INTERPRETING THE STARS IN PRECISE HYPOTHESIS TESTING[1]

by

James O. Berger and Julia Mortera
Purdue University Università di Roma

Technical Report # 90–38C

Department of Statistics
Purdue University

August, 1990
Revised October, 1990

[1]Research supported by the National Science Foundation, Grant DMS–8923071 and by MURS projects.

△

INTERPRETING THE STARS IN PRECISE HYPOTHESIS TESTING

by

James O. Berger and Julia Mortera
Purdue University Università di Roma

ABSTRACT

The problem of testing a precise null hypothesis is considered when available information is limited to knowledge of either the P-value or that the P-value is in some interval (e.g., the classical one or two "stars"). Because of the recognized conflict between classical and Bayesian measures of evidence in testing a precise null hypothesis, the interpretation of a P-value or of "stars", from a Bayesian perspective, is explored. This is done by treating the P-value or the "stars" as the data, and computing corresponding posterior probabilities or Bayes factors. Of particular interest are lower bounds on these measures over wide classes of prior distributions. Comparisons are also made between classical meta-analysis techniques for combining many tests of statistical significance and lower bounds on Bayes factors and posterior probabilities.

Key words: P-values; Point null hypothesis; Bayes factor; Posterior probability; robust Bayesian analysis; meta-analysis.

1. INTRODUCTION

1.1 The Problem

Let X be a random variable having density $f(x|\theta)$, where θ is an unknown parameter assuming values in $\Theta \subset \mathbf{R}$. Suppose that the observed data is of the type $x \in A$, for some set A and that it is desired to test $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$. (In Berger and Delampady (1987) it was shown that point nulls are a good approximation to interval nulls of the form $H_0 : |\theta - \theta_0| \leq \varepsilon$, providing $\varepsilon < \frac{1}{2}\sigma_{\hat{\theta}}$, where $\sigma_{\hat{\theta}}$ is the standard error associated with the m.l.e. $\hat{\theta}$. Thus the results of this paper also hold, approximately, for such interval nulls.)

The set A typically arises as the set of x corresponding to a specified P-value (observed significance level), or corresponding to knowledge that the P-value is in some interval (α, α'). The latter arises, for example, when all that is available is a report of "statistical significance " at, say, the "one star" level ($0.01 < P$-value < 0.05) or the "two star" level ($0.001 < P$-value < 0.01). These are equivalent to observing $x \in A$, for

$$A = \{x : q_1 \leq |T(X)| \leq q_2\},$$

where $T(X)$ is the test statistic used in defining the P-value.

The particular interest in studying Bayesian measures of evidence here is that such measures, for the case of given x, are well known to indicate substantially less evidence against H_0 than is commonly thought to be implied by the corresponding P-value. (See, for instance, Edwards, Lindman, and Savage (1963), Berger and Sellke (1987), and Berger and Delampady (1987).) This fact is often strikingly demonstrated by showing that the lower bound on the Bayes factor, over *all* "sensible" prior distributions, is much larger than the corresponding P-value. For instance, a typical result in this regard is that if x is observed for which the P-value is 0.05, then the evidence for H_0 (the Bayes factor) is *at least* 1: $2\frac{1}{2}$ for *any* unimodal, symmetric prior. Of course, P-values and Bayes factors are not measured on the same scale, but many consider a P-value of 0.05 to be strong evidence against H_0, while few would consider odds of 1:2 $\frac{1}{2}$ to be more than mild evidence against H_0.

Whether or not the same phenomenon holds when one is only given A (and not x itself) is the primary question that is studied here. The main tool is again the computation of lower bounds on Bayes factors over wide classes of prior distributions. Rather surprisingly, essentially closed form answers can be obtained for common observation sets A, allowing easy conversion of P-values or "stars" to, at least, lower bounds on Bayes factors. While these lower bounds cannot be thought of as complete answers to the problem, they are at least as "objective" as P-values and are typically considerably more representative of the actual evidence against H_0.

The problem of combining P-values or "stars" from several studies is also considered. Lower bounds on Bayes factors are proposed as alternatives to meta-analytic P-values.

Several works have previously considered lower bounds on Bayes factors in hypothesis testing. These include Edwards, Lindman, and Savage (1963), Berger (1985, 1986), Berger and Delampady (1987), Berger and Sellke (1987), Casella and Berger (1987), Delampady (1989a, 1989b), and Delampady and Berger (1990). Berger and Sellke (1987) is the only previous paper that contains any results concerning observational sets, A.

Other relevant work on Bayes factors and on their relationship to P-values includes

Lindley (1957, 1977), Good (1958, 1967, 1983, 1984, 1986), Jeffreys (1961), Pratt (1965), Lempers (1971), Rubin (1971), Zellner (1971, 1984), DeGroot (1973), Dickey (1977), Smith and Spiegelhalter (1980), Shafer (1982), Hill (1982), and Kass and Vaidyanathan (1991).

1.2 Measures of Evidence

The three measures of evidence that will be considered are as follows:

P-value: Let $T(X)$ be a test statistic, large absolute values of which are considered to be evidence against H_0. If x_0 is observed, the P-value is

$$p = P(|T(X)| \geq |T(x_0)| \; |\theta_0). \tag{1.1}$$

Bayes factor: Let $g(\theta)$ be a density (with respect to Lebesgue measure) on $\{\theta : \theta \neq \theta_0\}$. The Bayes factor for H_0 versus H_1, given the observational set A, is

$$B_g(A) = \frac{P(A|H_0)}{P(A|H_1)} = \frac{P(A|\theta_0)}{m_g(A)}, \tag{1.2}$$

where

$$\begin{aligned} P(A|\theta) &= \int_A f(x|\theta)dx, \\ m_g(A) &= \int_{\theta \neq \theta_0} P(A|\theta)g(\theta)d\theta. \end{aligned} \tag{1.3}$$

When $A = \{x\}, P(A|\theta)$ is replaced in (1.2) and (1.3) by the density $f(x|\theta)$.

Posterior probability of H_0: Let π_0 be the prior probability of H_0. Then the posterior probability of H_0, given the set A, is

$$P(H_0|A) = \left[1 + \frac{(1-\pi_0)}{\pi_0}\frac{1}{B_g(A)}\right]^{-1}. \tag{1.4}$$

Bounds on B and P: We will assume that $g(\theta)$, the prior distribution of θ given H_1, is known only to lie within some (large) class, $\mathcal{G}$, of distributions. Of particular interest will then be the lower bounds on (1.2) and (1.4) as g varies over $\mathcal{G}$, defined by

$$\underline{B}(A,\mathcal{G}) = \inf_{g\in\mathcal{G}} B_g(A) = \frac{P(A|\theta_0)}{\sup_{g\in\mathcal{G}} m_g(A)} \tag{1.5}$$

and

$$\underline{P}(A,\mathcal{G}) = \inf_{g\in\mathcal{G}} P(H_0|A) = \left[1 + \frac{(1-\pi_0)}{\pi_0}\frac{1}{\underline{B}(A,\mathcal{G})}\right]^{-1}. \tag{1.6}$$

As an example of the type of result that will be established, suppose that – in a normal mean testing problem – the only information available is "one-star" significance, i.e., that $0.01 < p < 0.05$. The lower bound on the Bayes factor over the class of all symmetric (about θ_0) and unimodal g is then $1/4$, indicating that the odds for H_0 are no worse than $1:4$.

