

Robustness of the Shrinkage Estimator for the Relative Potency in the Combination of Multivariate Bioassays

Ding-Geng (Din) Chen

School of Social Work

University of North Carolina-Chapel Hill, NC 27599, USA

Department of Statistics, University of Pretoria, Pretoria, South Africa

Email: DrDG.Chen@gmail.com

September 20, 2016

Abstract

This paper investigates the robustness of the shrinkage Bayesian estimator for the relative potency parameter in the combinations of multivariate bioassays proposed in Chen et al.(1999), which incorporated prior information on the model parameters based on Jeffreys' rules. This investigation is carried out for the families of t -distribution and Cauchy-distribution based on the characteristics of bioassay theory since the t -distribution approaches the normal distribution which is the most commonly used distribution in the applications of bioassay as the degrees of freedom increases and the t -distribution approaches the Cauchy-distribution as the degrees of freedom approaches 1 which is also an important distribution in bioassay. A real data is used to illustrate the application of this investigation. This analysis further supports the application of the shrinkage Bayesian estimator to the theory of bioassay along with the empirical Bayesian estimator.

Key Words: Bayesian analysis; Noninformative Prior; Posterior likelihood function; Posterior mode; Relative Potency.

1 Introduction

Estimation of relative potency in biological assay (i.e. bioassay) is very important in pharmaceutical and toxicological sciences. A variety of methods for estimating relative potency can be found in the literature and the references in Finney (1978) and Govindarajulu (2001). Most of these methods can be formulated in terms of a general linear model for data following a normal distribution, or in a generalized linear model for other types of data. In the situation of multi-response experiments, multivariate statistics can be used as in Laska et al. (1985), Srivastava (1986) and Chen et al. (1999).

The combination of bioassays arises when the same or similar experiment is performed by multi-laboratories and multi-centers in the calibration of national/international standards for a particular substance (Rose and Gaines-Das, 1998). With the studies of combination of bioassays, most results appear in Bennett (1962), Armitage (1970), Meisner et al. (1986), William (1988), Chen et al. (1999), Xiong and Chen (2007), and Chen (2007).

Chen et al. (1999) proposed a novel empirical Bayesian estimator (EBE) and a shrinkage Bayesian estimator (SBE) for the relative potency from the combination of several multivariate bioassays by incorporating prior information on the model parameters based on Jeffreys' rules. The EBE can account for any extra-variability among the bioassays, and if the extra-variability is zero, the EBE reduces to the maximum likelihood estimator for combinations of bioassays. The SBE estimator turned out to be a weighted average of the prior information and the estimator from each bioassay with the weights depending on the prior variance, which is not uncommon in Bayesian framework. The key advantage of this shrinkage Bayesian estimator is that it can be written in closed form and therefore no iterative process is involved to yield the estimate.

However any prudent Bayesian analysis should include an investigation of the robustness of the posterior distribution to the specifications of commonly used prior distributions. For the proposed shrinkage Bayesian estimator, Chen et al.(1999) embedded the Bayesian

theory into a bioassay framework by using the Jeffreys' theory for the model parameters. This technique takes the advantage that some information is known about the log relative potency parameter μ , while little information is known about the other parameters from the model specification of the combination of multivariate bioassays. Therefore, the robust investigation for the shrinkage Bayesian estimator is given to different specifications of the prior distribution for the parameter of μ .

In this paper, the families of t - and Cauchy-distributions will be investigated for the robustness analysis since practically t -distribution is considerably important, especially for the heavy-tail distribution. As the degrees of freedom increases, the t -distribution approaches the normal distribution, which is the most commonly situation in the applications of bioassay as illustrated in Finney (1978) and Chen et al. (1999). On the other hand, as the degrees of freedom approaches 1, the t -distribution approaches the Cauchy distribution, which is also an important distribution in bioassay since the log relative potency estimator is a ratio of two normal variables, which is distributed as a Cauchy distribution.

This paper is organized as follows. Section 2 introduces the bioassay model and shrinkage Bayesian estimator based on Jeffreys' prior from Chen et al. (1999) and the investigation of two families of t -distribution and Cauchy-distribution is given in Section 3. A real data analysis is illustrated in Section 4 with further discussions given in Section 5.

2 Combinations of Bioassays and Shrinkage Estimator

2.1 Model Specification

The experimental model for i th ($i = 1, \dots, k$) multivariate bioassay in the combinations of k multivariate bioassays to estimate the relative potency of the test (t) preparation to the standard (s) preparation is usually of the following form:

$$\mathbf{y}_{si} = \boldsymbol{\alpha}_{si} + \boldsymbol{\beta}_{si}x_{si} + \boldsymbol{\epsilon}_{si}$$

$$\mathbf{y}_{ti} = \boldsymbol{\alpha}_{ti} + \boldsymbol{\beta}_{ti}x_{ti} + \boldsymbol{\epsilon}_{ti} \quad (2.1)$$

where $\mathbf{y}_{si}, \mathbf{y}_{ti} \in \mathbf{R}^p$, $p \geq 1$, are multivariate responses in the dose-response relationship, $\boldsymbol{\alpha}_{si}, \boldsymbol{\alpha}_{ti}, \boldsymbol{\beta}_{si}, \boldsymbol{\beta}_{ti} \in \mathbf{R}^p$ are the multivariate parameters, and x_{si}, x_{ti} are the scalar inputs as function of dose levels.

This model can then be casted in a multivariate regression framework

$$\mathbf{Y}_i = \boldsymbol{\Psi}_i \mathbf{X}_i + \boldsymbol{\epsilon}_i, \quad (2.2)$$

where $\mathbf{Y}_i = (\mathbf{y}_{s1}, \dots, \mathbf{y}_{sn_{si}}, \mathbf{y}_{t1}, \dots, \mathbf{y}_{tn_{ti}})$ is the $p \times n_i$ matrix of responses with $n_i = n_{si} + n_{ti}$, where n_{si}, n_{ti} are the number of observations in the i th bioassay for the standard and test preparations, respectively; and $\boldsymbol{\Psi}_i = (\boldsymbol{\alpha}_i, \boldsymbol{\delta}_i, \boldsymbol{\beta}_i)$ is the $p \times 3$ matrix of parameters. For all $i = 1, \dots, k$, $\boldsymbol{\epsilon}_i$'s are assumed *i.i.d.* $\sim N_p(0, \mathbf{V})$, where \mathbf{V} is $p \times p$ positive definite symmetric matrix.

