Bayesian Analysis (2007) **2**, Number 1, pp. 45–58

Comparing Normal Means: New Methods for an Old Problem

José M. Bernardo* and Sergio Pérez†

Abstract. Comparing the means of two normal populations is an old problem in mathematical statistics, but there is still no consensus about its most appropriate solution. In this paper we treat the problem of comparing two normal means as a Bayesian decision problem with only two alternatives: either to accept the hypothesis that the two means are equal, or to conclude that the observed data are, under the assumed model, incompatible with that hypothesis. The combined use of an information-theory based loss function, the *intrinsic discrepancy* (Bernardo and Rueda 2002), and an objective prior function, the *reference prior* (Bernardo 1979; Berger and Bernardo 1992), produces a new solution to this old problem which has the invariance properties one should presumably require.

Keywords: Bayes factor, BRC, comparison of normal means, intrinsic discrepancy, precise hypothesis testing, reference prior, two sided tests.

1 Introduction

1.1 Problem statement

Comparing the expected values μ_x and μ_y of two normal populations $\mathrm{N}(x \,|\, \mu_x, \sigma_x)$ and $\mathrm{N}(y \,|\, \mu_y, \sigma_y)$, given the information provided by two independent random samples, $\boldsymbol{x} = \{x_1, \ldots, x_n\}$ and $\boldsymbol{y} = \{y_1, \ldots, y_m\}$ of possibly different sizes, is surely one of the oldest non-trivial problems in mathematical statistics.

In this paper we formally treat the problem of comparing two normal means as a decision problem with only two alternatives: either a_0: to accept the (null) hypothesis that the two means are equal, and hence work as if $\mu_1 = \mu_2$; or a_1: to conclude that, under the assumed model, the observed data are incompatible with such hypothesis. Within this framework, the solution obviously depends on both the loss function and the prior distribution. In Section 2, a number of options are analyzed, and it is argued that the combined use of an invariant information-theory based loss function, the *intrinsic discrepancy* (Bernardo and Rueda 2002), and an objective prior function, the *reference prior* (Bernardo 1979; Berger and Bernardo 1992) may be expected to produce an appropriate solution, which is invariant under reparametrization. In Section 3, the problem of comparing two normal means with common variance is solved from this point of view, and its extension to the case of possibly different variances is briefly discussed.

*Universitat de València, Valencia, Spain, http://www.uv.es/bernardo
†Colegio de Postgraduados, Montecillo, Mexico, mailto:sergiop@colpos.mx

 ba0005

1.2 Notation

Probability distributions are described through their probability density functions, and no notational distinction is made between a random quantity and the particular values that it may take. Bold italic roman fonts are used for observable random vectors (typically data) and bold italic greek fonts for unobservable random vectors (typically parameters); lower case is used for variables and upper case calligraphic for their domain sets. The standard mathematical convention of referring to functions, say $f_{\boldsymbol{z}}(\cdot)$ and $g_{\boldsymbol{z}}(\cdot)$, respectively by $f(\boldsymbol{z})$ and $g(\boldsymbol{z})$ is often used. Thus, the conditional probability density of observable data given $\boldsymbol{\omega}$ is represented by either $p_{\boldsymbol{z}}(\cdot \,|\, \boldsymbol{\omega})$ or $p(\boldsymbol{z} \,|\, \boldsymbol{\omega})$, with $p(\boldsymbol{z} \,|\, \boldsymbol{\omega}) \geq 0$, and $\int_{\boldsymbol{\mathcal{Z}}} p(\boldsymbol{z} \,|\, \boldsymbol{\omega})\, d\boldsymbol{z} = 1$, and the posterior density of a non-observable parameter vector $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ given data $\boldsymbol{z}$ is represented by either $\pi_{\boldsymbol{\theta}}(\cdot \,|\, \boldsymbol{z})$ or $\pi(\boldsymbol{\theta} \,|\, \boldsymbol{z})$, with $\pi(\boldsymbol{\theta} \,|\, \boldsymbol{z}) \geq 0$ and $\int_{\boldsymbol{\Theta}} \pi(\boldsymbol{\theta} \,|\, \boldsymbol{z})\, d\boldsymbol{\theta} = 1$. Density functions of specific distributions are denoted by appropriate names. In particular, if x has a normal distribution with mean μ and standard deviation σ, its probability density function will be denoted by $\mathrm{N}(x \,|\, \mu,\, \sigma)$; if λ has a gamma distribution with mean α/β and variance α/β^2, its density function will be denoted by $\mathrm{Ga}(\lambda \,|\, \alpha, \beta)$; if t has a noncentral Student t distribution with non centrality parameter δ and ν degrees of freedom, its density function will be denoted by $\mathrm{NcSt}(t \,|\, \delta, \nu)$.

In the problem considered, available data $\boldsymbol{z}$ typically consist of two random samples $\boldsymbol{z} = \{\boldsymbol{x}, \boldsymbol{y}\}$, $\boldsymbol{x} = \{x_1, \ldots, x_n\}$ and $\boldsymbol{y} = \{y_1, \ldots, y_m\}$, of possibly different sizes n and m, respectively drawn from $\mathrm{N}(x \,|\, \mu_x, \sigma_x)$ and $\mathrm{N}(y \,|\, \mu_y, \sigma_y)$. Standard notation is used for the sample means and variances, respectively denoted by $\bar{x} = \sum_{j=1}^{n} x_j/n$, $\bar{y} = \sum_{j=1}^{m} y_j/m$ and $s_x^2 = \sum_{j=1}^{n} (x_j - \bar{x})^2/n$, $s_y^2 = \sum_{j=1}^{m} (y_j - \bar{y})^2/m$.

2 Structure of the decision problem

2.1 Precise hypothesis testing

Assume that available data $\boldsymbol{z}$ have been generated from an unknown element of the family of probability distributions for $\boldsymbol{z} \in \mathcal{Z}$, $\mathcal{M} = \{p_{\boldsymbol{z}}(\cdot \,|\, \boldsymbol{\phi}, \boldsymbol{\omega}),\ \boldsymbol{\phi} \in \Phi,\ \boldsymbol{\omega} \in \Omega\}$, and suppose that it is desired to evaluate whether or not these data may be judged to be compatible with the (null) hypothesis $H_0 \equiv \{\boldsymbol{\phi} = \boldsymbol{\phi}_0\}$. This may be treated as a decision problem with only two alternatives:

$$\begin{cases} a_0: & \text{to accept } H_0 \text{ and work } \textit{as if } \boldsymbol{\phi} = \boldsymbol{\phi}_0 \\ a_1: & \text{to claim that the observed data are incompatible with } H_0 \end{cases} \tag{1}$$

Notice that, with this formulation, H_0 is generally a composite hypothesis, described by the family of probability distributions $\mathcal{M}_0 = \{p_{\boldsymbol{z}}(\cdot \,|\, \boldsymbol{\phi}_0, \boldsymbol{\omega}_0),\ \boldsymbol{\omega}_0 \in \Omega\}$, for $\boldsymbol{z} \in \mathcal{Z}$. Simple nulls are included as a particular case where there are no nuisance parameters.