1.3 Preview

Section 2 discusses the classes of prior distributions and the types of sets A that will be considered, and presents some needed analytical tools. The lower bounds on the Bayesian measures of evidence for the class of all prior distributions are given in Section 3, and for the class of all symmetric unimodal distributions in Section 4. Lower bounds on the Bayesian measures for combining tests of statistical significance are developed in Section 5, as an alternative to classical meta-analysis techniques. Section 6 gives some closing remarks, as well as a table which summarizes results on the subject.

2. PRELIMINARIES

2.1 Classes of Prior Distributions

The following classes of prior distributions for g will be considered:

$\mathcal{G}_A = \{\text{all distributions}\}$,

$\mathcal{G}_S = \{\text{all symmetric distributions about } \theta_0\}$,

$\mathcal{G}_U = \{\text{all unimodal distributions with mode at } \theta_0\}$,

$\mathcal{G}_{US} = \{\mathcal{G}_S \cap \mathcal{G}_U = \text{all symmetric unimodal distributions with mode at } \theta_0\}$,

$\mathcal{G}_{Un} = \{\text{all symmetric uniform distributions on } [\theta_0 - r, \theta_0 + r], r \in \mathbb{R}\}$.

Lemmas 2.1 and 2.2, whose proofs are given in Appendix I, give simple formula for computing (1.5) when using classes $\mathcal{G}_A$ and $\mathcal{G}_{US}$. We will repeatedly use these lemmas.

Lemma 2.1: *For the class $\mathcal{G}_A$ and any set A,*

$$\underline{B}(A, \mathcal{G}_A) = \frac{P(A|\theta_0)}{P(A|\theta^*)}, \tag{2.1}$$

where θ^* is that θ which maximizes $P(A|\theta)$. (Note that (2.1) is the likelihood ratio of θ_0 to θ^*, based on "data" A.)

Lemma 2.2: *For the class $\mathcal{G}_{US}$ and any set A,*

$$\underline{B}(A, \mathcal{G}_{US}) = \frac{P(A|\theta_0)}{\sup_r \frac{1}{2r} H_A(r)}, \tag{2.2}$$

where

$$H_A(r) = \int_{\theta_0 - r}^{\theta_0 + r} P(A|\theta) d\theta. \tag{2.3}$$

2.2 Observational Sets A

In Sections 3 and 4, attention will be confined to the canonical normal mean problem, where

$$X \sim \mathcal{N}(\theta, 1), \;\; p = P - \text{value} \;\equiv 2[1 - \Phi(|x - \theta_0|)], \tag{2.4}$$

Φ being the standard normal c.d.f.; any univariate normal problem with known variance can obviously be transformed into this canonical form. For this problem, typical observational sets, A, of interest include:

a) $A_0 = \{\theta_0 - q, \theta_0 + q\}$, a two point set where $q \equiv \Phi^{-1}(1 - \frac{1}{2}p)$, Φ^{-1} being the inverse standard normal c.d.f.

b) $A_1 = (\theta_0 + q_1, \theta_0 + q_2)$, corresponding to $\alpha_2 < p < \alpha_1$ and sign of x positive; thus $(\theta_0 + q_i)$ is the $(1 - \frac{1}{2}\alpha_i)^{th}$ quantile of X under H_0. (We assume that a two-sided P-value was calculated.)

c) $A_2 = (\theta_0 - q_2, \theta_0 - q_1)$, corresponding to $\alpha_2 < p < \alpha_1$ and sign of x negative; thus $(\theta_0 - q_i)$ is the $(\frac{1}{2}\alpha_i)^{th}$ quantile of X under H_0.

d) $A_3 = A_1 \cup A_2 = (\theta_0 - q_2, \theta_0 - q_1) \cup (\theta_0 + q_1, \theta_0 + q_2)$, corresponding to $\alpha_2 < p < \alpha_1$, sign of x not known.

The motivations for b) and c) are that, in statistical reports, one is sometimes told the "direction" of the effect (i.e., the sign of x) in addition to just the number of "stars." Interestingly, for symmetric (about θ_0) priors g, it turns out to be irrelevant whether or not one knows the sign of x, as Lemma 2.3 shows. Lemma 2.4 establishes a result that will be used several times later, namely that if A_3 is the observational set then the Bayes factor for any prior is equal to the Bayes factor for a symmetrized version of that prior.

Lemma 2.3: *If $g \in \mathcal{G}_S$ and X has density $f(|x-\theta|)$, then*

$$B_g(A_3) = B_g(A_1) = B_g(A_2).$$

Proof. Write $g(\theta)$ as $h(|\theta - \theta_0|)$. Then

$$\begin{aligned} m_g(\theta_0 + y) &= \int f(|\theta_0 + y - \theta|)h(|\theta - \theta_0|)d\theta \\ &= \int f(|\theta - \theta_0 - y|)h(|\theta - \theta_0|)d\theta \\ &= \int f(|\theta_0 - \eta - y|)h(|\theta_0 - \eta|)d\eta \\ &= m_g(\theta_0 - y). \end{aligned}$$

Hence $m_g(A_1) = m_g(A_2)$ (slightly abusing notation by defining $m_g(A) = \int_A m_g(x)dx$). Also, it is clear that $P(A_1|\theta_0) = P(A_2|\theta_0)$. The result follows. □

Lemma 2.4: *If X has density $f(|x-\theta|)$, then*

$$B_g(A_3) = B_{g^*}(A_3),$$

where

$$g^*(\theta) = \frac{1}{2}[g(\theta) + g(2\theta_0 - \theta)] \in \mathcal{G}_S.$$

Proof. Note first that

$$\begin{aligned} &m_g(x) + m_g(2\theta_0 - x) \\ &= \int [f(|x-\theta|) + f(|2\theta_0 - x - \theta|)g(\theta)d\theta \\ &= \int f(|x-\theta|)g(\theta)d\theta + \int f(|x-\eta|)g(2\theta_0 - \eta)d\eta \\ &= 2\int f(|x-\theta|)g^*(\theta)d\theta \\ &= 2m_{g^*}(x). \end{aligned}$$

Thus

$$m_g(A_3) = m_g(A_1) + m_g(A_2) = 2m_{g^*}(A_1).$$

Since $P(A_1|\theta_0) = P(A_2|\theta_0)$, it follows that

$$
\begin{aligned}
B_g(A_3) &= \frac{P(A_1|\theta_0) + P(A_2|\theta_0)}{m_g(A_3)} \\
&= \frac{2P(A_1|\theta_0)}{2m_{g^*}(A_1)}. \\
&= B_{g^*}(A_1).
\end{aligned}
$$

Since g^* is symmetric about θ_0, Lemma 2.3 yields the conclusion. □

3. LOWER BOUNDS ON BAYES FACTORS OVER ALL PRIORS

This section is concerned with finding $\underline{B}(A, \mathcal{G}_A)$, the lower bound on the Bayes factor when A is the observation set and *any* prior g is allowed. The formula for $\underline{B}(A, \mathcal{G}_A)$, in general, is given by (2.1). In the following sections, we specialize to the various A discussed in Section 2, and to the canonical normal situation defined by (2.4). For simplicity, results will be stated only for the case $\theta_0 = 0$.