Different assumptions on the model parameters can produce different bioassay models. There are two commonly-known bioassays with one as parallel-line bioassay and another as slope-ratio bioassay. In the parallel-line bioassays where the two slopes in equation (2.1) on both test preparation (i.e. $\boldsymbol{\beta}_{ti}$) and standard preparation (i.e. $\boldsymbol{\beta}_{si}$) are equal (denoted by $\boldsymbol{\beta}_i$) where $\boldsymbol{\beta}_i = \boldsymbol{\beta}_{si} = \boldsymbol{\beta}_{ti}$, (see for example, Carter and Hubert (1985) for the test of the parallelism), but with different intercepts so that $\boldsymbol{\alpha}_i = \boldsymbol{\alpha}_{si}$; $\delta_{ij} = \alpha_{tij} - \alpha_{ij} = \mu_{ij}\beta_{ij}$, for all bioassays $i = 1, \dots, k$, and all responses $j = 1, \dots, p$, where μ_{ij} are scalar parameters, which are the log relative potencies of the i th bioassay and j th response variable. The slope-ratio bioassay has slightly different parameterizations where the slopes are different but the intercepts are the same, such that $\boldsymbol{\alpha}_i = \boldsymbol{\alpha}_{si} = \boldsymbol{\alpha}_{ti}$, $\boldsymbol{\delta}_i = \boldsymbol{\delta}_{ti} = \mu_{ij}\boldsymbol{\beta}_{si}$ where μ_{ij} are the relative potencies in i th bioassay and j th response variable. We will illustrate parallel-line bioassay in this paper and the similar discussions can be made for slope-ratio bioassay.

In the situation of homogeneity, $\mu_{ij} = \mu$ for all i and j which can be tested by the likelihood-ratio test. The $p \times n_i$ random matrix $\boldsymbol{\epsilon}_i$ consists of n_i columns of multivariate

normal random vectors $N_p(0, \mathbf{V})$; and \mathbf{X}_i is the $3 \times n_i$ design matrix defined as $\mathbf{X}_i = \begin{pmatrix} 1 & \cdots & 1 & 1 & \cdots & 1 \\ 0 & \cdots & 0 & 1 & \cdots & 1 \\ x_{s1} & \cdots & x_{sn_{si}} & x_{t1} & \cdots & x_{tn_{ti}} \end{pmatrix}$, where x 's are log dose levels. Then the multivariate regression technique can be used to obtain the estimators

$$\hat{\Psi}_i = \mathbf{Y}_i \mathbf{X}_i' (\mathbf{X}_i \mathbf{X}_i')^{-1} \quad (2.3)$$

$$\mathbf{S}_i = m_i^{-1} \mathbf{Y}_i \left[\mathbf{I} - \mathbf{X}_i' (\mathbf{X}_i \mathbf{X}_i')^{-1} \mathbf{X}_i \right] \mathbf{Y}_i' \quad (2.4)$$

as independent estimators of Ψ_i and \mathbf{V} , where $m_i = n_i - 3$ and \mathbf{X}' is the transpose of matrix \mathbf{X} . Since our interest is in the relative potency parameter μ_{ij} , the second and third columns of $\hat{\Psi}_i$ will be the relevant ones, while the parameters in the first column, $\hat{\alpha}_i$, are the nuisance parameters. Therefore, we will restrict our attention to $\begin{bmatrix} \hat{\delta}_i \\ \hat{\beta}_i \end{bmatrix} \sim N_{2p} \left(\begin{bmatrix} \delta_i \\ \beta_i \end{bmatrix}, \mathbf{V} \otimes \mathbf{A}_i \right)$, where \mathbf{A}_i is the lower 2×2 portion of $(\mathbf{X}_i \mathbf{X}_i')^{-1}$.

To block diagonalize the matrix $\mathbf{V} \otimes \mathbf{A}_i$ in order to make $\hat{\delta}_i$ and $\hat{\beta}_i$ independent, let's define $\mathbf{T}_i = \begin{pmatrix} t_{11i} & t_{12i} \\ 0 & t_{22i} \end{pmatrix}$, such that $\mathbf{T}_i \mathbf{A}_i \mathbf{T}_i' = \mathbf{I}$. Then, $t_{11i} = \frac{a_{22i}^{1/2}}{(a_{11i} a_{22i} - a_{12i}^2)^{1/2}}$, $t_{12i} = \frac{-a_{12i} a_{22i}^{-1/2}}{(a_{11i} a_{22i} - a_{12i}^2)^{1/2}}$ and $t_{22i} = a_{22i}^{-1/2}$. Let $\mathbf{u}_i = t_{11i} \hat{\delta}_i + t_{12i} \hat{\beta}_i$ and $\mathbf{v}_i = t_{22i} \hat{\beta}_i$, then the distribution of \mathbf{u}_i and \mathbf{v}_i is

$$\begin{bmatrix} \mathbf{u}_i \\ \mathbf{v}_i \end{bmatrix} \sim N_{2p} \left(\begin{bmatrix} t_{11i} \delta_i + t_{12i} \beta_i \\ t_{22i} \beta_i \end{bmatrix}, \mathbf{V} \otimes \mathbf{I}_2 \right). \quad (2.5)$$

In addition, \mathbf{S}_i in equation (2.4) is distributed as $\mathbf{S}_i \sim m_i^{-1} W_p(\mathbf{V}, m_i) = W_p(\mathbf{V}/m_i, m_i)$, which is the Wishart distribution, a multivariate generalization of the Chi-square distribution, with the pdf given by $C_i |\mathbf{V}|^{-\frac{m_i}{2}} |\mathbf{S}_i|^{\frac{m_i-p-1}{2}} \text{etr}(-\frac{m_i}{2} \mathbf{V}^{-1} \mathbf{S}_i)$, where etr is the notation for exponent of the trace, $|\mathbf{V}|$ is the determinant of matrix \mathbf{V} and C_i is a constant of integration chosen so that the total probability is 1 (Srivastava and Carter, 1983). The pooled estimator for \mathbf{V} is then $\mathbf{S} = (\sum_{i=1}^k m_i \mathbf{S}_i) / \sum_{i=1}^k m_i$.

