Bayesian and Robust Bayesian Analysis Under a General Class of Balanced Loss Functions

Mohammad Jafari Jozani,[a,1] Éric Marchand,[a,2] Ahmad Parsian,[b,3]

a Université de Sherbrooke, Département de mathématiques, Sherbrooke, QC, CANADA, J1K 2R1

b School of Mathematics, Statistics and Computer Science, University of Tehran, Tehran, IRAN

ABSTRACT

For estimating an unknown parameter θ, we introduce and motivate the use of balanced loss functions of the form $L_{\rho,\omega,\delta_0}(\theta,\delta) = \omega\rho(\delta_0,\delta)+(1-\omega)\rho(\theta,\delta)$, as well as weighted versions $q(\theta)L_{\rho,\omega,\delta_0}(\theta,\delta)$, with $q(\cdot)$ being a positive weight function, where $\rho(\theta,\delta)$ is an arbitrary loss function, δ_0 is a chosen a priori "target" estimator of θ, and the weight ω takes values in $[0,1)$. A general development with regards to Bayesian estimators under L_{ρ,ω,δ_0} is given, namely by relating such estimators to Bayesian solutions for the unbalanced case, i.e., L_{ρ,ω,δ_0} with $\omega = 0$. Illustrations are given for various choices of ρ, such as for absolute value, entropy, linex, intrinsic (i.e., model based), and a generalization of squared error losses. Finally, with regards to various robust Bayesian analysis criteria; which include posterior regret gamma-minimaxity, conditional gamma-minimaxity, and most stable, we again establish explicit connections between optimal actions derived under balanced and unbalanced losses.

Keywords and phrases: Balanced loss function; Bayes estimator; Robust Bayesian analysis; Posterior risk; Posterior regret gamma-minimax; Conditional gamma-minimax; Most stable estimation; Entropy loss; Intrinsic loss; Linex loss.

MSC: 62C05, 62C10, 62C15, 62C20, 62F10, 62F15, 62F30

[1]e-mail: mohammad.jafari.jozani@usherbrooke.ca

[2]Research supported by NSERC of Canada (corresponding author; e-mail: eric.marchand@usherbrooke.ca)

[3]Research supported by a grant of the Research Council of the University of Tehran. (e-mail: ahmad_p@khayam.ut.ac.ir)

1 Introduction

This paper concerns Bayesian point estimation of an unknown parameter θ of a model $X = (X_1, \ldots, X_n) \sim F_\theta$, under "Balanced" loss functions of the form:

$$L_{\rho,\omega,\delta_0}(\theta, \delta) = \omega\rho(\delta_0, \delta) + (1-\omega)\rho(\theta, \delta), \tag{1}$$

as well as weighted versions $q(\theta)L_{\rho,\omega,\delta_0}(\theta, \delta)$, with $q(\cdot)$ being a positive weight function (see Remark 2). Here ρ is an arbitrary loss function, while δ_0 is a chosen a priori "target" estimator of θ, obtained for instance from the criterion of maximum likelihood estimator, least-squares, or unbiasedness among others. Loss L_{ρ,ω,δ_0}, which depends on the observed value of $\delta_0(X)$, reflects a desire of closeness of δ to both: **(i)** the target estimator δ_0, and **(ii)** the unknown parameter θ; with the relative importance of these criteria governed by the choice of $\omega \in [0, 1)$.

Loss L_{ρ,ω,δ_0} may be viewed as a natural extension to Zellner's (1994) balanced loss function (in one dimension), the latter being specific to a squared error loss ρ and a least-squares δ_0. Subsequent findings, due to Dey, Ghosh and Strawderman (1999), as well as Jafari Jozani, Marchand and Parsian (2006); who also study the use of more general target estimators δ_0; involve the issues of Bayesianity, admissibility, dominance, and minimaxity and demonstrate how results for unbalanced ($\omega = 0$) squared error loss serve directly (and necessarily) in obtaining results for the balanced case with $\omega > 0$.

Here, for the more general loss in (1), we give posterior inference or Bayesian relationships between the cases $\omega = 0$ and $\omega > 0$. These connections are quite general, especially with respect to the choices of ρ and δ_0. Sections 2 and 4 are concerned with Bayesian point estimation and illustrations are provided for various choices of ρ, such as for absolute value, entropy, linex, intrinsic (i.e., model based), and a generalization of squared error losses. Section 3 is concerned with robust Bayesian analysis, and namely establishes connections with regards to optimizing posterior risk procedures, such as the criteria of posterior regret gamma-minimax, conditional gamma-minimax, and most stable estimation.