The foundations of decision theory (see *e.g.*, Bernardo and Smith 1994, Ch. 2, and references therein) dictate that to solve this decision problem utility functions $u\{a_i, (\boldsymbol{\phi}, \boldsymbol{\omega})\}$ for the two alternatives a_0 and a_1, and a joint prior distribution $\pi(\boldsymbol{\phi}, \boldsymbol{\omega})$

for the unknown parameters $(\boldsymbol{\phi}, \boldsymbol{\omega})$ must be specified, and that H_0 should be rejected if, and only if, the posterior expected utility from rejecting, $\overline{u}(a_1 \,|\, \boldsymbol{z})$, is larger than the posterior utility from accepting, $\overline{u}(a_0 \,|\, \boldsymbol{z})$, that is if, and only if,

$$\overline{u}(a_1 \,|\, \boldsymbol{z}) - \overline{u}(a_0 \,|\, \boldsymbol{z}) = \int_{\Phi} \int_{\Omega} [u\{a_1, (\boldsymbol{\phi}, \boldsymbol{\omega})\} - u\{a_0, (\boldsymbol{\phi}, \boldsymbol{\omega})\}] \, \pi(\boldsymbol{\phi}, \boldsymbol{\omega} \,|\, \boldsymbol{z}) \, d\boldsymbol{\phi} \, d\boldsymbol{\omega} > 0,$$

where, using Bayes theorem, $\pi(\boldsymbol{\phi}, \boldsymbol{\omega} \,|\, \boldsymbol{z}) \propto p(\boldsymbol{z} \,|\, \boldsymbol{\phi}, \boldsymbol{\omega}) \, \pi(\boldsymbol{\phi}, \boldsymbol{\omega})$ is the joint posterior which corresponds to the prior $\pi(\boldsymbol{\phi}, \boldsymbol{\omega})$. Thus, only the difference $u\{a_1, (\boldsymbol{\phi}, \boldsymbol{\omega})\} - u\{a_0, (\boldsymbol{\phi}, \boldsymbol{\omega})\}$, must be specified. This difference may usefully be written as

$$u\{a_1, (\boldsymbol{\phi}, \boldsymbol{\omega})\} - u\{a_0, (\boldsymbol{\phi}, \boldsymbol{\omega})\} = \ell\{\boldsymbol{\phi}_0, (\boldsymbol{\phi}, \boldsymbol{\omega})\} - u_0,$$

where $\ell\{\boldsymbol{\phi}_0, (\boldsymbol{\phi}, \boldsymbol{\omega})\}$ may be interpreted (Bernardo and Rueda 2002) as the non-negative terminal loss suffered by accepting $\boldsymbol{\phi} = \boldsymbol{\phi}_0$ given $(\boldsymbol{\phi}, \boldsymbol{\omega})$, and where $u_0 > 0$ is the utility of accepting H_0 when it is true. The corresponding Bayes criterion is to reject H_0 if, and only if,

$$t(\boldsymbol{\phi}_0 \,|\, \boldsymbol{z}) = \int_{\Theta} \int_{\Omega} \ell\{\boldsymbol{\phi}_0, (\boldsymbol{\phi}, \boldsymbol{\omega})\} \, \pi(\boldsymbol{\phi}, \boldsymbol{\omega} \,|\, \boldsymbol{z}) \, d\boldsymbol{\phi} \, d\boldsymbol{\omega} > u_0,$$

that is, if the posterior expected loss, the *test statistic* $t(\boldsymbol{\phi}_0 \,|\, \boldsymbol{z})$ is large enough.

2.2 The intrinsic discrepancy loss

As one would expect, the optimal decision depends heavily on the particular loss function $\ell\{\boldsymbol{\phi}_0, (\boldsymbol{\phi}, \boldsymbol{\omega})\}$ which is assumed to describe the preferences of the decision maker. Specific problems may require specific loss functions, but conventional loss functions may be used to proceed when one does not have any particular application in mind.

2.2.1 Conventional loss functions

A common conventional loss function is the *step* loss function induced by assuming step utility functions for both actions of the form $u\{a_1, (\boldsymbol{\phi}, \boldsymbol{\omega})\} = a$ if $\boldsymbol{\phi} \neq \boldsymbol{\phi}_0$, and zero otherwise, and $u\{a_0, (\boldsymbol{\phi}, \boldsymbol{\omega})\} = b$ if $\boldsymbol{\phi} = \boldsymbol{\phi}_0$, and zero otherwise. Whatever the loss function might be, one may (or may not) hold a sharp prior which makes values of $\boldsymbol{\phi}$ close to the null value $\boldsymbol{\phi}_0$ comparatively very likely. Notice, however, that a step loss function *forces* the use of a *non-regular* "spiked" proper prior which places a lump of probability $p_0 > 0$ at $\boldsymbol{\phi} = \boldsymbol{\phi}_0$, for otherwise the optimal decision would always be to reject H_0. This leads to rejecting H_0 if (and only if) its posterior probability is too small or, equivalently, if (and only if) the *Bayes factor* against H_0, is sufficiently large. This will be appropriate wherever preferences are well described by a step loss function, and prior information is available to justify an informative, spiked prior. It may be argued that many scientific applications of precise hypothesis testing fail to meet one or both of these conditions.

Another example of a conventional loss function is the ubiquitous *quadratic* loss function. With a quadratic loss function, H_0 should be rejected if (and only if) the

posterior expected Euclidean distance of $\boldsymbol{\phi}_0$ from the true value $\boldsymbol{\phi}$ is too large. In marked contrast with step loss functions, the quadratic loss function (as many other continuous loss functions) may safely be used with (typically improper) 'noninformative' priors. However, most conventional continuous loss functions, such as the quadratic loss, depend dramatically on the particular parametrization used. Yet, since the model parametrization is arbitrary, one would expect that for any one-to-one function $\boldsymbol{\psi}(\boldsymbol{\phi})$, the conditions under which $\boldsymbol{\phi} = \boldsymbol{\phi}_0$ must be rejected should be *precisely the same* as the conditions under which $\boldsymbol{\psi} = \boldsymbol{\psi}(\boldsymbol{\phi}_0)$ must be rejected. This requires the use of a loss function which is invariant under one-to-one reparametrizations.