2.2 Shrinkage Bayesian Estimator for the Log Relative Potency

The shrinkage estimator was derived in Chen et al. (1999) based on Jeffreys' prior (Hoadley, 1970) of $\pi(\mu, \boldsymbol{\beta}_i, \mathbf{V}) \propto \frac{h(\mu)}{|\mathbf{V}|^{(p+1)/2}}$ with $h(\mu) \sim N(\mu_0, \tau^2)$, for some μ_0 and $\tau^2 > 0$.

It was shown that $\mu | X_i, Y_i, \boldsymbol{\beta}_i, \mathbf{V} \sim N(\mu_p, \tau_p^2)$, where

$$\mu_p = \frac{\tau^2 \sum_{i=1}^k t_{11i} \boldsymbol{\beta}_i' \mathbf{V}^{-1} (t_{11i} \hat{\boldsymbol{\delta}}_i + t_{12i} \hat{\boldsymbol{\beta}}_i - t_{12i} \boldsymbol{\beta}_i) + \mu_0}{\tau^2 \sum_{i=1}^k t_{11i}^2 \boldsymbol{\beta}_i' \mathbf{V}^{-1} \boldsymbol{\beta}_i + 1}$$

$$\tau_p = \left(\sum_{i=1}^k t_{11i}^2 \boldsymbol{\beta}_i' \mathbf{V}^{-1} \boldsymbol{\beta}_i + \frac{1}{\tau^2} \right)^{-\frac{1}{2}}.$$

Then the shrinkage Bayesian estimator (SBE) for the log relative potency, μ , can be obtained using the estimates in equations (2.3) and (2.4) for $\boldsymbol{\beta}_i$ and \mathbf{V} as follows:

$$\mu_{SBE} = \frac{\tau^2 \sum_{i=1}^k t_{11i}^2 \hat{\boldsymbol{\beta}}_i' \mathbf{S}_i^{-1} \hat{\boldsymbol{\delta}}_i + \mu_0}{\tau^2 \sum_{i=1}^k t_{11i}^2 \hat{\boldsymbol{\beta}}_i' \mathbf{S}_i^{-1} \hat{\boldsymbol{\beta}}_i + 1} \quad (2.6)$$

$$\tau_{SBE} = \left(\sum_{i=1}^k t_{11i}^2 \hat{\boldsymbol{\beta}}_i' \mathbf{S}_i^{-1} \hat{\boldsymbol{\beta}}_i + \frac{1}{\tau^2} \right)^{-\frac{1}{2}}. \quad (2.7)$$

Therefore the credible region with $(1-\alpha)100\%$ coefficient for μ can be expressed as $\mu_{SBE} \pm z_{\alpha/2} \tau_{SBE}$ where $z_{\alpha/2}$ is the usual percentile point such that $1 - \Phi(z_{\alpha/2}) = \alpha/2$, and where $\Phi(\cdot)$ is the cumulative distribution function for a standard normal random variable.

The corresponding noninformative shrinkage Bayesian estimator can be obtained by letting τ approaches infinite as follows:

$$\hat{\mu} = \frac{\sum_{i=1}^k t_{11i}^2 \hat{\boldsymbol{\beta}}_i' \mathbf{S}_i^{-1} \hat{\boldsymbol{\delta}}_i}{\sum_{i=1}^k t_{11i}^2 \hat{\boldsymbol{\beta}}_i' \mathbf{S}_i^{-1} \hat{\boldsymbol{\beta}}_i} \quad (2.8)$$

$$\hat{\tau} = \left(\sum_{i=1}^k t_{11i}^2 \hat{\boldsymbol{\beta}}_i' \mathbf{S}_i^{-1} \hat{\boldsymbol{\beta}}_i \right)^{-\frac{1}{2}}. \quad (2.9)$$

Similar $(1-\alpha)100\%$ credible region for μ can be expressed as $\hat{\mu} \pm z_{\alpha/2} \hat{\tau}$.

The advantage for this shrinkage Bayesian estimator and its interval estimator described above is that they can be calculated explicitly from the model parameter estimates which have been obtained from equations (2.3) and (2.4). No iterative process is involved.

3 Robust Analysis for the Shrinkage Bayesian Estimator

Since, in most practical situations, some information can be known about the relative potency, μ and relatively little information is known about β_i and \mathbf{V} , then by Jeffreys' rules, a suitable prior distribution would be $\pi(\mu, \beta_i, \mathbf{V}) \propto \frac{h(\mu)}{|\mathbf{V}|^{(p+1)/2}}$, where $h(\mu)$ is a proper prior distribution for μ . Then the posterior distribution of μ , conditional on $\mathbf{X}_i, \mathbf{Y}_i, \beta_i$ and \mathbf{V} can be shown from equation (2.5) as follows:

$$\begin{aligned} \mu | \mathbf{X}_i, \mathbf{Y}_i, \beta_i, \mathbf{V} &\propto \pi(\mu, \beta_i, \mathbf{V}) \mathbf{L}(\mu, \beta_i, \mathbf{V}) \\ &\propto h(\mu) \exp \left\{ -\frac{1}{2} \sum_{i=1}^k (t_{11i} \hat{\delta}_i + t_{12i} \hat{\beta}_i - t_{11i} \mu \beta_i - t_{12i} \beta_i)' \mathbf{V}^{-1} (t_{11i} \hat{\delta}_i + t_{12i} \hat{\beta}_i - t_{11i} \mu \beta_i - t_{12i} \beta_i) \right\}. \end{aligned} \quad (3.1)$$

The posterior function (3.1) incorporates all the information from the data by the likelihood function and prior information from $h(\mu)$. Parallel to the maximum likelihood theory, an estimate of the log relative potency can be obtained by maximizing the posterior likelihood function, i.e. estimating the log relative potency parameter by the mode of the posterior likelihood function. Another alternative is to estimate the log relative potency parameter by the *median* or *mean* of the posterior likelihood function. This estimator is usually referred as the *posterior Bayesian estimator* (Mood et al. 1973). However if the prior distribution is normal, these two alternatives produce same estimators, since the posterior distribution is also normally distributed. Actually, if $h(\mu)$ is the normal distribution $N(\mu_0, \tau^2)$ and after simple manipulations, the same results in Section 2.2 can be derived.