2 Bayes estimation under balanced loss

We consider here Bayes estimation under balanced loss function L_{ρ,ω,δ_0} as in (1). When $\omega = 0$, we simply use L_0 instead of $L_{\rho,0,\delta_0}$ unless we want to emphasize the role of ρ . We show how the Bayes estimator of θ under balanced loss function L_{ρ,ω,δ_0} can be derived or expressed in terms of a Bayes estimator of θ under the loss L_0. We work with ρ and δ_0 such that $\rho(\delta_0, \delta) < \infty$ for at least one $\delta \neq \delta_0$, and $\rho(\delta_0, \delta) \geq 0$ with equality if and only if $\delta = \delta_0$. We assume throughout necessary conditions for finite expected posterior loss.

Lemma 1 *For estimating θ under balanced loss function L_{ρ,ω,δ_0} in (1) and for a prior π, the* Bayes estimator $\delta_{\omega,\pi}(X)$ *corresponds to the* Bayes solution $\delta^*(X)$ *with respect to (π_x^*, L_0), for all x, where*

$$\pi_x^*(\theta) = \omega 1_{\{\delta_0(x)\}}(\theta) + (1-\omega)\pi_x(\theta), \tag{2}$$

i.e., a mixture of a point mass at $\delta_0(x)$ and the posterior $\pi_x(\theta)$.[4]

Proof. Denoting $\mu_x(\cdot)$ and $\nu_x(\cdot)$ as dominating measures of π_x and π_x^* respectively, we have by definition of $\delta_{\omega,\pi}(x)$, $\delta^*(x)$, and L_0

$$\begin{aligned}
\delta_{\omega,\pi}(x) &= argmin_\delta \int_\Theta L_{\rho,\omega,\delta_0}(\theta,\delta)\pi_x(\theta)d\mu_x(\theta) \\
&= argmin_\delta \int_\Theta \{\omega\rho(\delta_0,\delta) + (1-\omega)\rho(\theta,\delta)\}\pi_x(\theta)d\mu_x(\theta) \\
&= argmin_\delta \int_{\Theta\cup\{\delta_0(x)\}} L_0(\theta,\delta)\pi_x^*(\theta)d\nu_x(\theta) = \delta^*(x). \qquad \square
\end{aligned}$$

Remark 1 *Strictly speaking, $\delta^*(X)$ is not a Bayes estimator under loss L_0 (hence the terminology Bayes solution), because π_x^* is not the posterior associated with π, while $\delta_{\omega,\pi}(X)$ is indeed Bayes under loss L_{ρ,ω,δ_0} and prior π.*

Remark 2 *Lemma 1 also applies for weighted versions of balanced loss function (1), i.e., $L^q_{\rho,\omega,\delta_0}(\theta,\delta) = q(\theta)L_{\rho,\omega,\delta_0}(\theta,\delta)$, with $q(\theta)$ being a positive weight function. Indeed, it is easy to see that the Bayes estimator of θ under loss $L^q_{\rho,\omega,\delta_0}$ using the prior π_0 is equivalent to the Bayes estimator of θ under loss L_{ρ,ω,δ_0} using prior $\pi \equiv q \times \pi_0$.*

We pursue with various applications of Lemma 1.

Example 1 *(A generalization of squared error loss). In (1), the choice $\rho(\theta,\delta) = \rho_1(\theta,\delta) = \tau(\theta)(\delta-\theta)^2$, with $\tau(\cdot) > 0$, leads to loss $\omega\tau(\delta_0)(\delta-\delta_0)^2 + (1-\omega)\tau(\theta)(\delta-\theta)^2$. Since under L_0*

[4] We do not emphasize the dependence of $\pi_x^*(\theta)$ on δ_0 with a notation of the type $\pi^*_{x,\delta_0}(\theta)$

and a prior π, the Bayes estimator is given by $\delta_{0,\pi}(x) = \frac{E_{\pi_x}(\theta\tau(\theta))}{E_{\pi_x}(\tau(\theta))}$ (subject to the finiteness conditions $E_{\pi_x}(\theta^i\tau(\theta)) < \infty$; $i = 0, 1$; for all x), Lemma 1 tells us that

$$\delta_{\omega,\pi}(x) = \frac{E_{\pi_x^*}(\theta\tau(\theta))}{E_{\pi_x^*}(\tau(\theta))} = \frac{\omega\delta_0(x)\tau(\delta_0(x)) + (1-\omega)E_{\pi_x}(\theta\tau(\theta))}{\omega\tau(\delta_0(x)) + (1-\omega)E_{\pi_x}(\tau(\theta))}.$$

The case ($\tau(\theta) = 1$, δ_0 = least square estimator) leads to balanced squared error loss (and $\delta_{\omega,\pi}(x) = \omega\delta_0(x) + (1-\omega)E_{\pi_x}(\theta)$) as introduced by Zellner (1994), and as further analyzed by Dey, Ghosh and Strawderman (1999). The case $L_{\rho_1,\omega,\delta_0}(\theta,\delta)$ with $\tau(\theta) = 1$ and arbitrary δ_0, as well as its weighted version $q(\theta)L_{\rho_1,\omega,\delta_0}(\theta,\delta)$, were investigated by Jafari Jozani, Marchand and Parsian (2006) with respect to classical decision theory criteria such as Bayesianity, dominance, admissibility, and minimaxity. Gómez-Déniz (2008) investigates the use of such balanced loss functions to credibility premiums, while Ahmadi et al. (2008a) studies them (as well as Balanced linex loss; see Example 5) for inference on k-records.

