This achieves its minimum at

$$\hat{\delta} = \sqrt{2k}\ z_p(1+o(1)),$$

from which it follows easily that

$$\lim_{k\to\infty} \frac{h_k(t_p|0)}{h_k(t_p|\hat{\delta})} = \exp\{-\frac{1}{2}z_p^2\}.$$

In establishing the first equality in (3.2), it can be shown that

$$\underline{B}^* = \frac{\exp\{-\frac{1}{2}|\underset{\sim}{x} - \underset{\sim}{\mu}_0|^2\}}{\sup\limits_{g_1 \in \mathcal{G}^*} \int_0^\infty [\int \exp\{-\frac{1}{2}|(\underset{\sim}{x} - \underset{\sim}{\mu}_0) - \theta \underset{\sim}{\tau}|^2\} g_2(\underset{\sim}{\tau}) d\underset{\sim}{\tau}] g_1(\theta) d\theta}$$
$$= \frac{h_k(t|0)}{\sup\limits_{g_1 \in \mathcal{G}^*} \int_0^\infty h_k(t|\theta^2) g_1(\theta) d\theta}.$$

Since the denominator is less than or equal to $h_k(t|\hat{\delta})$, it is immediate that $\underline{B}^* \geq \underline{B}(T(\underset{\sim}{x}), \mathcal{G}_A)$. To establish that the reverse inequality holds as $k \to \infty$, choose

$$g_1^*(\theta) = \frac{1}{k} \hat{\delta}^{k/2} \theta^{k-1} 1_{(0,\sqrt{\hat{\delta}})}(\theta)$$

(which is in the class $\mathcal{G}^*$). Clearly, for any $\varepsilon > 0$,

$$\lim_{k\to\infty} \frac{\int h_k(t|\theta^2) g_1^*(\theta) d\theta}{h_k(t|\hat{\delta})} = \lim_{k\to\infty} \int_0^{\sqrt{\hat{\delta}}} \frac{h_k(t|\theta^2)}{h_k(t|\hat{\delta})} \frac{1}{k} \hat{\delta}^{k/2} \theta^{k-1} d\theta$$
$$= \lim_{k\to\infty} \int_{\sqrt{\hat{\delta}-\varepsilon}}^{\sqrt{\hat{\delta}}} \frac{h_k(t|\theta^2)}{h_k(t|\hat{\delta})} \frac{1}{k} \hat{\delta}^{k/2} \theta^{k-1} d\theta,$$

the last step using the facts that $h_k(t|\theta^2)/h_k(t|\hat{\delta}) \leq 1$ and

$$\int_0^{\sqrt{\hat{\delta}-\varepsilon}} \frac{1}{k} \hat{\delta}^{k/2} \theta^{k-1} d\theta = \left(1 - \frac{\varepsilon}{\hat{\delta}}\right)^{k/2} = \left(1 - \frac{\varepsilon}{\sqrt{2k}\; z_p(1 + o(1))}\right)^{k/2} \to 0.$$

But, on $(\sqrt{\hat{\delta} - \varepsilon}, \sqrt{\hat{\delta}})$, $h_k(t_p|\theta^2)/h_k(t_p, \hat{\delta})$ converges to 1 uniformly (employing again the asymptotic normality), from which it can be concluded that

$$\lim_{k\to\infty} \frac{\int h_k(t|\theta^2) g_1^*(\theta) d\theta}{h_k(t|\hat{\delta})} = 1.$$

Thus

$$\lim_{k\to\infty} \underline{B}^* \leq \lim_{k\to\infty} \frac{h_k(t|0)}{\int h_k(t|\theta^2) g_1^*(\theta) d\theta}$$
$$= \lim_{k\to\infty} \frac{h_k(t|0)}{h_k(t|\hat{\delta})},$$

which establishes the desired inequality and completes the proof.

References

Berger, J. (1985). *Statistical Decision Theory and Bayesian Analysis.* New York: Springer.

Berger, J. (1986). Are P-values reasonable measures of accuracy. In *Pacific Statistical Congress*, Ed., I.S. Francis et al., pp. 21-27. North-Holland, Amsterdam.

Berger, J. (1992). A comparison of minimal Bayesian tests of precise hypotheses. *Rassegna di Metodi Statistici ed Applicazioni*, **7**, Pitagora Editrice, Bologna.

Berger, J. & Bernardo, J.M. (1992). On the development of the reference prior method. *Bayesian Statistics 4*, Ed., J.M. Bernardo, et. al.. Oxford University Press, London.

Berger, J. & Delampady, M. (1987). Testing precise hypotheses (with discussion). *Statist. Science* **2**, 317-352.

Berger, J. & Mortera, J. (1991). Interpreting the stars in precise hypothesis testing. *Int. Statist. Review.* **59**, 337-353.

Berger, J. & Sellke, T. (1987). Testing a point null hypothesis: The irreconcilability of P-values and evidence (with discussion). *J. Am. Statist. Assoc.* **82**, 112-139.

Bertolino, F., Piccinato, L. & Racugno, W. (1990). A marginal likelihood approach to analysis of variance. *The Statistician* **39**, 415-424.

Delampady, M. (1989a). Lower bounds on Bayes factors for interval hypotheses. *J. Am. Statist. Assoc.* **84**, 120-124.

Delampady, M. (1989b). Lower bounds on Bayes factors for invariant testing situations. *J. Mult. Anal.* **28**, 227-246.