For the *posterior Bayesian estimator*, numerical integration is involved to obtain this estimator. Therefore, to be concise for this robustness analysis and also parallel to the conventional likelihood estimation, the mode of the posterior likelihood function (3.1) is used in this paper for the robustness analysis to different specifications of the prior distribution.

This estimator is then referred to as the shrinkage estimator (SBE). The same analysis can be carried out for *posterior Bayesian estimator*.

3.1 Robustness of Shrinkage Estimator under t -Prior Density

If μ is distributed as a t -distribution with $n(n > 0)$ degrees of freedom, location parameter μ_0 and scale parameter τ , then the density function is:

$$h(\mu) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \frac{1}{\sqrt{n\pi}} \frac{1}{\tau} \frac{1}{\left[1 + \frac{1}{n} \left(\frac{\mu - \mu_0}{\tau}\right)^2\right]^{\frac{n+1}{2}}}, \quad -\infty < \mu < \infty, \quad (3.2)$$

where $E(\mu) = \mu_0$ and $Var(\mu) = \frac{n}{n-2} \tau^2$ ($-\infty < \mu_0 < \infty$, $\tau > 0$). It can be shown that the larger n value or the smaller τ value, the more peaked for the distribution and as $n \rightarrow \infty$, $Var(\mu) \rightarrow \tau^2$, which is the smallest value for $Var(\mu)$. Theoretically, for sufficient larger value of n , the t -distribution is back to normal distribution.

The shrinkage estimator can be obtained by taking the derivative of (3.1) with respect to μ and letting it to be zero. The following equation is obtained:

$$2a\mu^3 + (-b - 4a\mu_0)\mu^2 + (-1 - n + 2b\mu_0 + 2an\tau^2 + 2a\mu_0^2)\mu - bn\tau^2 + n\mu_0 + \mu_0 - b\mu_0^2 = 0. \quad (3.3)$$

It can be shown that there is only one real root for this equation which is

$$\mu(n; \mu_0, \tau) = \frac{1}{6} \frac{J(n; \mu_0, \tau)^{\frac{2}{3}} - 6ag_2 + g_1^2 - g_1J(n; \mu_0, \tau)^{\frac{1}{3}}}{aJ(n; \mu_0, \tau)^{\frac{1}{3}}}, \quad (3.4)$$

where:

$$\begin{aligned} J(n; \mu_0, \tau) &= 9ag_1g_2 - 54a^2g_3 - g_1^3 + 3\sqrt{3}a(8ag_2^3 - g_1^2g_2^2 - 36ag_1g_2g_3 + 108a^2g_3^2 + 4g_1^3g_3)^{\frac{1}{2}} \\ g_1 &= -b - 4a\mu_0 \\ g_2 &= -1 - n + 2b\mu_0 + 2an\tau^2 + 2a\mu_0^2 \\ g_3 &= -bn\tau^2 + n\mu_0 + \mu_0 - b\mu_0^2 \\ a &= -\frac{1}{2} \sum_{i=1}^k t_{11i}^2 \boldsymbol{\beta}_i' \mathbf{V}^{-1} \boldsymbol{\beta}_i \\ b &= -\sum_{i=1}^k t_{11i} \boldsymbol{\beta}_i' \mathbf{V}^{-1} (t_{11i} \hat{\boldsymbol{\delta}}_i + t_{12i} \hat{\boldsymbol{\beta}}_i - t_{12i} \boldsymbol{\beta}_i). \end{aligned}$$

Since the parameters β_i and \mathbf{V} are unknown, the true value of this shrinkage estimator is unknown. However, independent estimates of β_i and \mathbf{V} are available from (2.3) and (2.4). Thus, an empirical shrinkage estimator can be formed by replacing the estimators from (2.3) and (2.4) into (3.4) as follows:

$$\hat{\mu}(n; \mu_0, \tau) = \frac{1}{6} \frac{\hat{J}(n; \mu_0, \tau)^{\frac{2}{3}} - 6\hat{a}\hat{g}_2 + \hat{g}_1^2 - \hat{g}_1\hat{J}(n; \mu_0, \tau)^{\frac{1}{3}}}{\hat{a}\hat{J}(n; \mu_0, \tau)^{\frac{1}{3}}}, \quad (3.5)$$

It can be easily proven from (3.5) that

$$\hat{\mu}(n; \mu_0, \tau) \rightarrow \frac{\hat{b}\tau^2 - \mu_0}{2\hat{a}\tau^2 - 1} = \frac{\hat{b} - \mu_0/\tau^2}{2\hat{a} - 1/\tau^2} \quad \text{as } n \rightarrow \infty. \quad (3.6)$$

which is the shrinkage Bayesian estimator (μ_{SBE}) obtained in equation (2.6). Therefore $\hat{\mu}(n; \mu_0, \tau) \rightarrow \frac{\hat{b}}{2\hat{a}} = \hat{\mu}$ (i.e. the estimator in equation (2.8)) as $n \rightarrow \infty$ and $\tau \rightarrow \infty$ which is insensitive to the different specifications of the prior density and is just a function of the data. That is to say, in this situation, this empirical shrinkage estimator from t -family is very robust to various specifications of the prior parameters. In fact, as $n \rightarrow \infty$, the t -distribution is back to normal distribution and also, as $\tau \rightarrow \infty$, the empirical shrinkage estimator $\hat{\mu}(n; \mu_0, \tau)$ approaches the noninformative shrinkage Bayesian estimator, $\hat{\mu}$, proposed in Chen et al. (1999), which is

$$\hat{\mu} = \frac{\sum_{i=1}^k t_{11i}^2 \hat{\beta}_i' \mathbf{S}_i^{-1} \hat{\delta}_i}{\sum_{i=1}^k t_{11i}^2 \hat{\beta}_i' \mathbf{S}_i^{-1} \hat{\beta}_i}. \quad (3.7)$$

This estimator $\hat{\mu}$ represents a noninformative or uniform prior and can be rewritten as $\hat{\mu} = \sum_{i=1}^k \frac{t_{11i}^2 \hat{\beta}_i' \mathbf{S}_i^{-1} \hat{\beta}_i}{\sum_{i=1}^k t_{11i}^2 \hat{\beta}_i' \mathbf{S}_i^{-1} \hat{\beta}_i} \hat{\mu}_i$, which is the weighted average of estimators of log relative potency from i th bioassay, $\hat{\mu}_i$, with the weights depending only on the maximum likelihood estimates of the model parameters. Therefore, in the situation of no prior information, this estimator ($\hat{\mu}$) is recommended.