3.1 Lower Bounds When p Is Given

When p is given and $\theta_0 = 0$ in the canonical normal problem, the observation set is the two-point set $A_0 = \{-q, q\}$. Thus (2.1) yields

$$\underline{B}(A_0, \mathcal{G}_A) = \frac{2\varphi(q)}{[\varphi(q - \theta^*) + \varphi(q + \theta^*)]}, \tag{3.1}$$

where φ is the standard normal density and θ^* is the point maximizing the denominator.

An extremely accurate approximation to θ^*, and hence to $\underline{B}(A_0, \mathcal{G}_A)$, can be given. Indeed, consider

$$\tilde{\theta}^* = q - 2q \exp\{-2q^2\}, \tag{3.2}$$

$$\underline{\tilde{B}}(A_0, \mathcal{G}_A) = \frac{2\exp\{-\frac{1}{2}q^2\}}{1 + \exp\{-2q^2\} + 2q^2 \exp\{-4q^2\}}. \tag{3.3}$$

Lemma A.1 in Appendix II shows that $\tilde{\theta}^*$ and $\underline{\tilde{B}}$ are asymptotically (large q, i.e., small p) correct. Note that θ^* will typically be very close to q (see also Table 1).

Numerical comparison between the actual $\underline{B}(A_0, \mathcal{G}_A)$ and the approximation in (3.3) showed that the latter was accurate up to the 5th significant digit for $q > 1.5$ and to the 2nd significant digit for $1 < q \leq 1.5$. Also, the expression in (3.3) is extremely accurate even if the last term in the denominator is dropped.

In Table 1, comparison between the P-value, $\underline{B}(A_0, \mathcal{G}_A)$, and the corresponding $\underline{P}(A_0, \mathcal{G}_A)$ for $\pi_0 = 1/2$ is given. Note that, for $p = 0.05$, the lower bound on the Bayes factor corresponds to odds of roughly 3:10 for the null hypothesis H_0. Overall, $\underline{B}$ appears to be between 5 and 9 times larger than the corresponding p.

Table 1. Comparison between $p, \underline{B}$, and $\underline{P}$ when p is given and $\mathcal{G} = \{\text{all priors}\}$.

P-value (p)	q	$\underline{B}(A_0, \mathcal{G}_A)$	$\underline{P}(A_0, \mathcal{G}_A)$	θ^*
.001	3.291	.0089	.0088	3.291
.01	2.576	.0725	.0676	2.576
.05	1.96	.2928	.2265	1.958
.1	1.645	.5145	.3397	1.630

If p is given and the sign of x is known, the lower bounds are the same as those given in Berger and Sellke (1987), since learning p is equivalent in this case to learning x.

3.2 Lower Bounds When $\alpha_2 < p < \alpha_1$ and the Sign of X is Unknown

Suppose that, instead of observing p, one only learns that $\alpha_2 < p < \alpha_1$; in the canonical normal problem this is equivalent to observing that

$$x \in A_3 = (-q_2, -q_1) \cup (q_1, q_2). \tag{3.4}$$

Lemma 2.1 then implies that

$$\underline{B}(A_3, \mathcal{G}_A) = P(A_3|0)/P(A_3|\theta^*), \tag{3.5}$$

where θ^* is that θ which maximizes

$$P(A_3|\theta) = \Phi(\theta - q_1) - \Phi(\theta - q_2) + \Phi(\theta + q_2) - \Phi(\theta + q_1). \tag{3.6}$$

As in the case where p is given, an accurate approximation to θ^* can be found. Indeed,

$$\tilde{\theta}^* = \frac{1}{2}(q_1 + q_2) - (q_2 - q_1)^{-1} e^{-q_1 q_2}(e^{-q_1^2} - e^{-q_2^2}) \tag{3.7}$$

is shown in Lemma A.2 of Appendix II to be asymptotically correct as $q_1 \to \infty$. Use of $\tilde{\theta}^*$ in (3.5) also yields a very accurate approximation to the lower bound on the Bayes factor, a bound to be denoted by $\tilde{\underline{B}}(A_3, \mathcal{G}_A)$; Table 2 presents exact and approximate Bayes factors for a variety of quantiles.

Table 2: $\underline{B}(A_3, \mathcal{G}_A)$ ($\underline{\tilde{B}}(A_3, \mathcal{G}_A)$) for various $A_3 = (-q_2, -q_1) \cup (q_1, q_2)$.

	q_2 1.01	1.26	1.51	1.76	2.01	2.26	2.51	2.76	3.01
q_1 1.00	1.3513 (1.6173)	1.1358 (1.2207)	0.9645 (0.9917)	0.8270 (0.8357)	0.7146 (0.7193)	0.6281 (0.6291)	0.5581 (0.5585)	0.5031 (0.5032)	0.4602 (0.4603)
1.25		0.9460 (0.9697)	0.7925 (0.7989)	0.6678 (0.6695)	0.5671 (0.5675)	0.4867 (0.4869)	0.4235 (0.4235)	0.3742 (0.3742)	0.3362 (0.3362)
1.50			0.6512 (0.6525)	0.5364 (0.5367)	0.4441 (0.4442)	0.3713 (0.3713)	0.3147 (0.3147)	0.2713 (0.2713)	0.2383 (0.2383)
1.75				0.4297 (0.4297)	0.3453 (0.3453)	0.2798 (0.2798)	0.2299 (0.2299)	0.1924 (0.1924)	0.1645 (0.1645)
2.00					0.2681 (0.2681)	0.2096 (0.2096)	0.1661 (0.1661)	0.1342 (0.1342)	0.1110 (0.1110)
2.25						0.1573 (0.1573)	0.1197 (0.1197)	0.0928 (0.0928)	0.0738 (0.0738)
2.50							0.0868 (0.0868)	0.0642 (0.0642)	0.0487 (0.0487)
2.75								0.0450 (0.0450)	0.0324 (0.0324)
3.00									0.0219 (0.0219)

Of course, the basic interest here again resides in the comparison between P-values and $\underline{B}$. Table 3 presents such comparison for various "stars": included is $\underline{P}(A_3, \mathcal{G}_A)$, the lower bound on the posterior probability when $\pi_0 = \frac{1}{2}$, and θ^*.

Table 3. Comparison between $p, \underline{B}$, and $\underline{P}$ for observation set A_3 and $\mathcal{G} = \{\text{all priors}\}$.

Stars	P-value	(q_1, q_2)	$\underline{B}(A_3, \mathcal{G}_A)$	$\underline{P}(A_3, \mathcal{G}_A)$	θ^*
$*$	$.01 < p < .05$	(1.96, 2.576)	.1653	.1419	2.267
	$p < .05$	$(1.96, \infty)$	.05	.0476	∞
$**$	$.001 < p < .01$	(2.576, 3.291)	.0322	.0312	2.933
	$p < .01$	$(2.576, \infty)$	.01	.0099	∞
$***$	$.0001 < p < .001$	(3.291, 3.891)	.0038	.0038	3.591
	$p < .001$	$(3.891, \infty)$	.001	.001	∞

Observe that, for "$*$", "$**$", and "$* * *$" significance, $\underline{B}$ is 3 to 4 times larger than the upper bound of the P-value interval, implying substantially less doubt of H_0 than is commonly associated with "stars". In contrast, the lower bounds $\underline{B}$ associated with the observations $p < \alpha$ (or $A_3 = (-\infty, -q_1) \cup (q_1, \infty)$) exactly equal the upper bound on the P-value. Indeed, by letting $\theta^* \to \infty$, one can show that $\underline{B}(A_3, \mathcal{G}_A)$ equals the P-value for such A_3. (Priors that are point masses at large θ^* are obviously not sensible; it is of interest, however, that near uniform priors also yield the same $\underline{B}$.)

