The Korean Communications in Statistics **Vol. 14** No. 1, 2007, pp. 11–21

Reference-Intrinstic Analysis for the Difference between Two Normal Means

Eun Jin Jang[1], Dal Ho Kim[2] and Kyeong Eun Lee[3]

Abstract

In this paper, we consider a decision-theoretic oriented, objective Bayesian inference for the difference between two normal means with unknown common variance. We derive the Bayesian reference criterion as well as the intrinsic estimator and the credible region which correspond to the intrinsic discrepancy loss and the reference prior. We illustrate our results using real data analysis as well as simulation study.

Keywords: Intrinsic expected loss; Bayesian reference criterion; reference prior; intrinsic estimator; credible region.

1. Introduction

The problem of making inferences about the difference of normal means has been extensively discussed in the statistical literature. Specifically, we assume the data $x_{ij}(i=1,2; j=1,\ldots,n_i)$ are independent and normally distributed with means μ_i and unknown common variance σ^2. As well known, the statistical inferences about $\theta=\mu_1-\mu_2$ are based on t statistic

$$t^* = \frac{(\bar{x}_1-\bar{x}_2)-\theta}{s_p\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}},$$

where $s_p^2 = [\sum_{j=1}^{n_1}(x_{1j}-\bar{x}_1)^2 + \sum_{j=1}^{n_2}(x_{2j}-\bar{x}_2)^2]/(n_1+n_2-2)$ is the pooled variance.

1) Ph.D Graduate, Department of Statistics, Kyungpook National University, Daegu 702-701, Korea.
2) Professor, Department of Statistics, Kyungpook National University, Daegu 702-701, Korea. Correspondence : dalkim@knu.ac.kr
3) Full-time Lecturer, Department of Statistics, Kyungpook National University, Daegu 702-701, Korea.

Recently Bernardo (1999) and Bernardo and Rueda (2002) introduced a new model selection criterion, called the Bayesian Reference Criterion (BRC). Bernardo (1999) takes a decision theoretic approach to developing an objective Bayes solution to test for nested hypotheses. Furthermore, Bernardo and Juárez (2003) addressed intrinsic point estimation and Bernardo (2005) defined intrinsic credible region based on the information-theory based loss function as well as reference prior from an objective Bayesian viewpoint. The procedures merge the use of the reference algorithm (Berger and Bernardo, 1992; Bernardo, 1979) to derive noninformative priors, with the intrinsic discrepancy (Bernardo and Rueda, 2002; Bernardo and Juárez, 2003) as loss function to obtain an objective answer for statistical problem.

The purpose of this article is to develop a decision-theoretic oriented, objective Bayesian answer to the problems of sharp hypothesis testing and both point and region estimation for the difference between two normal means with unknown common variance.

The contents of the remaining paper are as follows. In Section 2, we describe the reference-intrinsic methodology. Also we derive the Bayesian reference criterion, the intrinsic estimator and the credible region for the difference between two normal means with unknown common variance. Comparisons with alternative approaches are carried out in Section 3, using both real data and simulated data. Some concluding remarks are given in Section 4.

2. Reference-Intrinsic Analysis

2.1. The Methodology

Suppose that available data $\boldsymbol{x}$ consist of a random sample $\boldsymbol{x} = \{x_1, \ldots, x_n\}$ from the family $M \equiv \{p(\boldsymbol{x}|\boldsymbol{\theta}, \boldsymbol{\lambda}), \boldsymbol{x} \in X, \boldsymbol{\theta} \in \Theta, \boldsymbol{\lambda} \in \Lambda\}$ where $\boldsymbol{\theta}$ is some vector of interest and $\boldsymbol{\lambda}$ is some vector of nuisance parameters.

The intrinsic discrepancy loss $\delta_{\boldsymbol{x}}\{\widetilde{\boldsymbol{\theta}}, (\boldsymbol{\theta}, \boldsymbol{\lambda})\}$, introduced by Bernardo and Rueda (2002), is basically used to measure the "distance" between the probability density $p(\boldsymbol{x}|\boldsymbol{\theta}, \boldsymbol{\lambda})$ and the family of probability densities $M \equiv \{p(\boldsymbol{x}|\boldsymbol{\theta}, \boldsymbol{\lambda}), \boldsymbol{\lambda} \in \Lambda\}$, defined as