The above outcomes strongly suggests that in the t -distribution and normal distribution, the shrinkage Bayesian estimator proposed in Chen et al. (1999) is robust for different specifications of prior distributions and approaches the noninformative shrinkage

Bayesian estimator. However, the noninformative shrinkage Bayesian estimator $\hat{\mu}$ is prior-independent and possesses the advantages that $\hat{\mu}$ is the weighted average of estimators of log relative potency from each bioassay with the weights depending only on the maximum likelihood estimates of the model parameters. It can be used to estimate the log relative potency in the bioassay theory.

3.2 Robustness of Shrinkage Estimator under Cauchy Distribution

Not only is the Cauchy prior distribution the extreme situation in the Section 3.1 when $n = 1$, but also it is special in the theory of bioassay. The well-known fact in bioassay theory is that the estimated log relative potency is the ratio of two normal variables. But the ratio of two normal variables is distributed as a Cauchy distribution. From this point of view, the Cauchy distribution is relatively more important. In this section, the Cauchy distribution is considered as the prior for the robustness analysis.

If μ is distributed by a Cauchy distribution, then the density function is $h(\mu) = \frac{1}{\pi\tau\{1+(\frac{\mu-\mu_0}{\tau})^2\}}$, where $\tau > 0$ is the scale parameter and μ_0 ($-\infty < \mu_0 < \infty$) is the location parameter. The same procedures as in Section 3.1 yields the shrinkage estimator:

$$\mu(\mu_0, \tau) = -\frac{1 - J_{\mu_0, \tau}^{\frac{2}{3}} - 12a + 12a^2\tau^2 - 4a^2\mu_0^2 + 4ab\mu_0 - b^2 - 4a\mu_0 J_{\mu_0, \tau}^{\frac{1}{3}} - b J_{\mu_0, \tau}^{\frac{1}{3}}}{6 a J_{\mu_0, \tau}^{\frac{1}{3}}}, \quad (3.8)$$

where:

$$\begin{aligned} J_{\mu_0, \tau} &= -[36a^2\mu_0 - 18ab + 72a^3\tau^2\mu_0 - 36a^2b\tau^2 + 8a^3\mu_0^3 - 12a^2b\mu_0^2 + 6ab^2\mu_0 - b^3 \\ &\quad - 6\sqrt{3}a(-16a + 4ab\mu_0 + 80a^3\tau^2\mu_0^2 + 20a\tau^2b^2 + 32a^4\tau^4\mu_0^2 + 8a^2b^2\tau^4 \\ &\quad + 16a^4\tau^2\mu_0^4 + \tau^2b^4 + 48a^2\tau^2 - 4a^2\mu_0^2 - b^2 - 48a^3\tau^4 + 16a^4\tau^6 - 80a^2b\tau^2\mu_0 \\ &\quad - 32a^3\tau^2b\mu_0^3 + 24a^2\tau^2\mu_0^2b^2 - 8a\tau^2b^3\mu_0 - 32a^3\tau^4b\mu_0)^{\frac{1}{2}}] \\ a &= -\frac{1}{2} \sum_{i=1}^k t_{11i}^2 \boldsymbol{\beta}_i' \mathbf{V}^{-1} \boldsymbol{\beta}_i \\ b &= -\sum_{i=1}^k t_{11i} \boldsymbol{\beta}_i' \mathbf{V}^{-1} (t_{11i} \hat{\boldsymbol{\delta}}_i + t_{12i} \hat{\boldsymbol{\beta}}_i - t_{12i} \boldsymbol{\beta}_i). \end{aligned}$$

Then the empirical shrinkage estimator can be obtained by replacing the estimators from (2.3) and (2.4). This results in:

$$\hat{\mu}(\mu_0, \tau) = -\frac{1 - \hat{J}_{\mu_0, \tau}^{\frac{2}{3}} - 12\hat{a} + 12\hat{a}^2\tau^2 - 4\hat{a}^2\mu_0^2 + 4\hat{a}\hat{b}\mu_0 - \hat{b}^2 - 4\hat{a}\mu_0\hat{J}_{\mu_0, \tau}^{\frac{1}{3}} - \hat{b}\hat{J}_{\mu_0, \tau}^{\frac{1}{3}}}{6\hat{a}\hat{J}_{\mu_0, \tau}^{\frac{1}{3}}}, \quad (3.9)$$

Similarly, it can be proven that

$$\hat{\mu}(\mu_0, \tau) \rightarrow \hat{\hat{\mu}} \quad \text{as} \quad \tau \rightarrow \infty. \quad (3.10)$$

In this case the empirical shrinkage estimator again approaches the noninformative shrinkage Bayesian estimator proposed in Chen et al. (1999).

3.3 The Ideal Shrinkage Estimator

The results from both Sections 3.1 and 3.2 imply that as $\mu_0 \rightarrow \hat{\hat{\mu}}$, the empirical shrinkage estimator $\hat{\mu}(n; \mu_0, \tau)$ in (3.5) and $\hat{\mu}(\mu_0, \tau)$ in (3.9) will both be robust irrespective of the prior variance τ^2 . That is to say, if the data contain sufficient information to make the prior distribution concentrate on the prior means, or the prior distribution is chosen to match the data “perfectly”, then the empirical shrinkage estimator is very robust to different specifications of prior. Otherwise, if the prior disagrees sharply with the information contained in the data, the empirical shrinkage estimator is less robust.

Therefore, to estimate the log relative potency in combination of bioassays, the plausible procedure is to use the estimators from (2.3) and (2.4) to estimate a prior mean and prior standard deviation on any $k - 1$ of k bioassays. Next utilize them as the prior information to form the final shrinkage estimator for the last bioassay, which is robust and has all the properties previously described. Hence, it is named as the ideal shrinkage estimator. This ideal shrinkage estimator is superior to the general shrinkage estimator and the noninformative shrinkage estimator because the prior distribution is formulated from the data, yielding a more accurate estimator.