3.3 Lower Bounds When $\alpha_2 < p < \alpha_1$ and the Sign of X is Known

Without loss of generality assume that the known sign of X is positive, so observing $\alpha_2 < p < \alpha_1$ corresponds to observing that $x \in A_1$, with $A_1 = (q_1, q_2)$. For the canonical normal problem it follows that

$$P(A_1|\theta) = \Phi(\theta - q_1) - \Phi(\theta - q_2). \tag{3.9}$$

It is then trivial to prove that the θ that maximizes $P(A_1|\theta)$ is

$$\theta^* = (q_1 + q_2)/2 \tag{3.10}$$

and thus

$$\underline{B}(A_1, \mathcal{G}_A) = \frac{[\Phi(q_2) - \Phi(q_1)]}{\left[2\Phi\left(\frac{1}{2}(q_2 - q_1)\right) - 1\right]}. \tag{3.11}$$

Values of $\underline{B}(A_1, \mathcal{G}_A)$ are typically about half those of $\underline{B}(A_3, \mathcal{G}_A)$ (since the mass of the half of A_3 that is "opposite" θ^* is essentially wasted), and hence we do not present separate numerical results.

4. LOWER BOUNDS ON BAYES FACTOR OVER THE SYMMETRIC UNIMODAL PRIORS

This section is concerned with finding $\underline{B}(A, \mathcal{G}_{US})$, the lower bound on the Bayes factor when A is the observation set and any symmetric (about θ_0) unimodal prior is allowed. The formula for $\underline{B}(A, \mathcal{G}_{US})$ is given in (2.2) and (2.3). In the following sections we specialize to the various A discussed in Section 2, and to the canonical normal situation defined by (2.4). For simplicity, results will be stated only for the case $\theta_0 = 0$.

By Lemma 2.3, $\underline{B}(A, \mathcal{G}_{US})$ will be the same for $A_1 = (q_1, q_2)$ and for $A_3 = (-q_2, -q_1) \cup (q_1, q_2)$. We will formally present the results for the A_3 case. Also, by Lemma 2.4, the bounds obtained for A_3 are the same whether $\mathcal{G}_{US}$ or $\mathcal{G}_U$, the class of all unimodal priors, is used.

Lemma 2.2 yields

$$\underline{B}(A_3, \mathcal{G}_{US}) = \frac{P(A_3|0)}{\sup_r \frac{1}{2r} H_{A_3}(r)}, \tag{4.1}$$

where

$$P(A_3|0) = 2(\Phi(q_2) - \Phi(q_1))$$

and (see Appendix III)

$$H_{A_3}(r) = 2\left[\Psi_1(q_2, r) - \Psi_1(q_1, r) + \Psi_2(q_2, r) - \Psi_2(q_1, r)\right], \tag{4.2}$$

with (again letting φ denote the standard normal density)

$$\Psi_1(q,r) = (q+r)\Phi(q+r) - (q-r)\Phi(q-r),$$

$$\Psi_2(q,r) = \varphi(q+r) - \varphi(-q+r).$$

A closed form is not available for r^*, the value maximizing $\frac{1}{2r}H_{A_3}(r)$. However, r^* can be computed from the following iterative expression (explained in Appendix III):

$$r_{i+1} = 2\log\left[-\Psi_2(r_i)/\Psi_1(r_i)\right], \tag{4.3}$$

where

$$\Psi_1(r) = q_2\left[\Phi(q_2+r) - \Phi(q_2-r)\right] - q_1\left[\Phi(q_1+r) - \Phi(q_1-r)\right],$$

$$\Psi_2(r) = \varphi(q_2)\left[e^{-q_2 r} - e^{q_2 r}\right] - \varphi(q_1)\left[e^{-q_1 r} - e^{q_1 r}\right].$$

At each step, including choice of the initial r_0, r_i should be constrained to lie within the interval $\left(\frac{1}{2}[q_1+q_2], q_2+4\right)$. Convergence is usually achieved within three iterations of (4.3).

Table 4 presents, for various interval P-values, the lower bounds on the Bayes factors and posterior probabilities. The values of r^* that result from (4.3) are also recorded. Recall that the prior resulting in $\underline{B}$ and $\underline{P}$ is the Uniform $[-r^*, r^*]$ prior.

Table 4. Comparison between $p, \underline{B}$, and $\underline{P}$ for observation sets A_3 (or A_1) and $\mathcal{G} = \{$all symmetric, unimodal priors$\}$.

Stars	P-value	(q_1, q_2)	$\underline{B}(A_3, \mathcal{G}_{US})$	$\underline{P}(A_3, \mathcal{G}_{US})$	r^*
$*$	$.01 < p < .05$	$(1.96, 2.576)$	.2532	.202	3.206
	$p < .05$	$(1.96, \infty)$	.05	.0476	∞
$**$	$.001 < p < .01$	$(2.576, 3.291)$	.0591	.0558	4.058
	$p < .01$	$(2.576, \infty)$	.01	.0099	∞
$***$	$.0001 < p < .001$	$(3.291, 3.891)$	.0081	.0081	4.841
	$p < .001$	$(3.891, \infty)$	.001	.001	∞

Whereas, in Table 3, $\underline{B}$ was typically 3 to 4 times the size of the corresponding "star" P-value, here $\underline{B}$ is 5 to 8 times larger. We feel that these numbers are more reasonable than those for $\mathcal{G}_A$, since the restriction to $\mathcal{G}_{US}$ (or, equivalently for A_3, to $\mathcal{G}_U$) eliminates unreasonable priors.

Once again, by letting $r \to \infty$, it can be shown that, when A is defined by $p \leq \alpha$, then $\underline{B}(A, \mathcal{G}_{US}) = \alpha$. While interesting, this should not be taken as proof that α is sensible in such situations; opinions about θ under H_1 will rarely be uniform over a huge range.

The limiting case of the above results in which the P-value is given, i.e., where $A_0 = \{-q, q\}$, can be reduced by Lemma 2.3 to the case in which x is given. The ensuing computations for $\underline{B}$ and $\underline{P}$ are extensively discussed in Berger and Sellke (1987).

5. COMBINED REPORTED SIGNIFICANCE LEVELS

5.1 Introduction

In this section, a comparison is made between classical meta-analysis techniques for combining many tests of statistical significance and lower bounds on Bayes factors and posterior probabilities. (See Good (1958) and Hedges and Olkin (1985) for an introduction and references.) The observations in this case will be assumed to be either a vector of P-values, $\tilde{p} = (p_1, \ldots, p_m)$, or a combination of m_1 P-values, $\tilde{p}_1 = (p_1, \ldots, p_{m_1})$, and m_2 P-value intervals, $\alpha_j < p_j < \beta_j$, for $j = m_1 + 1, \ldots, m_1 + m_2$. The $m = m_1 + m_2$ p-values (or intervals) are assumed to arise from independent studies concerning θ, so that the likelihood for θ is

$$f(\theta) = \prod_{i=1}^{m} P_i(A_i|\theta), \tag{5.1}$$

where A_i is the observation set corresponding to the i^{th} P-value and P_i is the probability (or density) function from the i^{th} study.