$$\delta_{\boldsymbol{x}}\{\widetilde{\boldsymbol{\theta}}, (\boldsymbol{\theta}, \boldsymbol{\lambda})\} = \inf_{\widetilde{\boldsymbol{\lambda}} \in \Lambda} \delta\{p(\boldsymbol{x}|\boldsymbol{\theta}, \boldsymbol{\lambda}), p(\boldsymbol{x}|\widetilde{\boldsymbol{\theta}}, \widetilde{\boldsymbol{\lambda}})\},$$

where $\delta\{p(\boldsymbol{x}|\boldsymbol{\theta}, \boldsymbol{\lambda}), p(\boldsymbol{x}|\widetilde{\boldsymbol{\theta}}, \widetilde{\boldsymbol{\lambda}})\} = \min\{k(\widetilde{\boldsymbol{\theta}}, \widetilde{\boldsymbol{\lambda}}|\boldsymbol{\theta}, \boldsymbol{\lambda}), k(\boldsymbol{\theta}, \boldsymbol{\lambda}|\widetilde{\boldsymbol{\theta}}, \widetilde{\boldsymbol{\lambda}})\}$ and $k(\boldsymbol{\theta}, \boldsymbol{\lambda}|\widetilde{\boldsymbol{\theta}}, \widetilde{\boldsymbol{\lambda}}) = \int_X p(\boldsymbol{x}|\widetilde{\boldsymbol{\theta}}, \widetilde{\boldsymbol{\lambda}}) \log\{p(\boldsymbol{x}|\widetilde{\boldsymbol{\theta}}, \widetilde{\boldsymbol{\lambda}})/p(\boldsymbol{x}|\boldsymbol{\theta}, \boldsymbol{\lambda})\}d\boldsymbol{x}$, that is the Kullback-Leibler divergences.

Computation of intrinsic loss functions in regular models may be simplified by

$$\delta_{\boldsymbol{x}}\{\widetilde{\boldsymbol{\theta}}, (\boldsymbol{\theta}, \boldsymbol{\lambda})\} = \min\Big\{ \inf_{\widetilde{\boldsymbol{\lambda}} \in \Lambda} k(\widetilde{\boldsymbol{\theta}}, \widetilde{\boldsymbol{\lambda}} | \boldsymbol{\theta}, \boldsymbol{\lambda}), \inf_{\widetilde{\boldsymbol{\lambda}} \in \Lambda} k(\boldsymbol{\theta}, \boldsymbol{\lambda} | \widetilde{\boldsymbol{\theta}}, \widetilde{\boldsymbol{\lambda}}) \Big\}. \tag{2.1}$$

In our problem, the directed divergence $k(\widetilde{\theta}, \widetilde{\lambda}, \widetilde{\sigma} | \theta, \lambda, \sigma)$ is

$$\begin{aligned} k(\widetilde{\theta}, \widetilde{\lambda}, \widetilde{\sigma} | \theta, \lambda, \sigma) &= \int p(\boldsymbol{x}|\theta, \lambda, \sigma^2) \log \frac{p(\boldsymbol{x}|\theta, \lambda, \sigma^2)}{p(\boldsymbol{x}|\widetilde{\theta}, \widetilde{\lambda}, \widetilde{\sigma}^2)} d\boldsymbol{x} \\ &= \frac{n_1}{2} \left\{ \log \frac{\widetilde{\sigma}^2}{\sigma^2} - 1 + \frac{\sigma^2}{\widetilde{\sigma}^2} + \frac{\{(\theta + \lambda) - (\widetilde{\theta} + \widetilde{\lambda})\}^2}{\widetilde{\sigma}^2} \right\} \\ &\quad + \frac{n_2}{2} \left\{ \log \frac{\widetilde{\sigma}^2}{\sigma^2} - 1 + \frac{\sigma^2}{\widetilde{\sigma}^2} + \frac{(\lambda - \widetilde{\lambda})^2}{\widetilde{\sigma}^2} \right\}. \end{aligned} \tag{2.2}$$