2.2.2 The intrinsic discrepancy loss function

Bernardo and Rueda (2002) and Bernardo (2005b) argue that an invariant loss function which is appropriate for general use in hypothesis testing is the *intrinsic discrepancy*, $\delta_{\boldsymbol{z}}\{H_0, (\boldsymbol{\phi}, \boldsymbol{\omega})\}$, defined as the minimum (Kullback-Leibler) logarithmic divergence between the distribution $p_{\boldsymbol{z}}(\cdot \,|\, \boldsymbol{\phi}, \boldsymbol{\omega})$ which is assumed to have generated the data, and the family of distributions $\mathcal{F}_0 \equiv \{p_{\boldsymbol{z}}(\cdot \,|\, \boldsymbol{\phi}_0, \boldsymbol{\omega}_0), \boldsymbol{\omega}_0 \in \Omega\}$ which corresponds to the hypothesis $H_0 \equiv \{\boldsymbol{\phi} = \boldsymbol{\phi}_0\}$ to be tested. Formally,

$$\delta_{\boldsymbol{z}}\{H_0, (\boldsymbol{\phi}, \boldsymbol{\omega})\} \equiv \inf_{\boldsymbol{\omega}_0 \in \boldsymbol{\Omega}} \kappa^*_{\boldsymbol{z}}\{p_{\boldsymbol{z}}(\cdot \,|\, \boldsymbol{\phi}_0, \boldsymbol{\omega}_0),\, p_{\boldsymbol{z}}(\cdot \,|\, \boldsymbol{\phi}, \boldsymbol{\omega})\}, \tag{2}$$

$$\kappa^*_{\boldsymbol{z}}\{p_{\boldsymbol{z}}(\cdot), q_{\boldsymbol{z}}(\cdot)\} \equiv \min\{\kappa\{p_{\boldsymbol{z}}(\cdot) \,|\, q_{\boldsymbol{z}}(\cdot)\},\, \kappa\{q_{\boldsymbol{z}}(\cdot) \,|\, p_{\boldsymbol{z}}(\cdot)\}\}, \tag{3}$$

$$\kappa\{q_{\boldsymbol{z}}(\cdot) \,|\, p_{\boldsymbol{z}}(\cdot)\} \equiv \int_{\mathcal{Z}} p_{\boldsymbol{z}}(\boldsymbol{z}) \log \frac{p_{\boldsymbol{z}}(\boldsymbol{z})}{q_{\boldsymbol{z}}(\boldsymbol{z})}\, d\boldsymbol{z}. \tag{4}$$

The *intrinsic discrepancy function* $\kappa^*_{\boldsymbol{z}}\{p_{\boldsymbol{z}}(\cdot), q_{\boldsymbol{z}}(\cdot)\}$ defined by (3), a measure of the disparity between the distributions $p_{\boldsymbol{z}}(\cdot)$ and $q_{\boldsymbol{z}}(\cdot)$, has many attractive properties. It is *symmetric*, *non-negative*, and it is zero if, and only if, $p_{\boldsymbol{z}}(\boldsymbol{z}) = q_{\boldsymbol{z}}(\boldsymbol{z})$ almost everywhere. It inherits the *additive* property of the logarithmic divergence, so that it is *additive* under independent observations; thus if $\boldsymbol{z} = \{x_1, \ldots, x_n\}$, $p_{\boldsymbol{z}}(\boldsymbol{z}) = \prod_{i=1}^n p_x(x_i)$, and $q_{\boldsymbol{z}}(\boldsymbol{z}) = \prod_{i=1}^n q_x(x_i)$, then

$$\kappa^*_{\boldsymbol{z}}\{p_{\boldsymbol{z}}(\cdot), q_{\boldsymbol{z}}(\cdot)\} = n\, \kappa^*_x\{p_x(\cdot), q_x(\cdot)\}.$$

Thus, the intrinsic loss to be suffered by deciding that a random sample of size n was generated from a particular distribution is n times the intrinsic loss to be suffered by deciding that a single observation was generated from that distribution.

The intrinsic discrepancy loss (2) is invariant under one-to one transformations of the parameters. Thus, for any one-to-one function $\boldsymbol{\psi} = \boldsymbol{\psi}(\boldsymbol{\phi})$ the intrinsic loss suffered from assuming that $\boldsymbol{\phi} = \boldsymbol{\phi}_0$ is precisely the same as that of assuming that $\boldsymbol{\psi} = \boldsymbol{\psi}(\boldsymbol{\phi}_0)$. Moreover, one may equivalently work with sufficient statistics: if $\boldsymbol{t} = \boldsymbol{t}(\boldsymbol{z})$ is a sufficient statistic for $p_{\boldsymbol{z}}(\cdot \,|\, \boldsymbol{\phi}, \boldsymbol{\omega})$, then $\delta_{\boldsymbol{z}}\{H_0, (\boldsymbol{\phi}, \boldsymbol{\omega})\} = \delta_{\boldsymbol{t}}\{H_0, (\boldsymbol{\phi}, \boldsymbol{\omega})\}$. The intrinsic loss may be safely be used with improper priors.

If, as it is usually the case, the parameter space $\boldsymbol{\Phi} \times \boldsymbol{\Omega}$ is convex, the two minimization

procedures in (2) and (3) may be interchanged (Juárez 2005) to have

$$\begin{aligned} &\delta_{\boldsymbol{z}}\{H_0, (\boldsymbol{\phi}, \boldsymbol{\omega})\} \\ &= \min\left\{ \inf_{\boldsymbol{\omega}_0 \in \boldsymbol{\Omega}} \kappa\{p_{\boldsymbol{z}}(\cdot \,|\, \boldsymbol{\phi}_0, \boldsymbol{\omega}_0) \,|\, p_{\boldsymbol{z}}(\cdot \,|\, \boldsymbol{\phi}, \boldsymbol{\omega})\}, \inf_{\boldsymbol{\omega}_0 \in \boldsymbol{\Omega}} \kappa\{p_{\boldsymbol{z}}(\cdot \,|\, \boldsymbol{\phi}, \boldsymbol{\omega}) \,|\, p_{\boldsymbol{z}}(\cdot \,|\, \boldsymbol{\phi}_0, \boldsymbol{\omega}_0)\} \right\} \end{aligned} \tag{5}$$

which is typically easier to compute than direct evaluation of (2) and (3).

As it is apparent from its definition, the intrinsic loss $\delta_{\boldsymbol{z}}\{H_0, (\boldsymbol{\phi}, \boldsymbol{\omega})\}$ is the *minimum expected log-likelihood ratio* (under repeated sampling) against H_0. This provides a direct *calibration* for its numerical values; thus, intrinsic loss values of about $\log(10)$ indicate some evidence against H_0, while intrinsic loss values of about $\log(100)$ indicate rather strong evidence against H_0.

2.3 The Bayesian Reference Criterion (BRC)

Any statistical procedure depends on the accepted assumptions, and those typically include many subjective judgements. If has become standard, however, to term 'objective' any statistical procedure whose results only depend on the quantity of interest, the model assumed and the data obtained. From this point of view, frequentist procedures are declared to be 'objective', and this has often be used as an argument against Bayesian solutions. Objective Bayesian solutions in this sense require the use of objective prior functions, that is formal priors which only depend on the quantity of interest and on the model assumed. See Berger (2006) (and ensuing discussion) for a recent analysis of this often polemical issue. The *reference prior* (Bernardo 1979; Berger and Bernardo 1992; Bernardo and Smith 1994; Bernardo 2005a), loosely defined as that prior which maximizes the missing information about the quantity of interest, provides a general solution to the problem of specifying an objective prior.