The canonical example that will be considered here is that in which the i^{th} study has observation $X_i \sim \mathcal{N}(\theta, \sigma_i^2)$; as usual we consider testing $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$. Also, we assume that the sign of $(X_i - \theta_0)$ is known for each experiment, i.e. that the "direction" of the effect is reported. Then, being given p_i corresponds to knowledge of $(x_i - \theta_0)/\sigma_i$, while being given $\alpha_j < p_j < \beta_j$ corresponds to knowledge that

$$q_{1j} < (x_j - \theta_0)/\sigma_j < q_{2j}. \tag{5.2}$$

Note that, when "nonsignificance" is reported (i.e., $p_j > \alpha_j$), the "direction" of the effect may not be stated. But $p_j > \alpha_j$ corresponds to a region of the form (5.2) with $q_{1j} = -q_{2j}$, so this situation is accommodated in what follows. Also, the general case in which the direction of the effect is unknown could be handled with minor modifications of the following formulas.

For the above normal situation, the likelihood for θ becomes (after some algebra),

$$\begin{aligned} f(\theta) &= \left[\prod_{i=1}^{m_1} \frac{1}{\sigma_i}\varphi\left(\frac{x_i - \theta}{\sigma_i}\right)\right] \prod_{j=m_1+1}^{(m_1+m_2)} \left[\Phi\left(q_{2j} + \frac{\theta_0 - \theta}{\sigma_j}\right) - \Phi\left(q_{1j} + \frac{\theta_0 - \theta}{\sigma_j}\right)\right] \\ &= K(\underset{\sim}{\sigma})\varphi\left(\frac{\tilde{x} - \theta}{\tilde{\sigma}}\right) \prod_{j=m_1+1}^{(m_1+m_2)} \left[\Phi\left(q_{2j} + \frac{\theta_0 - \theta}{\sigma_j}\right) - \Phi\left(q_{1j} + \frac{\theta_0 - \theta}{\sigma_j}\right)\right], \end{aligned} \tag{5.3}$$

where

$$\tilde{x} = \tilde{\sigma}^2 \sum_{i=1}^{m_1} \frac{x_i}{\sigma_i^2},\ \tilde{\sigma}^2 = \left(\sum_{i=1}^{m_1} \frac{1}{\sigma_i^2}\right)^{-1}, \tag{5.4}$$

$$K(\underset{\sim}{\sigma}) = \frac{\tilde{\sigma} \exp\left\{-\frac{1}{2} \sum_{i=1}^{m_1} (x_i - \tilde{x})^2/\sigma_i^2\right\}}{(2\pi)^{(m_1-1)/2} \left[\prod_{i=1}^{m_1} \sigma_i\right]}.$$

The following scenarios involving $\underset{\sim}{\sigma} = (\sigma_1, \ldots, \sigma_m)$ will be considered or discussed:

Case (i): $\underset{\sim}{\sigma}$ known;

Case (ii): $\underset{\sim}{\sigma}$ unknown, and no prior information about $\underset{\sim}{\sigma}$ is available;

Case (iii): $\underset{\sim}{\sigma}$ is known up to a multiplicative constant – this arises typically when the sample sizes, n_i, in each study are known, and the individual data in the studies have a common variance σ^2: then

$$\sigma_i^2 = \sigma^2/n_i, \qquad i = 1, \ldots, m. \tag{5.5}$$

For this case, (5.3) becomes, after defining $n = \sum_{i=1}^{m_1} n_i$ and making the transformation

$$\eta = \sqrt{n}(\theta - \theta_0)/\sigma, \tag{5.6}$$

$$\begin{aligned} f^*(\eta) = K(\underset{\sim}{\sigma})\varphi\left(\left[\sum_{i=1}^{m_1}\sqrt{\frac{n_i}{n}}\frac{(x_i-\theta_0)}{\sigma_i}\right]-\eta\right) \\ \times \prod_{j=m_1+1}^{(m_1+m_2)}\left[\Phi\left(q_{2j}-\sqrt{\frac{n_j}{n}}\,\eta\right)-\Phi\left(q_{1j}-\sqrt{\frac{n_j}{n}}\,\eta\right)\right]. \end{aligned} \tag{5.7}$$

One can then use $f^*(\eta)$ in the computation of lower bounds on the Bayes factor, noting that $\eta = 0$ corresponds to the null hypothesis $\theta = \theta_0$, and that all other quantities in (5.7) are known (with the exception of $K(\underset{\sim}{\sigma})$, which will always cancel in our Bayes factor computations).

5.2 Lower Bounds Over $\mathcal{G}_A$

The lower bound on B over all priors is still given by Lemma 2.1, namely (letting $\underset{\sim}{A} = (A_1, \ldots, A_{m_1+m_2})$)

$$\underline{B}(\underset{\sim}{A}, \mathcal{G}_A) = f(\theta_0)/f(\theta^*), \tag{5.8}$$

where θ^* maximizes $f(\theta)$. In general, determining θ^* can only be done numerically. For the case of independent studies, however, a lower bound on $\underline{B}$ is available. We state this first in generality, and then specialize to the normal problem.

Lemma 5.1: *Suppose* $f(\theta) = \prod_{i=1}^{m} f_i(\theta)$, *where* $f_i(\theta)$ *is the likelihood,* $P_i(A_i|\theta)$, *from the* i^{th} *study. Then*

$$\begin{aligned} \underline{B}(\underset{\sim}{A}, \mathcal{G}_A) &= f(\theta_0)/\sup_{\theta} f(\theta) \\ &> \prod_{i=1}^{m} \underline{B}_i(A_i, \mathcal{G}_A), \end{aligned} \tag{5.9}$$

where $\underline{B}_i$ *is the lower bound on the Bayes factor from the* i^{th} *study alone.*

Proof: This follows immediately from the fact that

$$\sup_{\theta} f(\theta) \leq \prod_{i=1}^{m} \sup_{\theta} f_i(\theta). \qquad \square$$

Corollary 5.2: *For the canonical normal problem, with* $f(\theta)$ *as in (5.3),*

$$\underline{B}(\underset{\sim}{A}, \mathcal{G}_A) \geq \left[\prod_{i=1}^{m_1} \exp\left\{ -\frac{(x_i - \theta_0)^2}{2\sigma_i^2} \right\} \right] \prod_{j=m_1+1}^{(m_1+m_2)} \frac{[\Phi(q_{2j}) - \Phi(q_{1j})]}{\left[2\Phi\left(\frac{1}{2}(q_{2j} - q_{1j})\right) - 1\right]}. \tag{5.10}$$

Proof: This is immediate from Lemma 5.1, using (3.11). $\square$

Example 1: A meta-analysis is desired of 8 independent studies on θ. Five studies have reported (two-sided) P-values of 0.04, 0.08, 0.05, 0.03, and 0.01, all associated with a positive effect. One study claims only "one-star" significance, one claims "two-star" significance, both with positive effects, and one claims no significance at the "one-star" level (sign of effect not noted).