Delampady, M. & Berger, J. (1990). Lower bounds on posterior probabilities for multinomial and chi-squared tests. *Ann. Statist.* **18**, 1295-1316.

Edwards, W., Lindman, H. & Savage, L.J. (1963). Bayesian statistical inference for psychological research. *Psychol. Rev.* **70**, 193-242. Reprinted (1984) in *Robustness of Bayesian Analysis*, Ed., J.B. Kadane, pp. 1-62. North-Holland, Amsterdam.

Jeffreys, H. (1961). *Theory of Probability*, 3rd ed. Oxford Univ. Press, London.

Lindley, D.V. (1957). A statistical paradox. *Biometrika* **44**, 187-192.

Moreno, E. & Cano, J.A. (1989). Testing a point null hypothesis: Asymptotic robust Bayesian analysis with respect to the priors given on a subsigma field. *Int. Statist. Rev.* **57**, 221-232.

Sivaganesan, S. & Berger, J. (1989). Ranges of posterior measures for priors with unimodal contaminations. *Ann. Statist.* **17**, 868-889.

Figure 1a. Graph of $B(x, g_{1,r}, g_{2,k})$ as a function of r and $k^*=1/k$, when n=4, s=1, and P-value = 0.05.

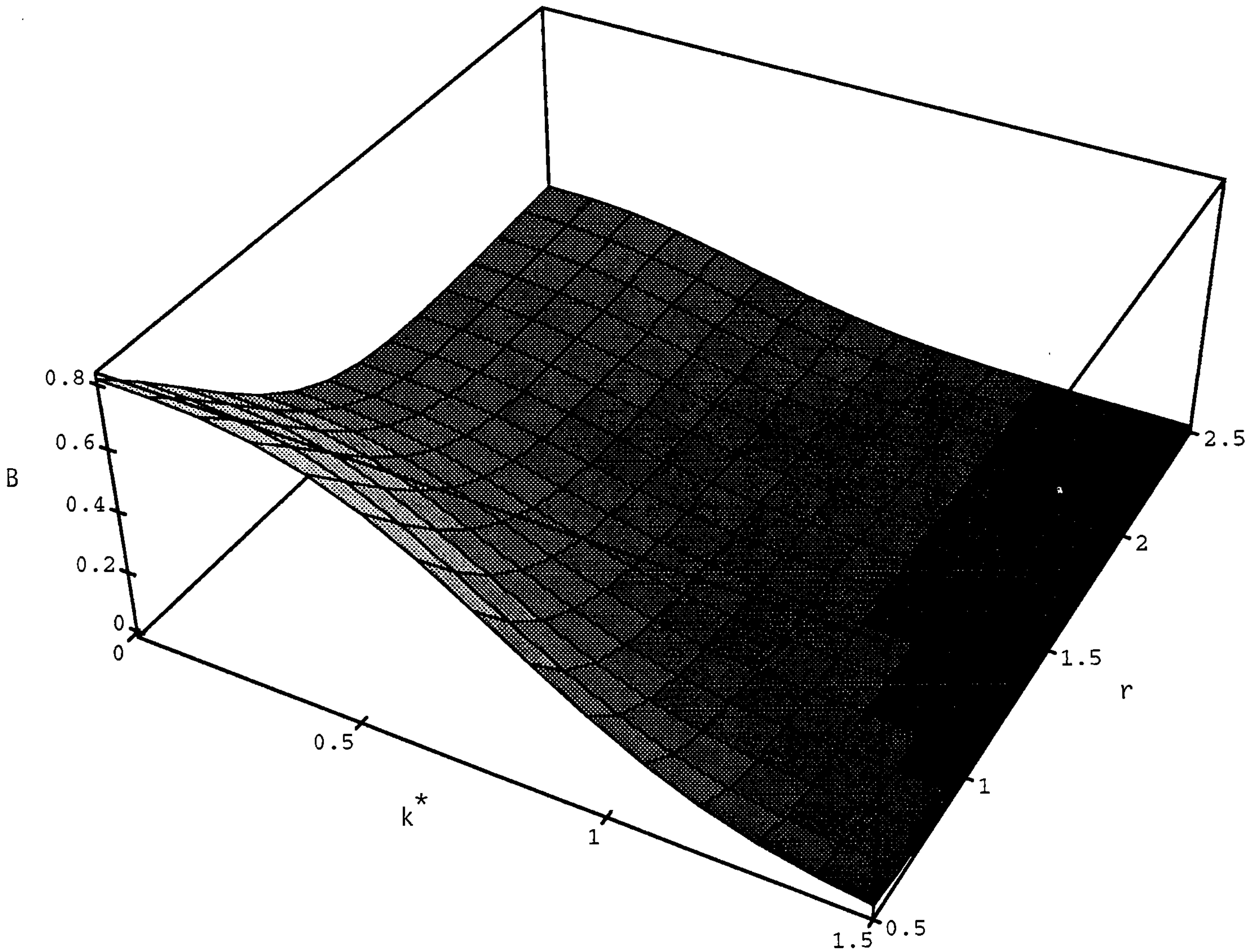

Figure 1b. Contours of $B(\underset{\sim}{x}, g_{1,r}, g_{2,k})$ as a function of r and $k^*=1/k$, when n=4, s=1, and P-value = 0.05.