2.3.1 The intrinsic statistic

The Bayesian reference criterion (BRC) is the normative Bayes solution to the decision problem of hypothesis testing described in Section 2.1 which corresponds to the use of the *intrinsic loss* function and the *reference prior* function.

Given model $\mathcal{M} = \{p_{\boldsymbol{z}}(\cdot \,|\, \boldsymbol{\phi}, \boldsymbol{\omega}),\ \boldsymbol{\phi} \in \Phi,\ \boldsymbol{\omega} \in \Omega\}$, this formally means to reject the hypothesis $H_0 \equiv \{\boldsymbol{\phi} = \boldsymbol{\phi}_0\}$ if, and only if

$$d(H_0 \,|\, \boldsymbol{z}) = \int_0^\infty \delta\, \pi(\delta \,|\, \boldsymbol{z})\, d\delta > \delta_0, \tag{6}$$

where $d(H_0 \,|\, \boldsymbol{z})$, termed the *intrinsic (test) statistic*, is the reference posterior expectation of the intrinsic loss $\delta_{\boldsymbol{z}}\{H_0, (\boldsymbol{\phi}, \boldsymbol{\omega})\}$ defined by (2), and where δ_0 is a context dependent positive utility constant, *the largest acceptable average log-likelihood ratio against H_0 under repeated sampling*. For scientific communication, δ_0 could conventionally be set to $\delta_0 = \log(100) \approx 4.6$.

3 Normal means comparison

3.1 Problem statement in the common variance case

Let available data $\boldsymbol{z} = \{\boldsymbol{x}, \boldsymbol{y}\}$, $\boldsymbol{x} = \{x_1, \ldots, x_n\}$, $\boldsymbol{y} = \{y_1, \ldots, y_m\}$, consist of two random samples of possibly different sizes n and m, respectively drawn from $\mathrm{N}(x \,|\, \mu_x, \sigma)$ and $\mathrm{N}(y \,|\, \mu_y, \sigma)$, so that the assumed model is

$$p(\boldsymbol{z} \,|\, \mu_x, \mu_y, \sigma) = \prod_{i=1}^{n} \mathrm{N}(x_i \,|\, \mu_x, \sigma) \prod_{j=1}^{m} \mathrm{N}(y_j \,|\, \mu_y, \sigma). \tag{7}$$

It is desired to test $H_0 \equiv \{\mu_x = \mu_y\}$, that is, whether or not these data could have been drawn from some member of the family of probability distributions

$$\begin{aligned} \mathcal{M}_0 &\equiv \{p(\boldsymbol{z} \,|\, \mu_0, \mu_0, \sigma_0), \quad \mu_0 \in \Re,\ \sigma_0 > 0\} \\ p(\boldsymbol{z} \,|\, \mu_0, \mu_0, \sigma_0) &= \prod_{i=1}^{n} \mathrm{N}(x_i \,|\, \mu_0, \sigma_0) \prod_{j=1}^{m} \mathrm{N}(y_j \,|\, \mu_0, \sigma_0). \end{aligned} \tag{8}$$

To implement the BRC criterion described above one should:

1. Compute the *intrinsic discrepancy* $\delta\{H_0, (\mu_x, \mu_y, \sigma)\}$ between the family of distributions $\mathcal{M}_0$ which define the hypothesis H_0 and the assumed model $p(\boldsymbol{z} \,|\, \mu_x, \mu_y, \sigma)$.
2. Determine the *reference* joint prior $\pi_\delta(\mu_x, \mu_y, \sigma)$ of the three unknown parameters when δ is the quantity of interest.
3. Derive the relevant *intrinsic statistic*, that is the reference posterior expectation $d(H_0 \,|\, \boldsymbol{z}) = \int_0^\infty \delta \, \pi_\delta(\delta \,|\, \boldsymbol{z}) \, d\delta$ of the intrinsic discrepancy $\delta\{H_0, (\mu_x, \mu_y, \sigma)\}$.

3.2 The intrinsic loss

Logarithmic divergences. The (Kullback-Leibler) logarithmic divergence of a normal distribution $\mathrm{N}(x \,|\, \mu_2, \sigma_2)$ from another normal distribution $\mathrm{N}(x \,|\, \mu_1, \sigma_1)$ is given by

$$\begin{aligned} \kappa\{\mu_2, \sigma_2 \,|\, \mu_1, \sigma_1\} &\equiv \kappa\{\mathrm{N}(x \,|\, \mu_2, \sigma_2) \,|\, \mathrm{N}(x \,|\, \mu_1, \sigma_1)\} \\ &\equiv \int_{-\infty}^{\infty} \mathrm{N}(x \,|\, \mu_1, \sigma_1) \log \frac{\mathrm{N}(x \,|\, \mu_1, \sigma_1)}{\mathrm{N}(x \,|\, \mu_2, \sigma_2)} \, dx \\ &= \frac{1}{2} \left(\frac{\mu_2 - \mu_1}{\sigma_2} \right)^2 + \frac{1}{2} \left(\frac{\sigma_1^2}{\sigma_2^2} - 1 - \log \frac{\sigma_1^2}{\sigma_2^2} \right). \end{aligned} \tag{9}$$

This is a nonnegative quantity which is zero if, and only if, $\mu_1 = \mu_2$ and $\sigma_1 = \sigma_2$. Using (9) and the additive property of the logarithmic divergences, the logarithmic divergence of $p(\boldsymbol{z} \,|\, \mu_0, \mu_0, \sigma_0)$ from $p(\boldsymbol{z} \,|\, \mu_x, \mu_y, \sigma)$ is

$$\begin{aligned} &\kappa\{p_{\boldsymbol{z}}(\cdot \,|\, \mu_0, \mu_0, \sigma_0) \,|\, p_{\boldsymbol{z}}(\cdot \,|\, \mu_x, \mu_y, \sigma)\} \\ &= n \, \kappa\{\mu_0, \sigma_0 \,|\, \mu_x, \sigma\} + m \, \kappa\{\mu_0, \sigma_0 \,|\, \mu_y, \sigma\} \\ &= \frac{n}{2} \left(\frac{\mu_0 - \mu_x}{\sigma_0} \right)^2 + \frac{m}{2} \left(\frac{\mu_0 - \mu_y}{\sigma_0} \right)^2 + \frac{n+m}{2} \left(\frac{\sigma^2}{\sigma_0^2} - 1 - \log \frac{\sigma^2}{\sigma_0^2} \right). \end{aligned} \tag{10}$$