As a function of $(\widetilde{\lambda}, \widetilde{\sigma}^2)$, the directed divergence $k(\widetilde{\theta}, \widetilde{\lambda}, \widetilde{\sigma} | \theta, \lambda, \sigma)$ is minimized when $(\widetilde{\lambda}, \widetilde{\sigma}^2)$ takes the value $(\lambda + \{n_1/(n_1+n_2)\}(\theta - \widetilde{\theta}), \sigma^2 + \{n_1 n_2/(n_1+n_2)^2\}(\theta - \widetilde{\theta})^2)$. Thus by substituting in (2.2), the minimum directed divergence is

$$\inf_{\widetilde{\lambda} \in R, \widetilde{\sigma}^2 > 0} k(\widetilde{\theta}, \widetilde{\lambda}, \widetilde{\sigma} | \theta, \lambda, \sigma) = \frac{n_1 + n_2}{2} \log \left[1 + \frac{n_1 n_2}{(n_1 + n_2)^2} \left(\frac{\theta - \widetilde{\theta}}{\sigma} \right)^2 \right].$$

Similarly, the directed divergence $k(\theta, \lambda, \sigma | \widetilde{\theta}, \widetilde{\lambda}, \widetilde{\sigma})$ is given by

$$\begin{aligned} k(\theta, \lambda, \sigma | \widetilde{\theta}, \widetilde{\lambda}, \widetilde{\sigma}) &= \frac{n_1}{2} \left\{ \log \frac{\sigma^2}{\widetilde{\sigma}^2} - 1 + \frac{\widetilde{\sigma}^2}{\sigma^2} + \frac{\{(\widetilde{\theta} + \widetilde{\lambda}) - (\theta + \lambda)\}^2}{\sigma^2} \right\} \\ &\quad + \frac{n_2}{2} \left\{ \log \frac{\sigma^2}{\widetilde{\sigma}^2} - 1 + \frac{\widetilde{\sigma}^2}{\sigma^2} + \frac{(\widetilde{\lambda} - \lambda)^2}{\sigma^2} \right\}. \end{aligned} \tag{2.3}$$

As a function of $(\widetilde{\lambda}, \widetilde{\sigma}^2)$, the directed divergence $k(\theta, \lambda, \sigma | \widetilde{\theta}, \widetilde{\lambda}, \widetilde{\sigma})$ is minimized when $(\widetilde{\lambda}, \widetilde{\sigma}^2)$ takes the value $(\lambda + \{n_1/(n_1+n_2)\}(\theta - \widetilde{\theta}), \sigma^2)$. Thus by substituting in (2.3), the minimum directed divergence is

$$\inf_{\widetilde{\lambda} \in R, \widetilde{\sigma}^2 > 0} k(\theta, \lambda, \sigma | \widetilde{\theta}, \widetilde{\lambda}, \widetilde{\sigma}) = \frac{n_1 + n_2}{2} \cdot \frac{n_1 n_2}{(n_1 + n_2)^2} \left(\frac{\theta - \widetilde{\theta}}{\sigma} \right)^2.$$

Hence, using (2.1) and the fact that, for all $x > 0$, $\log(1+x) \le x$, the intrinsic discrepancy loss $\delta_{\boldsymbol{x}}\{\widetilde{\theta}, (\theta, \lambda, \sigma)\}$ from using $\widetilde{\theta}$ as a proxy for θ is

$$\delta_{\boldsymbol{x}}\{\widetilde{\theta}, (\theta, \lambda, \sigma)\} = \frac{n_1 + n_2}{2} \log \left[1 + \frac{n_1 n_2}{(n_1 + n_2)^2} \left(\frac{\widetilde{\theta} - \theta}{\sigma} \right)^2 \right]. \tag{2.4}$$

2.2. The Bayesian Reference Criterion

The intrinsic discrepancy reference expected loss, or intrinsic expected loss, defined as

$$d(\widetilde{\boldsymbol{\theta}}|\boldsymbol{x}) = \int_{\Theta}\int_{\Lambda} \delta_{\boldsymbol{x}}\{\widetilde{\boldsymbol{\theta}}, (\boldsymbol{\theta}, \boldsymbol{\lambda})\}\pi(\boldsymbol{\theta}, \boldsymbol{\lambda}|\boldsymbol{x})d\boldsymbol{\lambda}d\boldsymbol{\theta}, \tag{2.5}$$

where $\pi(\boldsymbol{\theta}, \boldsymbol{\lambda}|\boldsymbol{x}) \propto p(\boldsymbol{x}|\boldsymbol{\theta}, \boldsymbol{\lambda})\pi(\boldsymbol{\theta}, \boldsymbol{\lambda})$, and $\pi(\boldsymbol{\theta}, \boldsymbol{\lambda})$ is the joint reference prior when $\boldsymbol{\theta}$ is the quantity of interest.