4 Data Analysis

Data was originally from Finney (1978). In illustrating the combination of multivariate bioassays, Meisner, Kushner and Laska (1986) artificially broke up the data into two multivariate bioassays with the first bioassay using the twelve bivariate responses from lines 1 and 3 and the second bioassay using the twelve bivariate responses from lines 2 and 4. The summary statistics are given in Table 1 for the two assays.

Table 1: Summary Statistics of the Two Assays

Assay i	$\hat{\beta}_i$	$\hat{\delta}_i$	$\begin{pmatrix} t_{11i} & t_{12i} \\ 0 & t_{22i} \end{pmatrix}$	$m_i S_i$
1	$\begin{pmatrix} 16.625 \\ -3.125 \end{pmatrix}$	$\begin{pmatrix} -8.5 \\ 2.3333 \end{pmatrix}$	$\begin{pmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{8} \end{pmatrix}$	$\begin{pmatrix} 1475.0417 & 126.625 \\ 126.625 & 68.5417 \end{pmatrix}$
2	$\begin{pmatrix} 27.125 \\ -2.5 \end{pmatrix}$	$\begin{pmatrix} -15.8333 \\ 1.16667 \end{pmatrix}$	$\begin{pmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{8} \end{pmatrix}$	$\begin{pmatrix} 6042.7083 & 218.5 \\ 218.5 & 80.8333 \end{pmatrix}$

As seen in Chen et al (1999), the test for equality of covariance can not be rejected since the test statistic is 4.107, which is less than $\chi_{3,0.05}^2=7.815$. Consequently we can use the pooled variance-covariance as follows:

$$S = \begin{pmatrix} 417.653 & 19.174 \\ 9.174 & 8.299 \end{pmatrix}.$$

Based on these values, the noninformative shrinkage estimate for the overall log relative potency $\hat{\mu}$ in equation (2.8) can be calculated to be -0.590 with 95% CI as (-1.039, -0.141), suggesting the test preparation is significantly less potent than the standard preparation in this combination of bioassays.

To graphically illustrate the robustness analysis in Section 3, Figure 1 displays this empirical shrinkage estimator $\hat{\mu}(n; 0, \tau)$ in equation (3.5) versus n and τ for fixed $\mu_0 = 0$ based on this data. It can be seen from Figure 1 that even though the shrinkage estimator $\hat{\mu}(n; \mu_0, \tau)$ differs for different specifications of values of n , μ_0 and τ , but as n and τ are

sufficiently large enough, the empirical shrinkage estimator $\hat{\mu}(n; \mu_0, \tau)$ quickly approaches the noninformative empirical Bayesian estimator $\hat{\mu}$ in equation (2.8) and appears to be robust irrespective of the specifications of n , μ_0 and τ in t -distribution family.

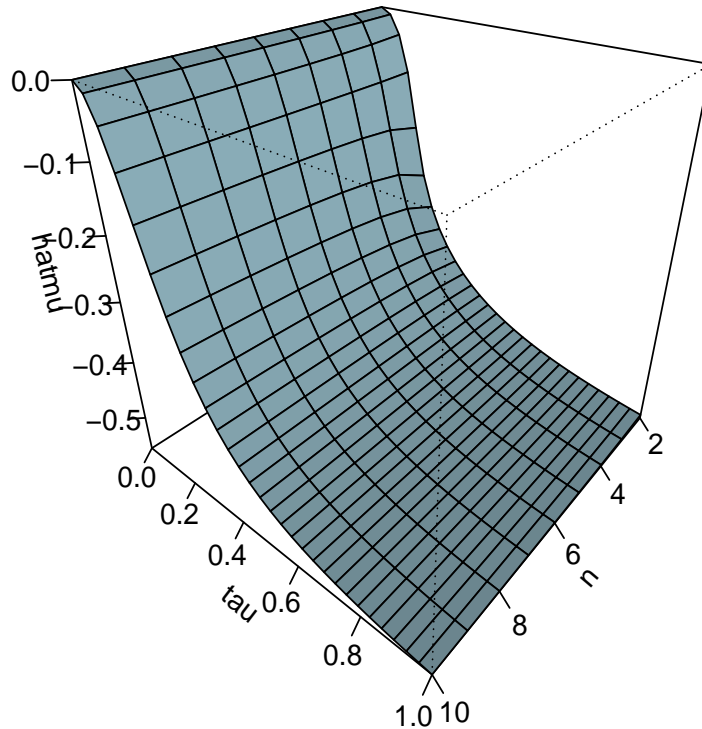


Figure 1: $\hat{\mu}(n; 0, \tau)$ as the function of n and τ with the prior of t -distribution.

Similar conclusion can be made for the empirical shrinkage estimator $\hat{\mu}(\mu_0, \tau)$ in equation (3.9) versus μ_0 and τ as shown in Figure 2. It can be seen from Figure 2 that if τ is sufficiently large, $\hat{\mu}(\mu_0, \tau)$ quickly approaches the noninformative empirical Bayesian estimator $\hat{\mu}$ in equation (2.8) and appears to be robust irrespective of the value at μ_0 . If the prior mean $\mu_0 \rightarrow \hat{\mu}$, $\hat{\mu}(\mu_0, \tau)$ is robust regardless of the values for τ .