The first five studies correspond to $(x_i - \theta_0)/\sigma_i$ equal to 2.05, 1.75, 1.96, 2.17, and 2.58, respectively. The sixth study, interpreted as $0.01 < p_6 < 0.05$ with a positive effect, corresponds to (two-sided P-value) $q_{16} = 1.96$, $q_{26} = 2.576$. The seventh study gives $q_{17} = 2.576$, $q_{27} = 3.291$. The eighth study of "no significance" corresponds to $q_{18} = -1.96$, $q_{28} = 1.96$.

Computation of (5.10) yields $\underline{B} \geq 1.70 \times 10^{-8}$. It is of interest to compare this with a classical solution. One popular classical solution (see Hedges and Olkin, 1985, for discussion and alternatives) is the "inverse normal method," which is based on the fact that the test statistic

$$Z = \frac{1}{\sqrt{m}} \sum_{i=1}^{m} \frac{(X_i - \theta_0)}{\sigma_i}$$

is $\mathcal{N}(0,1)$ under H_0. Unfortunately, for studies 6, 7, and 8 it is known only that $q_{1i} \leq (x_i - \theta_0)/\sigma_i \leq q_{2i}$. To compute Z we simply chose $(x_i - \theta_0)/\sigma_i$ to be the midpoint of each interval for these three studies. The value of Z was then $z = 5.55$, resulting in a (two-sided) P-value of 2.86 $\times 10^{-8}$.

Interestingly, the classical P-value is *larger* than $\underline{B}$. As we shall see in Example 2, the reason is probably that $\underline{B}$ is too crude a bound. □

The lower bounds (5.9) and (5.10) can be excessively low, especially when the $f_i(\theta)$ differ substantially. It is thus very desirable to identify situations in which $\underline{B}$ itself can actually be determined. One such is the case (iii) scenario in which $\sigma_i^2 = \sigma^2/n_i$, with the n_i being known. Then, from (5.7) it is clear that

$$\underline{B}(\underset{\sim}{A}, \mathcal{G}_A) = f^*(0)/\sup_{\eta} f^*(\eta)$$

can be computed. An important special case is when $m_1 = m$, in which case this becomes

$$\underline{B}(\underset{\sim}{A}, \mathcal{G}_A) = \exp\left\{-\frac{1}{2}\left[\sum_{i=1}^{m} \sqrt{\frac{n_i}{n}} \frac{(x_i - \theta_0)}{\sigma_i}\right]^2\right\}. \tag{5.11}$$

Example 2: Four studies produced (two-sided) P-values of 0.05, 0.001, 0.1, and 0.005, all with an effect in the positive direction. The sample sizes were 50, 100, 150, and 200, respectively. Here the $(x_i - \theta_0)/\sigma_i$ are 1.96, 3.29, 1.645, and 2.81, respectively, and $n = 500$.

Calculation yields $\underline{B}(\underset{\sim}{A}, \mathcal{G}_A) = 1.15 \times 10^{-5}$. The analogous bound from (5.10) is 3.25×10^{-6}, which is too small by a factor of 3.5.

The usual classical analysis here would be based on the statistic

$$T = \sum_{i=1}^{m} \sqrt{\frac{n_i}{n}} \cdot \frac{(X_i - \theta_0)}{\sigma_i},$$

which is $\mathcal{N}(0,1)$ under H_0. Computation yields $t = 4.77$, resulting in a two-sided P-value of 1.84×10^{-6}. As in Table 1, the classical P-value is thus only about 1/6 the value of $\underline{B}(\underset{\sim}{A}, \mathcal{G}_A)$.

5.3 Lower Bounds for $\mathcal{G}_{US}$

The analysis here is again based on Lemma 2.2. Indeed, for general $f(\theta)$ as in (5.1),

$$\underline{B}(\underset{\sim}{A}, \mathcal{G}_{US}) = f(\theta_0) / \sup_r \frac{1}{2r} \int_{\theta_0 - r}^{\theta_0 + r} f(\theta) d\theta. \tag{5.12}$$

This could be specialized to $f(\theta)$ in (5.3) if the σ_i were known, or to $f^*(\eta)$ in (5.7) if $\sigma_i^2 = \sigma^2/n_i$, with n_i known but σ^2 unknown; in the latter case, $\mathcal{G}_{US}$ would be defined as the class of all symmetric, unimodal densities in η (see (5.6)), rather than in θ. (One could, of course, question the suitability of assuming unimodality in η, and use instead, say, $\mathcal{G}_S$. Our recommendation, however, would be to use $\mathcal{G}_{US}$, since it is a large and reasonably "objective" class of priors.)

Simplification in (5.12) occurs for $f(\theta)$ in (5.3) or $f^*(\eta)$ in (5.7) when $m_1 = m$. Indeed, as in Berger and Sellke (1987), for the single study case it can then be shown that

$$\underline{B}(\underset{\sim}{A}, \mathcal{G}_{US}) = \begin{cases} 1 & \text{if } t \leq 1 \\ 2\varphi(t)/[\varphi(r^* + t) + \varphi(r^* - t)] & \text{if } t > 1, \end{cases}$$

where

$$t = \begin{cases} |\tilde{x} - \theta_0|/\tilde{\sigma} & \text{for } f(\theta), \\ \sum_{i=1}^{m} \sqrt{\frac{n_i}{n}} \frac{|x_i - \theta_0|}{\sigma_i} & \text{for } f^*(\eta), \end{cases}$$

and, for $t \geq 1.645, r^*$ can be approximated using the iterative formula

$$r_{i+1} = t + [2\log(r_i/\Phi(r_i - t)) - 1.838]^{1/2}, \tag{5.13}$$

beginning with $r_0 = t$.

Example 2 (continued). It was earlier calculated that $t = 4.77$, and iteration in (5.13) yields $r^* = 6.17$ (only 3 iterations being required for convergence). Thus $\underline{B}(\underset{\sim}{A}, \mathcal{G}_{US}) = 6.11 \times 10^{-5}$; compare with $\underline{B}(\underset{\sim}{A}, \mathcal{G}_A) = 1.15 \times 10^{-5}$.

6. CONCLUSIONS

The lower bounds on Bayes factors are not meant to be a substitute for actual subjective Bayes factors. They are, however, useful for getting some feel as to how "stars" or P-values convert into understandable measures of evidence for testing precise hypotheses. We recommend use of $\underline{B}(A, \mathcal{G}_{US})$ when possible, since it corresponds to a lower bound over "reasonable" priors.

The discrepancies between "stars" and $\underline{B}$ are generally less than the discrepancies between P-values and $\underline{B}$ given the actual data. Indeed, for the observational set corresponding to $\{p < \alpha\}$, we saw that $\underline{B} = \alpha$. Nevertheless, for many common reports, such as $.01 < p < .05$, $\underline{B}$ will often be 3 to 5 times larger than the "star" level.

We conclude with Table 5 that may help to summarize and locate existing formulas for $\underline{B}$. (The meta-analysis scenarios are not included, since there are so many different possibilities, only a few of which were covered in Section 5.) The table also indicates the complexity of the numerical calculation of $\underline{B}$. The table specifically refers only to the normal problem, though many of the results given are more general.