The intrinsic statistic $d(\boldsymbol{\theta}_0|\boldsymbol{x})$ is a measure of the evidence against the simplified model $p(\boldsymbol{x}|\boldsymbol{\theta} = \boldsymbol{\theta}_0, \boldsymbol{\lambda})$ provided by the data $\boldsymbol{x}$. The hypothesis $H_0 \equiv \{\boldsymbol{\theta} = \boldsymbol{\theta}_0\}$ should be rejected if (and only if) the posterior expected loss is sufficiently large. To decide whether or not the precise value $\boldsymbol{\theta}_0$ may be used as a proxy for the unknown value of $\boldsymbol{\theta}$, the Bayesian reference criterion (BRC) might be used as follows:

$$\text{Reject } H_0 \text{ iff } d(\boldsymbol{\theta}_0|\boldsymbol{x}) = \int_{\Theta}\int_{\Lambda} \delta_{\boldsymbol{x}}\{\boldsymbol{\theta}_0, (\boldsymbol{\theta}, \boldsymbol{\lambda})\}\pi(\boldsymbol{\theta}, \boldsymbol{\lambda}|\boldsymbol{x})d\boldsymbol{\lambda}d\boldsymbol{\theta} > d^*.$$

The values of d^* around 2.5 would imply a ratio of $e^{2.5} \approx 12$, providing mild evidence against the null; while values around $5(e^5 \approx 150)$ can be regarded as strong evidence against H_0; values of $d^* \geq 7.5(e^{7.5} \approx 1800)$ can be safely used to reject the null.

In our problem, the reference prior when θ is the parameter of interest is given by $\pi(\theta, \lambda, \sigma) \propto \sigma^{-1}$. Thus the reference posterior distribution of $(\theta, \lambda, \sigma)$ is

$$\pi(\theta, \lambda, \sigma|\boldsymbol{x}) \;\propto\; p(\boldsymbol{x}|\theta, \lambda, \sigma)\pi(\theta, \lambda, \sigma). \tag{2.6}$$

Integrating out the nuisance parameters $\{\lambda, \sigma\}$, it leads to the marginal reference posterior of θ

$$\pi(\theta|\boldsymbol{x}) = St(\theta|\bar{x}_1 - \bar{x}_2, s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}, n_1 + n_2 - 2), \tag{2.7}$$

which is the Student-t distribution with $n_1 + n_2 - 2$ degrees of freedom and location parameter $\bar{x}_1 - \bar{x}_2$, scale parameter $s_p\sqrt{(1/n_1) + (1/n_2)}$. Moreover, the marginal reference posterior density of the precision $r = \sigma^{-2}$ is

$$\pi(r|\boldsymbol{x}) = Ga(r|(n_1 + n_2 - 2)/2, s^2/2),$$

where $s^2 = \sum_{j=1}^{n_1}(x_{1j} - \bar{x}_1)^2 + \sum_{j=1}^{n_2}(x_{2j} - \bar{x}_2)^2$. Thus the posterior mean and variance of the precision r are

$$E[r|\boldsymbol{x}] = \frac{n_1 + n_2 - 2}{s^2}, \quad \text{Var}[r|\boldsymbol{x}] = \frac{2(n_1 + n_2 - 2)}{s^4}. \tag{2.8}$$

Using the intrinsic discrepancy loss (2.4) and the reference posterior (2.6), the intrinsic expected loss $d(\widetilde{\theta}|\boldsymbol{x})$ is given by

$$d(\widetilde{\theta}|\boldsymbol{x}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{0}^{\infty} \delta_{\boldsymbol{x}}\{\widetilde{\theta},(\theta,\lambda,\sigma)\}\pi(\theta,\lambda,\sigma|\boldsymbol{x})d\sigma d\lambda d\theta \tag{2.9}$$

$$= \frac{c(n_1+n_2)}{2}\int_{-\infty}^{\infty}\int_{0}^{\infty} \sigma^{-(n_1+n_2)}\log\left[1+\frac{n_1 n_2}{(n_1+n_2)^2}\left(\frac{\widetilde{\theta}-\theta}{\sigma}\right)^2\right]$$

$$\times \exp\left[-\frac{1}{2\sigma^2}\left\{s^2+\frac{n_1n_2}{n_1+n_2}\left((\bar{x}_1-\bar{x}_2)-\theta\right)^2\right\}\right]d\sigma d\theta,$$

where constant $c = 2^{-(n_1+n_2-3)/2}s^{n_1+n_2-2}\sqrt{(n_1n_2)/(n_1+n_2)}/\{\sqrt{\pi}\,\Gamma((n_1+n_2-2)/2)\}$. Thus, it induced the decision rule, the Bayesian reference criterion (BRC):

$$\text{Reject } \theta=\theta_0 \text{ iff } d(\theta_0|\boldsymbol{x}) > d^*.$$