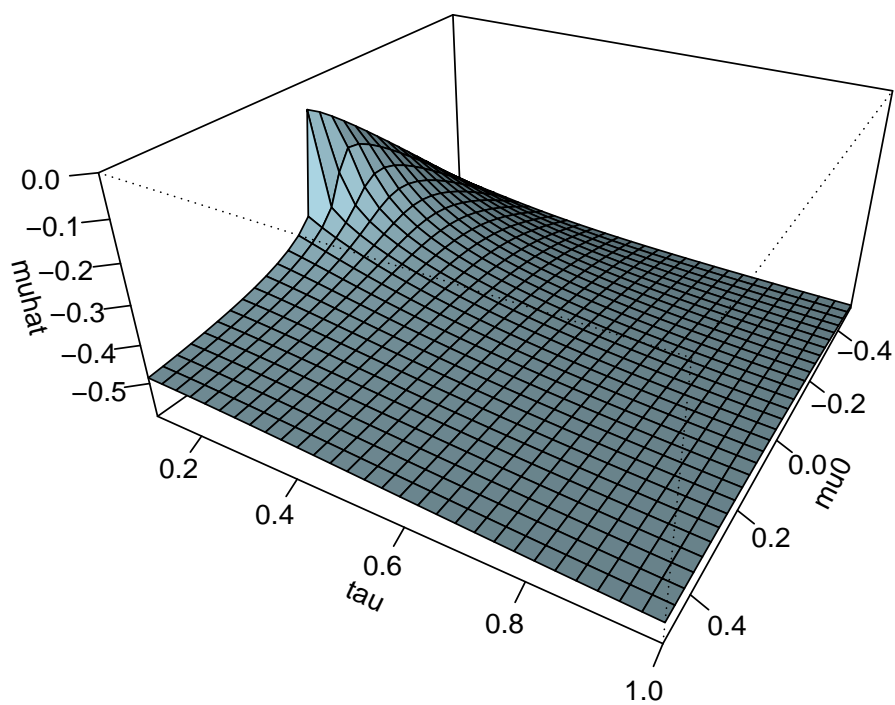


Figure 2: $\hat{\mu}(\mu_0, \tau)$ as the function of μ_0, τ with the prior of Cauchy distribution.

To numerically illustrate the robustness, Table 2 gives the values of the empirical shrinkage estimator $\hat{\mu}(n, \mu_0, \tau)$ in equation (3.5) for several specifications of n , μ_0 and τ . In the table, four values of $n = 1, 5, 10, 50$ are illustrated which are column-headed by $\hat{\mu}(1, \mu_0, \tau)$, $\hat{\mu}(5, \mu_0, \tau)$, $\hat{\mu}(10, \mu_0, \tau)$ and $\hat{\mu}(50, \mu_0, \tau)$. The first column $\hat{\mu}(1, \mu_0, \tau)$ is in fact $\hat{\mu}(\mu_0, \tau)$ in equation (3.9) since whenever $n = 1$, the t -distribution becomes the special case of Cauchy distribution and therefore $\hat{\mu}(n, \mu_0, \tau)$ in equation (3.5) would be $\hat{\mu}(\mu_0, \tau)$ in equation (3.9). The last column $\hat{\mu}(50, \mu_0, \tau)$ illustrates the empirical shrinkage estimator with large degrees of freedom from t -distribution which would be very similar to the normal prior. Therefore this last column would mimic the shrinkage Bayesian estimator μ_{SBE} in equation (2.6) as proposed in Chen et al (1999). The middle two columns illustrate the t -distribution between Cauchy and normal distributions.

In Table 2, three values of μ_0 are chosen where $\mu_0 = -0.5$ is close to the noninformative empirical estimator $\hat{\hat{\mu}} = -0.590$ and other two values of 0 and 0.5 are chosen farther away from $\hat{\hat{\mu}}$. Five values of τ are chosen from 0.1 (small), 0.5, 1, 2 to 3 (large). It can be clearly seen from Table 2 that all $\hat{\mu}(1, \mu_0, \tau)$, $\hat{\mu}(5, \mu_0, \tau)$, $\hat{\mu}(10, \mu_0, \tau)$ and $\hat{\mu}(50, \mu_0, \tau)$ rapidly approach the $\hat{\hat{\mu}} = -0.590$ irrespective of the different specifications of μ_0 and τ . This is especially true with $\mu_0 = -0.5$ which again indicates that when an appropriate prior is selected, the three shrinkage estimators of μ_{SBE} in equation (2.6), $\hat{\mu}(n, \mu_0, \tau)$ in equation (3.5) and $\hat{\mu}(\mu_0, \tau)$ in equation (3.9) could all approximate the noninformative shrinkage Bayesian estimator $\hat{\hat{\mu}}$ in equation (2.8).

This leads to the ideal shrinkage estimator proposed in Section 3.3 to use any $k - 1$ of the k bioassays as the prior information to compute a prior mean $\hat{\mu}_0$ using equation (2.8) and prior standard deviation $\hat{\tau}_0$ using equation (2.9) and then utilize these values to form the final shrinkage estimator using equations (2.6) and (2.7) with the k th bioassay. This can be easily implemented for this data. We take the first bioassay to compute a prior mean $\hat{\mu}_0$ using equation (2.8) and prior standard deviation $\hat{\tau}_0$ using equation (2.9) which are $\hat{\mu}_0 = -0.654$ and $\hat{\tau}_0 = 0.351$. We can then use these values as in the normal prior to

Table 2: Numerical illustration of the empirical Bayesian estimator $\hat{\mu}(n, \mu_0, \tau)$ for different specifications of $n = 1, 5, 10, 50$, $\mu_0 = -0.5, 0, 0.5$ and $\tau = 0.1, 0.5, 1.0, 2.0, 3.0$. The bolded values indicate $\hat{\mu}(n, \mu_0, \tau)$ approaches the noninformative shrinkage estimator $\hat{\mu} = -0.590$.

μ_0	τ	$\hat{\mu}(1, \mu_0, \tau)$	$\hat{\mu}(5, \mu_0, \tau)$	$\hat{\mu}(10, \mu_0, \tau)$	$\hat{\mu}(50, \mu_0, \tau)$
-0.5	0.1	-0.508	-0.512	-0.513	-0.514
-0.5	0.5	-0.564	-0.572	-0.573	-0.574
-0.5	1.0	-0.582	-0.585	-0.585	-0.586
-0.5	2.0	-0.588	-0.589	-0.589	-0.589
-0.5	3.0	-0.589	-0.590	-0.590	-0.590
0.0	0.1	-0.079	-0.093	-0.094	-0.094
0.0	0.5	-0.485	-0.487	-0.488	-0.488
0.0	1.0	-0.546	-0.557	-0.559	-0.561
0.0	2.0	-0.576	-0.581	-0.582	-0.583
0.0	3.0	-0.584	-0.586	-0.587	-0.587
0.5	0.1	-0.485	-0.165	0.274	0.319
0.5	0.5	-0.507	-0.451	-0.431	-0.408
0.5	1.0	-0.538	-0.537	-0.536	-0.536
0.5	2.0	-0.569	-0.574	-0.575	-0.576
0.5	3.0	-0.579	-0.583	-0.584	-0.584

calculate a final estimate using equations (2.6) and (2.7) from the second bioassay. In this situation, these two equations are actually reduced to:

$$\mu_{Ideal} = \frac{\hat{\tau}_0^2 t_{112}^2 \hat{\beta}_2' \mathbf{S}^{-1} \hat{\delta}_2 + \hat{\mu}_0}{\hat{\tau}_0^2 t_{112}^2 \hat{\beta}_2' \mathbf{S}^{-1} \hat{\beta}_2 + 1} \quad (4.11)$$

$$\tau_{Ideal} = \left(t_{112}^2 \hat{\beta}_2' \mathbf{S}^{-1} \hat{\beta}_2 + \frac{1}{\hat{\tau}_0^2} \right)^{-\frac{1}{2}}. \quad (4.12)$$