Similarly, the logarithmic divergence of $p(\boldsymbol{z} \,|\, \mu_x, \mu_y, \sigma)$ from $p(\boldsymbol{z} \,|\, \mu_0, \mu_0, \sigma_0)$ is

$$\begin{aligned}
&\kappa\{p_{\boldsymbol{z}}(\cdot \,|\, \mu_x, \mu_y, \sigma) \,|\, p_{\boldsymbol{z}}(\cdot \,|\, \mu_0, \mu_0, \sigma_0)\} \\
&\quad = \; n\, \kappa\{\mu_x, \sigma \,|\, \mu_0, \sigma_0\} + m\, \kappa\{\mu_y, \sigma \,|\, \mu_0, \sigma_0\} \\
&\quad = \; \frac{n}{2} \left(\frac{\mu_0 - \mu_x}{\sigma^2} \right)^2 + \frac{m}{2} \left(\frac{\mu_0 - \mu_y}{\sigma^2} \right)^2 + \frac{n+m}{2} \left(\frac{\sigma_0^2}{\sigma^2} - 1 - \log \frac{\sigma_0^2}{\sigma^2} \right) . \qquad (11)
\end{aligned}$$

The minimum of the logarithmic divergence (11) for all $\mu_0 \in \Re$ and $\sigma_0 > 0$ is reached when $\mu_0 = (n\mu_x + m\mu_y)/(n+m)$ and $\sigma_0 = \sigma$, and substitution yields

$$\begin{aligned}
&\inf_{\mu_0 \in \Re,\, \sigma_0 > 0} \kappa\{p_{\boldsymbol{z}}(\cdot \,|\, \mu_0, \mu_0, \sigma_0) \,|\, p_{\boldsymbol{z}}(\cdot \,|\, \mu_x, \mu_y, \sigma)\} \\
&\qquad = \; \frac{n\, m}{2(m+n)} \left(\frac{\mu_x - \mu_y}{\sigma} \right)^2 = \frac{h(n,m)}{4}\, \theta^2 \qquad (12)
\end{aligned}$$

where

$$\frac{1}{h(n,m)} = \frac{1}{2} \left(\frac{1}{n} + \frac{1}{m} \right), \quad h(n,m) = \frac{2\, n\, m}{m+n}\,, \qquad \theta = \frac{\mu_x - \mu_y}{\sigma}\,,$$

which only depends on $h(n,m)$, the harmonic mean of the two sample sizes, and θ^2, the squared standardized distance between the two means.

Similarly, the minimum of the logarithmic divergence (10) for all $\mu_0 \in \Re$ and $\sigma_0 > 0$ is reached when

$$\mu_0 = \frac{n\mu_x + m\mu_y}{n+m}, \quad \sigma_0^2 = \sigma^2 + \frac{mn}{(m+n)^2} \left(\mu_x - \mu_y \right)^2 \,,$$

and substitution yields

$$\begin{aligned}
&\inf_{\mu_0 \in \Re,\, \sigma_0 > 0} \kappa\{p_{\boldsymbol{z}}(\cdot \,|\, \mu_x, \mu_y, \sigma) \,|\, p_{\boldsymbol{z}}(\cdot \,|\, \mu_0, \mu_0, \sigma_0)\} \\
&\qquad = \; \frac{n+m}{2} \log \left[1 + \frac{m\, n}{(n+m)^2} \left(\frac{\mu_x - \mu_y}{\sigma} \right)^2 \right] \\
&\qquad = \; \frac{n+m}{2} \log \left[1 + \frac{h(n,m)}{2(n+m)}\, \theta^2 \right] . \qquad (13)
\end{aligned}$$

3.2.1 The intrinsic discrepancy loss function

Since the first minimized logarithmic discrepancy (Equation 12) may be written as $[(m+n)/2][h(n,m)/(2(m+n))]\, \theta^2$ and, for all positive w, $\log(1+w) < w$, the value of the second minimized logarithmic discrepancy (Equation 13) is always smaller than the first, and therefore, using (5), the required intrinsic loss function is given by (13), so that

$$\delta_{\boldsymbol{z}}\{H_0, (\mu_x, \mu_y, \sigma)\} = \frac{n+m}{2} \log \left[1 + \frac{h(n,m)}{2(n+m)}\, \theta^2 \right], \qquad (14)$$

a logarithmic transformation of the standardized distance $\theta = (\mu_x - \mu_y)/\sigma$ between the two means. The intrinsic loss (14) increases linearly with the total sample size $n + m$, and it is essentially quadratic in θ in a neighbourhood of zero, but it becomes concave for $|\,\theta\,| > (k+1)/\sqrt{k}$, where $k = n/m$ is the ratio of the two sample sizes, an eminently reasonable behaviour which conventional loss functions do not have. For equal sample sizes, $m = n$, this reduces to

$$\delta_{\boldsymbol{z}}\{H_0, (\mu_x, \mu_y, \sigma)\} = n\,\log\left[1 + \frac{1}{4}\left(\frac{\mu_x - \mu_y}{\sigma}\right)^2\right] = n\,\log\left[1 + \frac{\theta^2}{4}\right] \tag{15}$$

a linear function of the sample size n, which behaves as $\theta^2/4$ in a neighbourhood of the origin, but becomes concave for $|\,\theta\,| > 2$.

3.3 The intrinsic statistic

Reference analysis. The intrinsic loss $\delta_{\boldsymbol{z}}\{H_0, (\mu_x, \mu_y, \sigma)\}$ (Equation 14) is a simple piecewise invertible function of θ, the standardized difference of the means. Consequently, the required objective prior is the joint reference prior function $\pi_\theta(\mu_x, \mu_y, \sigma)$ when the standardized difference of the means, $\theta = (\mu_x - \mu_y)/\sigma$, is the quantity of interest.

This may simply be obtained using the orthogonal parametrization $\{\theta, \omega_1, \omega_2\}$, with

$$\omega_1 = \sigma\sqrt{2(m+n)^2 + m\,n\,\theta^2}, \qquad \omega_2 = \mu_y + \frac{n}{n+m}\sigma\,\theta.$$

Indeed, Fisher's information matrix in this parametrization is

$$F(\theta, \omega_1, \omega_2) = \frac{2(m+n)^2 + mn\theta^2}{m+n}\begin{pmatrix} \frac{2mn(m+n)^2}{(2(m+n)^2+mn\theta^2)^2} & 0 & 0 \\ 0 & \omega_1^{-2} & 0 \\ 0 & 0 & (m+n)^2\omega_1^{-2} \end{pmatrix} \tag{16}$$

and, therefore (see Bernardo and Smith 1994, Th. 5.30), the reference prior in that parametrization is $\pi_\theta(\theta, \omega_1, \omega_2) = \pi(\omega_2\,|\,\theta, \omega_1)\pi(\omega_1\,|\,\theta)\pi(\theta)$, where $\pi(\omega_2\,|\,\theta, \omega_1) = 1$, $\pi(\omega_1\,|\,\theta) = \omega_1^{-1}$, and $\pi(\theta)$, the marginal reference prior for the quantity of interest is

$$\pi(\theta) = \left(1 + \frac{m\,n}{2(m+n)^2}\,\theta^2\right)^{-1/2} = \left(1 + \frac{h(n,m)}{4(m+n)}\,\theta^2\right)^{-1/2}. \tag{17}$$