The above intrinsic expected loss (2.9) may be written in terms of the reference posterior of η,

$$d(\widetilde{\theta}|\boldsymbol{x}) = \int_{-\infty}^{\infty}\frac{n_1+n_2}{2}\log\left[1+\frac{\eta^2}{n_1+n_2}\right]\pi(\eta|\boldsymbol{x})d\eta, \tag{2.10}$$

where $\eta = (\theta-\widetilde{\theta})/\{\sigma\sqrt{(1/n_1)+(1/n_2)}\}$. But η may be written as $a+\beta$ where, as a function of θ and σ, $\beta = \{\theta-(\bar{x}_1-\bar{x}_2)\}/\{\sigma\sqrt{(1/n_1)+(1/n_2)}\}$ has a standard normal reference posterior from the conditional reference posterior of θ given σ, $\pi(\theta|\sigma,\boldsymbol{x}) = N(\theta|\bar{x}_1-\bar{x}_2,\sigma\sqrt{(1/n_1)+(1/n_2)})$, and $a = \{(\bar{x}_1-\bar{x}_2)-\widetilde{\theta}\}/\{\sigma\sqrt{(1/n_1)+(1/n_2)}\}$ is the constant. Hence, the conditional posterior distribution of η^2 given σ is noncentral χ^2 with one degree of freedom and non centrality parameter a^2,

$$\pi(\eta^2|\boldsymbol{x},\sigma) = \chi^2(\eta^2|1,a^2),\quad a^2 = \frac{\{(\bar{x}_1-\bar{x}_2)-\widetilde{\theta}\}^2}{\sigma^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}. \tag{2.11}$$

A simple asymptotic approximation to $d(\widetilde{\theta}|x)$, which provides a direct measure in a log-likelihood ratio scale of the expected loss associated to the use of $\widetilde{\theta}$, may easily be obtained. Indeed, a variation of the delta method shows that, under appropriate regularity conditions, the expectation of some function $y=g(x)$ of a random quantity x with mean μ_x and variance σ_x^2 may be approximated by

$$E[g(x)] \approx g\left[\mu_x+\frac{\sigma_x^2}{2}\frac{g''(\mu_x)}{g'(\mu_x)}\right]. \tag{2.12}$$

From the conditional posterior distribution of η^2, the conditional posterior mean of η^2 is $1+a^2$, and its conditional posterior variance is $2+4a^2$. The posterior expectation of $\log\{1+\eta^2/(n_1+n_2)\}$ required in (2.10) can be approximated by using (2.12).

Theorem 2.1 *The approximation of the intrinsic expected loss is*

$$d(\widetilde{\theta}|\boldsymbol{x}) \approx \frac{n_1+n_2}{2}\log\left[1+\frac{1}{n_1+n_2}(1+t^2)\right], \tag{2.13}$$

where $t = \{(\bar{x}_1-\bar{x}_2)-\widetilde{\theta}\}/\{s_p\sqrt{(1/n_1)+(1/n_2)}\}$.

Proof: From the posterior mean and variance of the precision $r=\sigma^{-2}$ in (2.8), the unconditional posterior mean of η^2 is

$$\begin{aligned} E(\eta^2|\boldsymbol{x}) &= E[E(\eta^2|\boldsymbol{x},r)|\boldsymbol{x}] \\ &= 1+\frac{\{(\bar{x}_1-\bar{x}_2)-\widetilde{\theta}\}^2}{\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}\cdot\frac{n_1+n_2-2}{s^2} \\ &= 1+t^2, \end{aligned}$$

where $t=\{(\bar{x}_1-\bar{x}_2)-\widetilde{\theta}\}/\{s_p\sqrt{(1/n_1)+(1/n_2)}\}$ and the unconditional posterior variance of η^2 is

$$\begin{aligned} \mathrm{Var}(\eta^2|\boldsymbol{x}) &= \mathrm{Var}[E(\eta^2|r,\boldsymbol{x})|\boldsymbol{x}]+E[\mathrm{Var}(\eta^2|r,\boldsymbol{x})|\boldsymbol{x}] \\ &= 2+4t^2+\frac{2t^4}{n_1+n_2-2}. \end{aligned}$$