And they are estimated as $\mu_{Ideal} = -0.590$ and $\tau_{Ideal} = 0.229$ with associated 95% CI as (-1.039,-0.141). Similarly if we take the second bioassay to compute the prior mean $\hat{\mu}_0$ and prior standard deviation $\hat{\tau}_0$, we can get $\hat{\mu}_0 = -0.543$ and $\hat{\tau}_0 = 0.302$. Using these values in the normal prior to calculate the final estimate using equations (4.11) and (4.12) on the first bioassay, we can obtain $\mu_{Ideal} = -0.590$ and $\tau_{Ideal} = 0.229$ with associated 95% CI as (-1.039,-0.141) which yields the exact numeric values.

All the calculations in this section are done in R (a free software available from <http://www.r-project.org>) and the R program can be requested from the author.

5 Discussion

In this paper the investigation of the robustness of the shrinkage Bayesian estimator was conducted for two commonly-known families of t -distribution and Cauchy-distribution. It was shown that the shrinkage estimator would change for different specifications of different prior distributions. But if the prior variance is large enough or the prior information matches the “data”, the shrinkage estimator is robust and approaches the noninformative shrinkage Bayesian estimator which is insensitive to different specifications of prior distribution and is only data-dependent. A real data analysis demonstrated these conclusions. As seen from the data analysis, the three shrinkage estimators from the normal distribution (in Chen et al. 1999), the t -distribution and the Cauchy-distribution approached to the noninformative shrinkage Bayesian estimate of $\hat{\mu} = -0.590$ within wide range of prior variance τ^2 .

If the prior information is indeed available, the shrinkage Bayesian estimator proposed in Chen et al. (1999) can be readily applied to the prior information directly and hence a Bayesian estimator can be calculated using the procedures in this paper. Otherwise, the ideal shrinkage estimator proposed in Section 3.3 could be adopted for the estimation of the relative potency for combination of bioassays. This ideal shrinkage estimator was demonstrated to be valid and applicable which also produced the same results as the noninformative shrinkage Bayesian estimator from the real data analysis.

In fact, the procedure in this paper is very general and can be used for any specification of the prior density. Not only can it be used for robust analysis, it can also be used to propose different Bayesian estimators under different prior specifications. As a summary, it can be concluded that the Bayesian approach is recommended in the theory of bioassay. Furthermore, this newly-developed Bayesian procedure is applicable for the combination of univariate bioassays (Xiong and Chen 2007), combination of multivariate bioassays, parabolic bioassays (Chen 2010), as well as combination of parabolic bioassays. It has been shown that it is superior to some conventional methods. The robust investigation in this paper makes a good conclusion for its properties and provides a strong recommendation for its use. Hence, the Bayesian method is a natural way to improve the theory of bioassay.

Acknowledgements

I am grateful for the useful suggestions and comments from the two reviewers, Professor Balakrishnan and Dr. Nicole Trabold which greatly improved the draft manuscript.

References

- [1] Armitage, P. (1970). The combination of assay results. *Biometrika*, 57: 665-666.
- [2] Bennett, B.M. (1962). On combining estimates of relative potency in bioassay. *Journal of Hygiene*, 60: 379-385.

- [3] Carter, E. M. and Hubert, J. J. (1985). Analysis of parallel-line assays with Multivariate responses. *Biometrics*, 41: 703-710.
- [4] Chen, D. G., Carter, E. M., Hubert J. J. and Kim, P. T.(1999). Empirical Bayesian estimation for combinations of multivariate bioassays. *Biometrics*, 55(4): 1035-1043.
- [5] Chen, D. G. (2007). Bootstrapping Estimation for Relative Potency in the Combinations of Bioassays. *Computational Statistics and Data Analysis*, 51: 4597-4604.
- [6] Chen, D. G. (2010). Estimate the relative potency in parabolic bioassay. *Journal of Advances and Applications in Statistical Sciences*. 2(1): 1-18.
- [7] Finney, D. J. (1978). *Statistical Method in Biological Assay*. Third Edition. C. Griffin, London.
- [8] Govindarajulu, Z. (2001). *Statistical Techniques in bioassay*. Karger PubliSBErs (USA).
- [9] Hoadley, B.(1970). A Bayesian look at inverse linear regression. *Journal of the American Statistical Association*, 65: 357-369.
- [10] Laska, E. M., Kushner, H. B. and Meisber, M(1985). Multivariate bioassay. *Biometrics*, 41: 547-554.
- [11] Meisner, M., Kushner, H. B. and Laska, E. M.(1986). Combining multivariate bioassays. *Biometrics*, 42: 421-427.
- [12] Mood, A. M., Graybill, F.A. and Boes, D.C.(1974). Introduction to the Theory of Statistics. McGraw-Hill, Inc.
- [13] Rose, M. P. and Gaines-Das, R. E.(1998). Characterisation, calibration and comparison by international collaborative study of international standards for the calibration of therapeutic preparations of FSH. *Journal of Endocrinology*, 158: 97-114.

- [14] Srivastava, M. S.(1986). Multivariate bioassay, combination of bioassays and Fieller's theorem. *Biometrics*, 42: 131-141.
- [15] Srivastava, M.S. and Carter, E.M.(1983). *An Introduction to Applied Multivariate Statistics*. North-Holland, New York.
- [16] Williams, D.A.(1988). An exact confidence region for a relative potency estimated from a multivariate bioassays. *Biometrics*, 44: 861-867.
- [17] Xiong, J. and Chen, D. G. (2007). A Shrinkage Estimator for Combination of Bioassays. *Acta Mathematicae Applicatae Sinica*, 23(3): 467-476.