In the original parametrization this becomes

$$\pi_\theta(\mu_x, \mu_y, \sigma) = \frac{1}{\sigma^2}\left(1 + \frac{h(n,m)}{4(m+n)}\left(\frac{\mu_x - \mu_y}{\sigma}\right)^2\right)^{-1/2}. \tag{18}$$

Using Bayes theorem to obtain the joint reference posterior

$$\pi_\theta(\mu_x, \mu_y, \sigma\,|\,\boldsymbol{z}) \propto p(\boldsymbol{z}\,|\,\mu_x, \mu_y, \sigma)\,\pi_\theta(\mu_x, \mu_y, \sigma),$$

changing variables to $\{\theta, \mu_y, \sigma\}$, and integrating out μ_y and σ, produces the (marginal) reference posterior density of the quantity of interest

$$\pi(\theta \,|\, \boldsymbol{z}) = \pi(\theta \,|\, t, m, n) \propto \left(1 + \frac{h(n,m)}{4(m+n)}\,\theta^2\right)^{-1/2} \mathrm{NcSt}\left(t \,\Big|\, \sqrt{\frac{h(n,m)}{2}}\,\theta,\; n+m-2\right) \quad (19)$$

where

$$t = \frac{\bar{x} - \bar{y}}{s\,\sqrt{2/h(n,m)}}\,, \quad s^2 = \frac{n\,s_x^2 + m\,s_y^2}{n+m-2}\,, \quad (20)$$

and $\mathrm{NcSt}(\cdot \,|\, \lambda, \nu)$ is the density of a noncentral Student distribution with noncentrality parameter λ and ν degrees of freedom. The reference posterior (19) is proper provided $n \geq 1$, $m \geq 1$, and $n + m \geq 3$. For further details on this derivation, see Pérez (2005).

Notice that (19) has the form $\pi(\theta \,|\, \boldsymbol{z}) \propto \pi(\theta)\; p(t \,|\, \theta)$. Indeed, the sample mean difference $\bar{x} - \bar{y}$ has a normal sampling distribution with mean $\mu_x - \mu_y$ and variance $h(n,m)\,\sigma^2/2$, the ratio $(n+m-2)s^2/\sigma^2$ has an independent χ^2 distribution with $n+m-2$ degrees of freedom and, therefore, t has a noncentral Student t sampling distribution with non-centrality parameter $\sqrt{h(n,m)/2}\;\theta$ and $n+m-2$ degrees of freedom. This result was to be expected, since the posterior distribution of θ only depends on the data through t and the sample sizes, and reference analysis in known to be consistent under marginalization. It may be verified that the naive prior $\pi(\mu_x, \mu_y, \sigma) = \sigma^{-1}$ produces a marginalization paradox of the type described in Dawid, Stone and Zidek (1973).

3.3.1 The intrinsic statistic

The reference posterior for θ may now be used to obtain the required intrinsic test statistic. Indeed, substituting into (14) yields

$$d(H_0 \,|\, \boldsymbol{z}) = d(H_0 \,|\, t, m, n) = \int_0^\infty \frac{n+m}{2}\; \log\left[1 + \frac{h(n,m)}{2(m+n)}\,\theta^2\right] \pi(\theta \,|\, t, m, n)\, d\theta, \quad (21)$$

where $\pi(\theta \,|\, t, m, n)$ is given by (19). This has no simple analytical expression but may easily be obtained by one-dimensional numerical integration.

3.3.2 Example

The derivation of the appropriate reference prior allows us to draw precise conclusions even when data are extremely scarce. As an illustration, consider a (minimal) sample of three observations with $\boldsymbol{x} = \{4, 6\}$ and $\boldsymbol{y} = \{0\}$, so that $n = 2$, $m = 1$, $\bar{x} = 5$, $\bar{y} = 0$, $s = \sqrt{2}$, $h(n,m) = 4/3$ and $t = 5/\sqrt{3}$. The corresponding exact reference posterior density (Equation 19) is represented in Figure 1.

If may be verified numerically that the reference posterior probability that $\theta < 0$ is

$$\Pr[\theta < 0 \,|\, \boldsymbol{x}, \boldsymbol{y}] = \int_{-\infty}^0 \pi(\theta \,|\, t, m, n)\, d\theta = 0.0438,$$

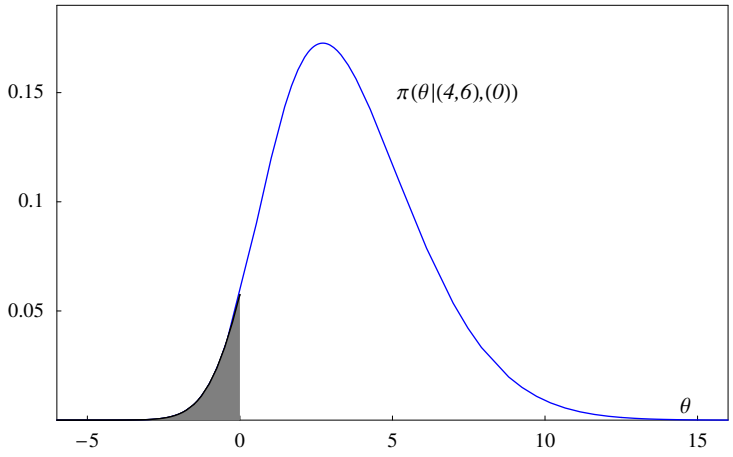

Figure 1: Reference posterior of the standardized difference $\theta = (\mu_x - \mu_y)/\sigma$ given the sample of three observations $\boldsymbol{z} = \{\{4, 6\}, \{0\}\}$.

directly suggesting some (mild) evidence against $\theta = 0$ and, hence, against $\mu_x = \mu_y$. On the other hand, using the formal procedure described above, the numerical value of intrinsic statistic to test $H_0 \equiv \{\mu_x = \mu_y\}$ is

$$d(H_0 \,|\, t, m, n) = \int_0^\infty \frac{3}{2} \, \log\left[1 + \frac{2}{9}\,\theta^2\right] \, \pi(\theta \,|\, t, m, n)\, d\theta = 1.193 = \log[6.776].$$

Thus, given the available data, the expected value of the average (under repeated sampling) of the log-likelihood ratio against H_0 is 1.193 (so that likelihood ratios may be expected to be about 6.8 against H_0), which provides a precise measure of the available evidence against the hypothesis $H_0 \equiv \{\mu_x = \mu_y\}$.