Thus using (2.12), progressively cruder, the approximated posterior expectation is

$$\begin{aligned} d(\widetilde{\theta}|\boldsymbol{x}) &= \int_{-\infty}^{\infty}\frac{n_1+n_2}{2}\log\left[1+\frac{\eta^2}{n_1+n_2}\right]\pi(\eta|\boldsymbol{x})d\eta \\ &\approx \frac{n_1+n_2}{2}\log\left[1+\frac{1}{n_1+n_2}\right. \\ &\qquad \left.\times\frac{(n_1+n_2)(n_1+n_2-1)(1+t^2)+(n_1+n_2-2)t^4}{(n_1+n_2-1)(n_1+n_2+1+t^2)}\right] \\ &\approx \frac{n_1+n_2}{2}\log\left[1+\frac{1}{n_1+n_2}(1+t^2)\right]. \end{aligned}$$

□

By Taylor's series expansion, simpler approximation is

$$d(\widetilde{\theta}|\boldsymbol{x}) \approx \frac{n_1+n_2}{2}\log\left[1+\frac{1}{n_1+n_2}(1+t^2)\right] \approx \frac{1}{2}(1+t^2).$$

Both the intrinsic statistic (2.9) and its approximation (2.13) are represented in Figure 2.1, calculated from simulated data with $\theta = 0$ and for several sample sizes of (n_1, n_2). We can see that for small (n_1, n_2), it is not possible to reject almost any value of the parameter and that the criterion becomes more discriminating as the sample size increases. Moreover, the approximated intrinsic statistic (2.13) works well even for moderated sample sizes.

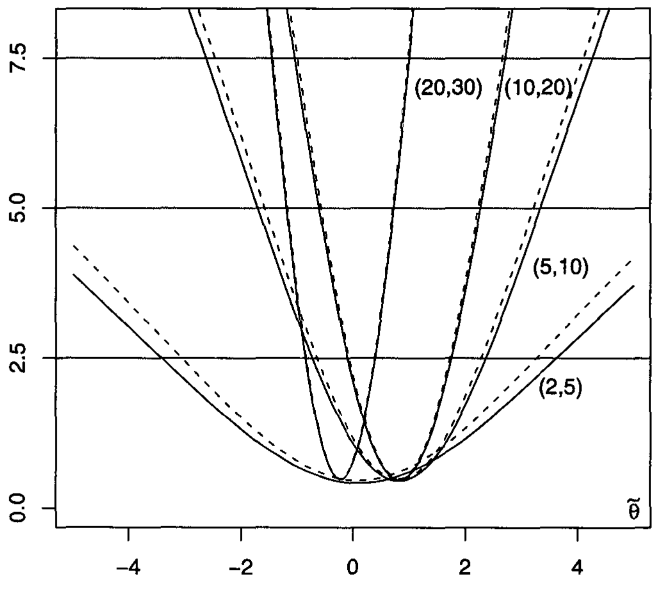

Figure 2.1: The intrinsic statistic $d(\widetilde{\theta}|\boldsymbol{x})$ for the difference between two normal means (solid line) and its approximation (dashed line), fixing $\theta = \mu_1 - \mu_2 = 0$.

2.3. The Intrinsic Estimation

Bayes estimates are those which minimize the expected posterior loss. The intrinsic estimate is the Bayes estimate which corresponds to the intrinsic discrepancy loss and the reference posterior distribution. Introduced by Bernardo and Juárez (2003), this is a completely general objective Bayesian estimator which is invariant under reparametrization. The intrinsic estimate of $\boldsymbol{\theta}$

$$\widetilde{\boldsymbol{\theta}}_{int}(\boldsymbol{x}) = \operatorname*{argmin}_{\widetilde{\boldsymbol{\theta}}\in\Theta} d(\widetilde{\boldsymbol{\theta}}|\boldsymbol{x})$$

is that parameter value which minimizes the reference posterior expected intrinsic loss (2.5).