Table 5. Summary and location of results for determining $\underline{B}$ in the normal, single study, problem.

prior class	p-given: sign x known	p-given: sign x not known	$\alpha_1 < p < \alpha_2$: sign x known	$\alpha_1 < p < \alpha_2$: sign x not known
$\mathcal{G}_A$	See Berger and Sellke (1987). Closed form answers.	See Section 3.1. Approximate closed form answers.	See Section 3.3. Closed form answers.	See Section 3.2. Approximate closed form answers.
$\mathcal{G}_{US}$	See Berger and Sellke (1987). Iterative closed form answers.	Equivalent by Lemma 2.3.	See Section 4. Iterative closed form answers.	Equivalent by Lemma 2.3.
$\mathcal{G}_U$		Equivalent by Lemma 2.4.		Equivalent by Lemma 2.4.

ACKNOWLEDGEMENT

We are grateful to M.J. Bayarri for helpful comments concerning the paper.

REFERENCES

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

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

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

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

Casella, G. and Berger, R.L. (1987). Reconciling Bayesian and frequentist evidence in the one-sided testing problem. *J. Amer. Statist. Assoc.*, **82**, 106–111.

DeGroot, M.H. (1973). Doing what comes naturally: Interpreting a tail area as a posterior probability or as a likelihood ratio. *J. Amer. Statist. Assoc.*, **68**, 966–969.

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

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

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

Dickey, J.M. (1977). Is the tail area useful as an approximate Bayes factor? *J. Amer. Statist. Assoc.*, **72**, 138–142.

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

Good, I.J. (1958). Significance tests in parallel and in series. *J. Amer. Statist. Assoc.*, **53**, 799–813.

Good, I.J. (1967). A Bayesian significance test for the multinomial distribution. *J. Roy.*

Statist. Soc. Ser. B, **29**, 399–431.

Good, I.J. (1983). *Good Thinking: The Foundations of Probability and Its Applications.* Univ. Minnesota Press, Minneapolis.

Good, I.J. (1984). Notes C140, C144, C199, C200 and C201. *J. Statist. Comput. Simulation*, **19**.

Good, I.J. (1986). A flexible Bayesian model for comparing two treatments. *J. Statist. Comput. Simulation*, **26**, 301–305.

Hedges, L.V. and Olkin, I. (1985). *Statistical Methods for Meta-Analysis*. Academic Press, Orlando.

Hill, B. (1982). Comment on "Lindley's paradox," by G. Shafer. *J. Amer. Statist. Assoc.*, **77**, 344–347.

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

Kass, R.E. and Vaidyanathan, S.K. (1991). Approximate Bayes factors and orthogonal parameters, with application to testing equality of binomial proportions. To appear in *J. Roy. Statist. Soc. B*, 1991.

Lempers, F.B. (1971). *Posterior Probabilities of Alternative Linear Models.* Univ. Rotterdam Press, Rotterdam.

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

Lindley, D.V. (1977). A problem in forensic science. *Biometrika*, **64**, 207–213.

Pratt, J.W. (1965). Bayesian interpretation of standard inference statements (with discussion). *J. Roy. Statist. Soc. Ser. B*, **27**, 169–203.

Rubin, H. (1971). A decision-theoretic approach to the problem of testing a null hypothesis.

In *Statistical Decision Theory and Related Topics* (S.S. Gupta and J. Yackel, eds.). Academic, New York.

Shafer, G. (1982). Lindley's paradox (with discussion). *J. Amer. Statist. Assoc.*, **77**, 325–351.

Smith, A.F.M. and Spiegelhalter, D.J. (1980). Bayes factors and choice criteria for linear models. *J. Roy. Statist. Soc. Ser. B*, **42**, 213–220.

Zellner, A. (1971). *An Introduction to Bayesian Inference in Economics.* Wiley, New York.

Zellner, A. (1984). Posterior odds ratios for regression hypotheses: General considerations and some specific results. In *Basic Issues in Econometrics* (A. Zellner, ed.), 275–305. Univ. Chicago Press, Chicago.

Appendix I

Proof of Lemma 2.1: This is immediate from the fact that $m_g(A)$ in (1.3) is a linear functional of g, and hence is maximized over "point mass" g.

Proof of Lemma 2.2: Any $g \in \mathcal{G}_{US}$ can be represented as

$$g(\theta) = \int_0^\infty \frac{1}{2r} 1_{(\theta_0 - r, \theta_0 + r)}(\theta) dF(r) \tag{A.1}$$

for some distribution function F, where $1_C(\theta)$ denotes the indicator function on the set C. Thus, for $g \in \mathcal{G}_{US}$, an application of Fubini's theorem yields

$$\begin{aligned} B_g(A) &= P(A|\theta_0) / \int_{\theta \neq \theta_0} P(A|\theta) g(\theta) d\theta \\ &= \frac{P(A|\theta_0)}{\int_0^\infty \frac{1}{2r} \left[\int_{\theta \neq \theta_0} P(A|\theta) 1_{(\theta_0 - r, \theta_0 + r)}(\theta) d\theta \right] dF(r)} \end{aligned}$$

$$= \frac{P(A|\theta_0)}{\int_0^{\infty} \frac{1}{2r} H_A(r) dF(r)}. \tag{A.2}$$

Since any g as in (A.1) can also be shown to be in $\mathcal{G}_{US}$, it follows that

$$\begin{aligned}
\underline{B}(A, \mathcal{G}_{US}) &= \inf_{\{\text{all } F\}} B_g(A) \\
&= \frac{P(A|\theta_0)}{\sup\limits_{\{\text{all } F\}} \int_0^{\infty} \frac{1}{2r} H_A(r) dF(r)} \\
&= \frac{P(A|\theta_0)}{\sup\limits_r [\frac{1}{2r} H_A(r)]},
\end{aligned}$$

by the same reasoning as in the proof of Lemma 2.1. □

Appendix II

Lemma A.1: *$\tilde{\theta}^*$ in (3.2) and $\underline{\tilde{B}}$ in (3.3) are asymptotically correct in the sense that, as $q \to \infty$,*

$$\theta^* = q - 2q \exp\{-2q^2\}(1 + o(1)), \tag{A.3}$$

$$\underline{B}(A_0, \mathcal{G}_a) = \frac{2 \exp\{-\frac{1}{2}q^2\}}{1 + \exp\{-2q^2\} + 2q^2 \exp\{-4q^2\}(1 + o(1))}. \tag{A.4}$$

Proof. Defining $h(\theta) = \log[\varphi(q - \theta) + \varphi(q + \theta)]$, computation yields

$$h'(\theta) = \frac{d}{d\theta} h(\theta) = q - \theta - 2q/(1 + e^{2q\theta}). \tag{A.5}$$

Setting equal to zero and replacing θ by

$$\theta = q - 2q \exp\{-2q^2\}(1 + \varepsilon)$$

shows that any maximizing θ must satisfy

$$(1 + \varepsilon)[\exp\{-2q^2\} + \exp\{-4q^2(1 + \varepsilon)e^{-2q^2}\}] = 1. \tag{A.6}$$

It is easy to check that for, say, $q > 100$, the left hand side of (A.6) is less than 1 when $\varepsilon = 0$, and is greater than 1 when $\varepsilon = q^{-1}$. Obviously the left hand side of (A.6) is continuous in ε, so the equation must have a solution with $0 < \varepsilon < q^{-1}$. This proves that (A.3) defines a solution to $h'(\theta) = 0$.