This (moderate) evidence against H_0 is *not* captured by the conventional frequentist analysis of this problem. Indeed, since the sampling distribution of t under H_0 is a standard Student distribution with $n + m - 2$ degrees of freedom, the p-value which corresponds to the two-sided test for H_0 is $2(1-T_{m+n-2}(|t|)$, where T_ν is the cumulative distribution function of an Student distribution with ν degrees of freedom (see, *e.g.*, DeGroot and Schervish 2002, Sec. 8.6). In this case, this produces a p-value of 0.21 which, contrary to the preceding analysis, suggests lack of sufficient evidence in the data against H_0.

3.3.3 Asymptotic approximations

For large sample sizes, $p(t \,|\, \theta)$ converges to a normal distribution $\mathrm{N}(t \,|\, \sqrt{h(n,m)/2}\,\theta,\, 1)$ and the reference posterior of θ converges to

$$\pi\big(\theta \,|\, t, m, n\big) \approx \mathrm{N}\left(\theta \,\middle|\, t\,\sqrt{\frac{2}{h(n,m)}},\, \sqrt{\frac{2}{h(n,m)}}\right) = \mathrm{N}\Big(\theta \,\Big|\, \frac{\bar{x} - \bar{y}}{s}\,,\, \sqrt{\frac{m+n}{m\,n}}\,\Big).$$

A good large sample approximation to the intrinsic statistic is given by

$$d(H_0 \,|\, \boldsymbol{z}) \approx \frac{n+m}{2} \log \left[1 + \frac{1}{n+m}(1+t^2) \right] \leq \frac{1}{2}(1+t^2), \tag{22}$$

where t is given by Equation (20).

3.4 Behaviour under repeated sampling

As it is usually the case with good objective Bayesian procedures, the behaviour under repeated sampling of the intrinsic test statistic is very attractive from a frequentist, repeated sampling perspective.

3.4.1 Sampling distribution of $d(H_0 \,|\, \boldsymbol{z})$ when H_0 is true

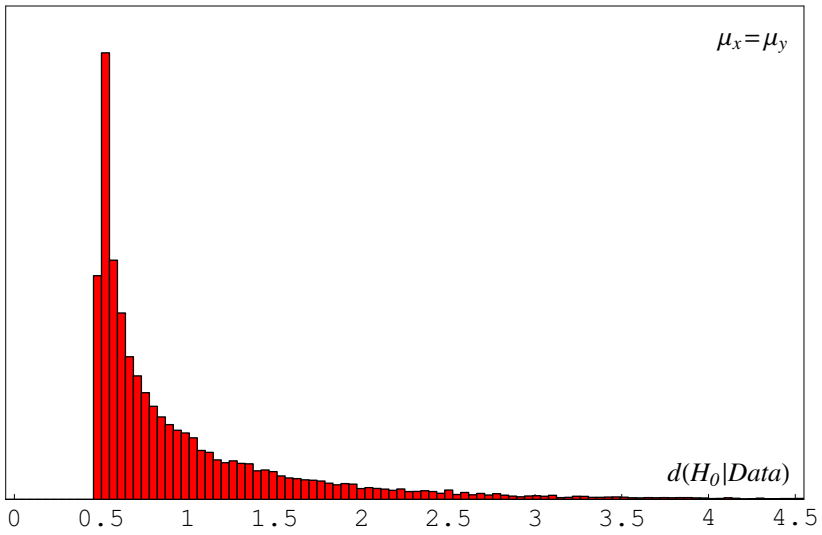

Figure 2: Sampling distribution of $d(H_0 \,|\, \boldsymbol{z})$ under H_0 obtained from 5000 simulations of standard normal random samples of sizes $n = 200$ and $m = 100$.

When H_0 is true, the sampling distribution of t is asymptotically normal $N(t \,|\, 0, 1)$ and, using (22), the sampling distribution of $d(H_0 \,|\, \boldsymbol{z})$ under the null is asymptotically

$$\frac{n+m}{2} \log \left[1 + \frac{1}{n+m} \left(1 + F^1_{n+m-2} \right) \right] \approx \frac{1}{2} \left(1 + \chi^2_1 \right) \tag{23}$$

where F^{α}_{β} is a Snedecor F with α and β degrees of freedom. In particular, the expected value of $d(H_0 \,|\, \boldsymbol{z})$ under sampling when H_0 is true converges to one, and its variance converges to $1/2$ as the sample sizes increase.

As a numerical illustration, Figure 2 represents the histogram of the $d(H_0 \,|\, \boldsymbol{z})$ values obtained from 5000 simulated samples $\boldsymbol{z}_i = \{\boldsymbol{x}_i, \boldsymbol{y}_i\}$, $i = 1, \ldots, 5000$, where $\boldsymbol{x}_i$ and $\boldsymbol{y}_i$ where random samples of sizes $n = 200$ and $m = 100$ respectively from standard normal $\mathrm{N}(x_j \,|\, 0, 1)$ and $\mathrm{N}(y_j \,|\, 0, 1)$ distributions. The resulting sample mean and variances were 0.997 and 0.504 to be compared with the asymptotic values derived from the limiting distribution $\frac{1}{2}(1 + \chi^2_1)$, namely 1 and 1/2.

3.4.2 Sampling distribution of $d(H_0 \,|\, \boldsymbol{z})$ when H_0 is not true

When H_0 is not true, the sampling distribution of the intrinsic test statistic t is asymptotically normal $N(t \,|\, \sqrt{h(n,m)/2}\,\theta, 1)$ and thus, using (22), the sampling distribution of $d(H_0 \,|\, \boldsymbol{z})$ when H_0 is false is asymptotically $\frac{1}{2}[1 + \chi^2_1(h(n,m)\theta^2/2)]$, where $\chi^2_1(\lambda)$ is a non-central chi-squared distribution with one degree of freedom and non-centrality parameter λ. In particular, the expected value of $d(H_0 \,|\, \boldsymbol{z})$ under sampling when H_0 is false (so that $\theta > 0$) is asymptotic to $1 + h(n,m)\theta^2/2$ as the sample sizes increase. Thus, the expected value of the test statistic when H_0 is false increases linearly with the harmonic mean $h(n,m)$ of the two sample sizes. The standard deviation of $d(H_0 \,|\, \boldsymbol{z})$ is asymptotic to $\sqrt{1/2 + h(n,m)\theta^2/2}$ as the sample sizes increase. Thus, for sufficiently large samples, the value of $d(H_0 \,|\, \boldsymbol{z})$ will be larger than any fixed threshold with probability one, which establishes the consistency of the proposed procedure.