Bayesian region estimation is typically based on posterior credible regions, *i.e.*, sets of $\boldsymbol{\theta}$ values with pre-specified posterior probabilities. The p-credible intrinsic region is the lowest posterior loss p-credible region which corresponds to the intrinsic discrepancy loss and the reference prior. An intrinsic p-credible region is a subset $R_p^{int} = R_p^{int}(\boldsymbol{x}, \Theta)$ on the parameter space such that,

$$\int_{R_p^{int}} p(\boldsymbol{\theta}|\boldsymbol{x})d\boldsymbol{\theta} = p, \ \forall\, \widetilde{\boldsymbol{\theta}}_i \in R_p^{int}, \forall\, \widetilde{\boldsymbol{\theta}}_j \notin R_p^{int}, \ \ d(\widetilde{\boldsymbol{\theta}}_i|\boldsymbol{x}) \le d(\widetilde{\boldsymbol{\theta}}_j|\boldsymbol{x}),$$

where $d(\widetilde{\boldsymbol{\theta}}_i|\boldsymbol{x})$ is the intrinsic expected loss (2.5).

In our problem, the intrinsic expected loss $d(\widetilde{\theta}|\boldsymbol{x})$ only depends on $\widetilde{\theta}$ through $\{(\bar{x}_1 - \bar{x}_2) - \widetilde{\theta}\}^2$, and increases with $\{(\bar{x}_1 - \bar{x}_2) - \widetilde{\theta}\}^2$ from (2.11); therefore, the intrinsic estimator of θ is

$$\begin{aligned}\widetilde{\theta}_{int}(\boldsymbol{x}) &= \operatorname*{argmin}_{\widetilde{\theta}\in R} d(\widetilde{\theta}|\boldsymbol{x}) \\ &= \operatorname*{argmin}_{\widetilde{\theta}\in R}\{(\bar{x}_1 - \bar{x}_2) - \widetilde{\theta}\}^2 \\ &= \bar{x}_1 - \bar{x}_2.\end{aligned}$$

Moreover, $d(\widetilde{\theta}|\boldsymbol{x})$ is symmetric around $\bar{x}_1 - \bar{x}_2$ and, hence, all intrinsic credible regions must be centered at $\bar{x}_1 - \bar{x}_2$. Since the marginal reference posterior of θ is the Student distribution given by (2.7), the intrinsic p-credible regions are just the usual Student-t HPD p-credible intervals

$$R_p^{int}(\boldsymbol{x}, R) = (\bar{x}_1 - \bar{x}_2) \pm t_{p,n_1+n_2-2}\, s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}, \tag{2.14}$$

where t_{p,n_1+n_2-2} is the $(p+1)/2$ quantile of a standard Student-t with $n_1 + n_2 - 2$ degrees of freedom.

It immediately follows from (2.14) that R_p^{int} consist of the set of $\widetilde{\theta}$ values such that $\{(\bar{x}_1 - \bar{x}_2) - \widetilde{\theta}\}/\{s_p\sqrt{(1/n_1) + (1/n_2)}\}$ belongs to a probability p centred interval of a standard Student-t with $n_1 + n_2 - 2$ degrees of freedom. But, as a function of the data $\boldsymbol{x}$, the sampling distribution of

$$t(\boldsymbol{x}) = \{(\bar{x}_1 - \bar{x}_2) - \widetilde{\theta}\} / \left(s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right)$$

is also a standard Student-t with $n_1 + n_2 - 2$ degrees of freedom. Hence, for all sample sizes, the expected coverage under sampling of the p-credible intervals is exactly p, and the intrinsic credible regions are exact frequentist confidence intervals.

3. Numerical Analysis

3.1. Example

The data "Calcium and Blood Pressure Story" in DASL(lib.stat.cmu.edu/DASL), which contains a subset of the data shown in Lyle *et al.* (1987), consist of blood pressure measurements on a subgroup of 21 African-American subjects, 10 who have taken calcium supplements and 11 who have taken placebo. The primary analysis variable is the blood pressure difference ("begin" minus "end"). Summary statistics are as follows:

Group	n	Mean	StdDev
Calcium	10	5.0000	8.7433
Placebo	11	−0.2727	5.9007

Here, $s_p = 7.385$ and $t = 1.634$; the positive t-value suggests calcium is beneficial for reducing blood pressure. The two-sided frequentist p-value is $p = 0.1187$.

The exact value of the expected intrinsic loss by numerical integration using (2.9) is 1.5396 and its approximation using (2.13) is 1.6913. According to the BRC using the threshold value $d^* = 2.5$, we cannot reject the hypothesis $H_0 : \theta = 0$, so there is insufficient evidence to indicate that the change in blood pressure between the treatment and placebo groups are different.