Next, observe that

$$h''(\theta) = \frac{d^2}{d\theta^2} h(\theta) = -1 + 4q^2 (e^{q\theta} + e^{-q\theta})^{-2}.$$

Since $\theta^* > 1$ for large q,

$$h''(\theta^*) < -1 + 4q^2 e^{-2q} < 0$$

for large q. Hence θ^* is a local maximum.

Finally, $h''(\theta) = 0$ clearly has at most 2 solutions, so that $h'(\theta) = 0$ can have at most three solutions. Besides θ^* in (A.3), it can be shown that$(-\theta^*)$ is a solution to $h'(\theta) = 0$, with $h(-\theta^*) = h(\theta^*)$ by symmetry. The third solution to $h'(\theta) = 0$ is $\theta = 0$, and since it lies between two local maxima it must be a minimum. Hence θ^* is a global maximum.

Equation (A.4) follows from a lengthy Taylors series argument, inserting (A.3) into (3.1). $\square$

Lemma A.2: *$\tilde{\theta}^*$ in (3.7) is an asymptotically correct approximation to θ^*, in the sense that, as $q_1 \to \infty$,*

$$\theta^* = \frac{1}{2}(q_1 + q_2) - (q_2 - q_1)^{-1} e^{-q_1 q_2} (e^{-q_1^2} - e^{-q_2^2})(1 + o(1)). \qquad \text{(A.7)}$$

Proof. Define $\delta = (q_2 - q_1)$ and

$$Q(q_1, \delta) = (\delta)^{-1} e^{-(2q_1^2 + q_1\delta)} (1 - e^{-[2q_1\delta + \delta^2]}),$$

so that (A.7) becomes

$$\theta^* = \frac{1}{2}(q_1 + q_2) - Q(q_1, \delta)(1 + o(1)). \tag{A.8}$$

From (3.6), calculation gives

$$\begin{aligned}\frac{d}{d\theta}P(A_3|\theta) &= \varphi(\theta - q_1) - \varphi(\theta - q_2) + \varphi(\theta + q_2) - \varphi(\theta + q_1)\\ &= \varphi(\theta - q_1)[1 - e^{-2q_1\theta}] - \varphi(\theta - q_2)[1 - e^{-2q_2\theta}]. \end{aligned} \tag{A.9}$$

Next, define $\epsilon(\theta)$ by

$$\theta = q_1 + \frac{1}{2}\delta - Q(q_1, \delta)(1 + \epsilon(\theta)); \tag{A.10}$$

from (A.8) it is clear that our goal is to show that $\epsilon(\theta^*) = o(1)$.

Using (A.10) in (A.9) yields

$$\frac{d}{d\theta}P(A_3|\theta) = \frac{1}{\sqrt{2\pi}}\exp\{-\frac{1}{2}[\frac{1}{4}\delta^2 + Q^2(1+\varepsilon)^2]\}\psi(\epsilon), \tag{A.11}$$

where

$$\begin{aligned}\psi(\epsilon) = &[1 - \exp\{-2q_1(q_1 + \frac{1}{2}\delta - Q(1+\epsilon))\}]\exp\{\frac{1}{2}\delta Q(1+\epsilon)\}\\ &- [1 - \exp\{-2(q_1+\delta)(q_1 + \frac{1}{2}\delta - Q(1+\epsilon))\}]\exp\{-\frac{1}{2}\delta Q(1+\epsilon)\}.\end{aligned}$$

Note first that

$$\begin{aligned}\frac{\partial}{\partial\delta}\log Q(q_1, \delta) &= -\frac{1}{\delta} - q_1 + \frac{2(q_1+\delta)}{[\exp\{2q_1\delta + \delta^2\} - 1]}\\ &< -\frac{1}{\delta} - q_1 + \frac{2(q_1+\delta)}{[2q_1\delta + \delta^2]} < 0\end{aligned}$$

for large q_1. Hence

$$Q(q_1, \delta) < \lim_{\delta \to 0} Q(q_1, \delta) = 2q_1 e^{-2q_1^2}.$$

If we consider $|\varepsilon(\theta)| < 1$, it follows that $q_1 Q(1+\epsilon) = o(1)$ uniformly in ϵ, so that

$$\begin{aligned}
\psi(\epsilon) &= [1 - \exp\{-2Q_1(q_1 + \frac{1}{2}\delta)\}(1+o(1))][1 + \frac{1}{2}\delta Q(1+\epsilon)(1+o(1))] \\
&\quad - [1 - \exp\{-2(q_1+\delta)(q_1 + \frac{1}{2}\delta)\}(1+o(1))][1 - \frac{1}{2}\delta Q(1+\epsilon)(1+o(1))] \\
&= \delta Q(1+\epsilon)(1+o(1)) - \exp\{-2q_1(q_1+\delta)\}(1+o(1)) \\
&= \exp\{-2q_1(q_1+\delta)\}(\varepsilon + o(1)).
\end{aligned}$$

It can be checked that the $o(1)$ term is uniform in $|\epsilon| < 1$ and is continuous in ϵ. The conclusion is that there exists an ϵ which is $o(1)$ for which $\psi(\epsilon)$ and hence (A.11) is zero, so that (A.10) defines a local maxima or minima. The other zeroes of $\frac{d}{d\theta}P(A_3|\theta)$ can be shown to be $(-\theta^*)$ and 0. For very large θ the derivative is negative, so θ^* must be a global maxima. $\square$

Appendix III

Proof of Equation (4.2): Using Lemma 2.3 we obtain

$$\begin{aligned}
H_{A_3}(r) &= \int_{-r}^{r} P(A_3|\theta)d\theta \\
&= 2\int_{-r}^{r} P(A_1|\theta)d\theta \\
&= 2\int_{-r}^{r} \Phi(\theta - q_1)d\theta - 2\int_{-r}^{r} \Phi(\theta - q_2)d\theta.
\end{aligned}$$

Since

$$\int_{-\infty}^{x} \Phi(y)dy = x\Phi(x) + \varphi(x),$$

it follows that

$$\begin{aligned}
H_{A_3}(r) = 2\Big\{ &[(r-q_1)\Phi(r-q_1) + \varphi(r-q_1)] - [(r-q_2)\Phi(r-q_2) + \varphi(r-q_2)] \\
&+ [(r+q_1)(1-\Phi(r+q_1)) - \varphi(r+q_1)] - [(r+q_2)(1-\Phi(r+q_2)) - \varphi(r+q_2)]\Big\}.
\end{aligned}$$

Algebra yields (4.2).

Explanation of Equation (4.3): The maximizing r is a solution to

$$0 = \frac{d}{dr} \log[\frac{1}{2r} H_{A_3}(r)] = -\frac{1}{r} + \frac{H'_{A_3}(r)}{H_{A_3}(r)},$$

where (using (4.2))

$$\begin{aligned} H'_{A_3}(r) &= \frac{d}{dr} H_{A_3}(r) = P(A_3|r) + P(A_3|-r) \\ &= 2\left[\Phi(r - q_1) - \Phi(r + q_1) + \Phi(r + q_2) - \Phi(r - q_2)\right]. \end{aligned}$$

Rearranging terms gives (4.3).