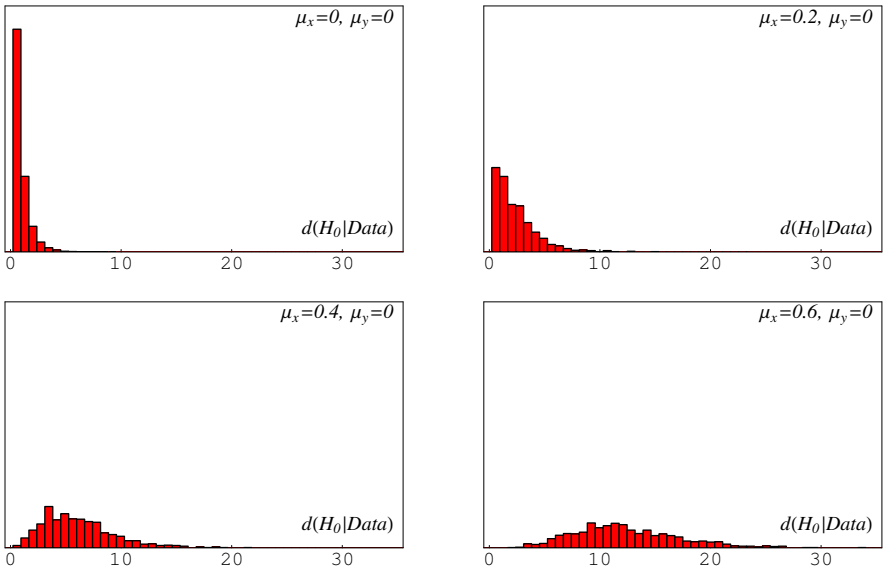

Figure 3: Sampling distribution of $d(H_0 \,|\, \boldsymbol{z})$ obtained from 5000 simulations of normal random samples of sizes $n = 200$ and $m = 100$ from $\mathrm{N}(x \,|\, \mu_x, 1)$ and $\mathrm{N}(y \,|\, 0, 1)$, for $\mu_x \in \{0.0, 0.2, 0.4, 0.6\}$.

As a numerical illustration, Figure 3 represents the histograms (scaled to have area equal to one) of the $d(H_0 \,|\, \boldsymbol{z})$ values obtained from 5000 simulated samples $\boldsymbol{z}_i = \{\boldsymbol{x}_i, \boldsymbol{y}_i\}$, for $i = 1, \ldots, 5000$, where $\boldsymbol{x}_i$ and $\boldsymbol{y}_i$ where random samples of sizes $n = 200$ and $m = 100$ respectively from standard normal $\mathrm{N}(x_j \,|\, \mu_x, 1)$, $\mu_x \in \{0.0, 0.2, 0.4, 0.6\}$, and $\mathrm{N}(y_j \,|\, 0, 1)$ distributions.

3.5 The general case

The methodology described above may be extended to the general case of possibly different variances. Thus, let available data $\boldsymbol{z} = \{\boldsymbol{x}, \boldsymbol{y}\}$, $\boldsymbol{x} = \{x_1, \ldots, x_n\}$, $\boldsymbol{y} = \{y_1, \ldots, y_m\}$, consist of two random samples respectively drawn from $\mathrm{N}(x \,|\, \mu_x, \sigma_x)$ and $\mathrm{N}(y \,|\, \mu_y, \sigma_y)$, and suppose that it is again desired to verify whether or not the observed data $\boldsymbol{z}$ are compatible with the hypothesis H_0 that the two means are equal, that is, whether $\boldsymbol{z}$ could have been drawn from a probability distribution of the family

$$\mathcal{M}_0 \equiv \{p(\boldsymbol{z} \,|\, \mu_0, \mu_0, \sigma_{x0}, \sigma_{y0}),\ \mu_0 \in \Re,\ \sigma_{x0} > 0, \sigma_{y0} > 0\}\,. \tag{24}$$

The relevant intrinsic loss function may then be found to be

$$\begin{aligned} &\delta_{\boldsymbol{z}}\{H_0,\ (\mu_x, \mu_y, \sigma_x, \sigma_y)\} \\ &\qquad \approx \frac{n}{2} \log\left[1 + \frac{\theta_1^2}{(1+r^2)^2}\right] + \frac{m}{2} \log\left[1 + \frac{\theta_2^2}{(1+r^{-2})^2}\right]\,, \end{aligned} \tag{25}$$

where $\theta_1 = (\mu_x - \mu_y)/\sigma_x$ and $\theta_2 = (\mu_x - \mu_y)/\sigma_y$ are the two standardized differences of the means, and $r = r(\sigma_x, \sigma_y, n, m) = (n\,\sigma_y)/(m\,\sigma_x)$ is a measure of the design balance. As one would expect, the intrinsic loss (25) reduces to (15) when $n = m$ and $\sigma_x = \sigma_y$.

Derivation of the exact form of the appropriate joint reference prior $\pi_\delta(\mu_x, \mu_y, \sigma_x, \sigma_y)$ when the quantity of interest is $\delta_{\boldsymbol{z}}\{H_0,\ (\mu_x, \mu_y, \sigma_x, \sigma_y)$, the intrinsic discrepancy between H_0 and the true model is, however, not trivial. Work in this direction is in progress.

References

Berger, J. O. (2006). The case for objective Bayesian analysis (with discussion). *Bayesian Analysis* 1: 385–402 and 457–464. 49

Berger, J. O. and Bernardo, J. M. (1992). On the development of reference priors (with discussion). In *Bayesian Statistics 4*, eds. J. M. Bernardo, J. O. Berger, D. V. Lindley and A. F. M. Smith, 61–77. Oxford, UK: Oxford University Press. 45, 49

Bernardo, J. M. (1979). Reference posterior distributions for Bayesian inference (with discussion). *Journal of the Royal Statistical Society, Series B* 41: 113–147. 45, 49

Bernardo, J. M. (2005a). Reference analysis. In *Handbook of Statistics 25*, eds. D. K. Dey and C. R. Rao, 17–90. Amsterdam: Elsevier. 49

Bernardo, J. M. (2005b). Intrinsic credible regions: An objective Bayesian approach to interval estimation (with discussion). *Test* 14: 317–384. 48

Bernardo, J. M. and Rueda, R. (2002). Bayesian hypothesis testing: A reference approach. *International Statistical Review* 70: 351–372. 45, 47, 48

Bernardo, J. M. and Smith, A. F. M. (1994). *Bayesian Theory*. Chichester, UK: Wiley (2nd. edition to appear in 2007). 46, 49, 52

DeGroot, M. H. and Schervish, M. J. (2002). *Probability and Statistics*, 3rd ed. Reading, MA: Addison-Wesley. 54

Dawid, A. P., Stone, M. and Zidek, J. V. (1973). Marginalization paradoxes in Bayesian and structural inference (with discussion). *Journal of the Royal Statistical Society, Series B* 35: 189–233. 53

Juárez, M. A. (2005). Normal correlation: An objective Bayesian approach. *University of Warwick, UK* (CRiSM Working Paper 05-15). 49

Pérez, S. (2005). Objective Bayesian Methods for Mean Comparison. *Universidad de Valencia, Spain* (Ph.D. Thesis). 53

Acknowledgments

The authors wish to thank the Editor and anonymous AE and referee for helpful comments to the first draft of this paper. Work by Professor Bernardo has been partially funded with Grant MTM2006-07801 of the MEC, Spain.