3.2. Simulation Study

In order to compare the performance under homogeneous repeated sampling of the BRC with their frequentist counterparts, ten thousand simulations were carried for several sample sizes (n_1, n_2). Table 3.1 summarises this comparison.

The third and fourth columns (BRC and t-test) report the relative number of times when the hypothesis $\theta = 0$ was rejected under each criterion, using the threshold values of $d^* = 2.4207$ for the BRC, and $\alpha = 0.05$ for the t-test.

Table 3.1: Comparison of the behaviour under repeated sampling of the BRC and t-test.

(n_1, n_2)	θ	BRC	t-test
	−3	0.7824	0.8175
(2,5)	0	0.0403	0.0480
	3	0.7816	0.8186
	−3	0.9986	0.9990
(5,10)	0	0.0388	0.0489
	3	0.9977	0.9985
	−3	1.0000	1.0000
(10,20)	0	0.0408	0.0487
	3	1.0000	1.0000
	−3	1.0000	1.0000
(20,30)	0	0.0438	0.0494
	3	1.0000	1.0000

From Table 3.1, we confirm that the frequentist p-value and the BRC p-value are comparable and the power of both tests increases with sample size. Moreover, underH_0, the BRC type-1 errors are 0.0403, 0.0388, 0.0408, 0.0438 for each sample sizes, which are smaller than the corresponding frequentist type-1 errors.

Table 3.2: The mean value and standard deviation of the intrinsic estimators, the coverage probability and expected length of the intrinsic 0.95-credible intervals for several sample sizes (n_1, n_2).

(n_1, n_2)	Inrinsic Estimators		Coverage	Expected
	Mean	StdDev	Probability	Length
(2,5)	−0.0052	0.8281	0.9498	4.0909
(5,5)	−0.0036	0.6295	0.9518	2.8275
(5,10)	−0.0042	0.5450	0.9506	2.3233
(10,10)	0.0058	0.4460	0.9494	1.8546
(10,20)	0.0083	0.3887	0.9482	1.5717
(20,30)	0.0020	0.2898	0.9513	1.1549

To explore the performance of both intrinsic estimator and intrinsic credible region, we simulated ten thousand data sets for several sample sizes (n_1, n_2) from a normal distributions $N(x_1|1,1)$ and $N(x_2|1,1)$, and calculated the mean value and standard deviation of the intrinsic estimators as well as both the coverage probability and the expected length of the intrinsic 0.95-credible intervals. The results are summarized in Table 3.2.

This table provides that the intrinsic statistic appears to be consistently closer to the true value of the parameter, the frequentist coverage of reference p-credible regions is indeed approximately equal to p for all sample sizes and the expected length decrease according to the increase of sample sizes.

4. Concluding Remarks

The reference-intrinsic approach described provides a powerful alternative to point and interval estimation and sharp hypothesis testing, with a clear interpretation in terms of information units. This study considers the reference-intrinsic approach for the difference between two normal means with unknown common variances. In future we will derive the Bayesian reference criterion, the intrinsic estimator and the credible region which corresponds to the intrinsic discrepancy loss and the reference prior in two different variances case, and show the possible extension of reference-intrinsic analysis in more complex setup.

References

Berger, J. O. and Bernardo, J. M. (1992). On the development of reference priors(with discussion). *Bayesian Statistics 4* (J. M. Bernardo, *et al.* eds.), 35–60, Oxford University Press, New York.

Bernardo, J. M. (1979). Reference posterior distributions for Bayesian inference. *Journal of the Royal Statistical Society*, Ser. B, **41**, 113–147.

Bernardo, J. M. (1999). Nested hypothesis testing: the Bayesian reference criterion. *Bayesian Statistics 6* (J. M. Bernardo, *et al.* eds.), 101–130, Oxford University Press, New York.

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

Bernardo, J. M. and Juárez, M. A. (2003). Intrinsic estimation. *Bayesian Statistics 7* (J. M. Bernardo, *et al.* eds.), 465–476, Oxford University Press, New York.

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

Lyle, R. M., Melby, C. L., Hyner, G. C., Edmondson, J. W., Miller, J. Z. and Weinberger, M. H. (1987). Blood pressure and metabolic effects of calcium supplementation in normotensive white and black Men. *Journal of the American Medical Association*, **257**, 1772–1776.

[Received October 2006, Accepted October 2006]