Example 2 *(Entropy balanced loss). The choice $\rho(\theta,\delta) = \frac{\theta}{\delta} - log\frac{\theta}{\delta} - 1$ in (1) leads to loss:*

$$\omega(\frac{\delta_0}{\delta} - log(\frac{\delta_0}{\delta}) - 1) + (1-\omega)(\frac{\theta}{\delta} - log(\frac{\theta}{\delta}) - 1). \tag{3}$$

Since under L_0 (i.e., $\omega = 0$ in (3)), and a prior π, the Bayes solution is given by $\delta_{0,\pi}(x) = E_{\pi_x}(\theta)$ (subject to the finiteness conditions: $E_{\pi_x}(\theta) < \infty$ and $E_{\pi_x}(log(\theta)) < \infty$, for all x), Lemma 1 tells us that $\delta_{\omega,\pi}(x) = E_{\pi_x^*}(\theta) = \omega\delta_0(x) + (1-\omega)E_{\pi_x}(\theta)$. As a special example, take the model $X \sim Gamma(\alpha,\theta)$; α known; $\theta > 0$; (density proportional to $x^{\alpha-1}e^{-x/\theta}$); and consider estimating θ under loss (3) with arbitrary δ_0 (e.g., $\delta_0(x) = \delta_{mle}(x) = \frac{x}{\alpha}$). With conjugate prior π such that $\lambda = \theta^{-1} \sim Gamma(\gamma, \frac{1}{h})$; $\gamma > -\alpha, h \geq 0$; and with the posterior distribution of θ^{-1} being $Gamma(\alpha+\gamma, \frac{1}{h+x})$; the unique Bayes estimator of θ under entropy balanced loss (3) is:

$$\delta_{\omega,\pi}(x) = \omega\delta_0(x) + (1-\omega)\frac{h+x}{\alpha+\gamma-1}.$$

Example 3 *(Stein type loss with $\rho(\theta,\delta) = (\frac{\delta}{\theta})^\beta - \beta log\frac{\delta}{\theta} - 1$; $\beta \neq 0$). This class of losses, which includes Example 2's loss (i.e., $\beta = -1$), as well as Stein's loss (i.e., $\beta = 1$) illustrates well Lemma 1. Under loss L_0, we have the Bayes solution $\delta_{0,\pi}(x) = (E_{\pi_x}(\frac{1}{\theta^\beta}))^{-\frac{1}{\beta}}$ (subject to risk finiteness conditions). Lemma 1 tells us that $\delta_{\omega,\pi}(x) = (E_{\pi_x^*}(\frac{1}{\theta^\beta}))^{-\frac{1}{\beta}} = \{\frac{\omega}{(\delta_0(x))^\beta} + (1-\omega)E_{\pi_x}(\frac{1}{\theta^\beta})\}^{-\frac{1}{\beta}}$.*

Example 4 *(Case with $\rho(\theta,\delta) = (\frac{\theta}{\delta}-1)^2$). Under loss L_0, for a prior π supported on $(0,\infty)$, we have the Bayes solution $\delta_{0,\pi}(x) = \frac{E_{\pi_x}(\theta^2)}{E_{\pi_x}(\theta)}$ (subject to the finiteness condition $E_{\pi_x}(\theta^2) < \infty$, for all x). Lemma 1 tells us that $\delta_{\omega,\pi}(x) = \frac{E_{\pi_x^*}(\theta^2)}{E_{\pi_x^*}(\theta)} = \frac{\omega\delta_0^2(x)+(1-\omega)E_{\pi_x}(\theta^2)}{\omega\delta_0(x)+(1-\omega)E_{\pi_x}(\theta)}$. Rodrigues and Zellner (1994) introduced and studied Bayesian estimators under such a loss function for a least square δ_0 and an exponential model.*

Example 5 *(Linex loss with $\rho(\theta,\delta) = e^{a(\delta-\theta)} - a(\delta-\theta) - 1$; $a \neq 0$; (e.g., Zellner, 1986; Parsian and Kirmani, 2002)). Under loss L_0, we have the Bayes solution $\delta_{0,\pi}(x) = -\frac{1}{a} log E_{\pi_x}(e^{-a\theta})$ (subject to risk finiteness conditions). Hence, using Lemma 1, we have $\delta_{\omega,\pi}(x) = -\frac{1}{a} log E_{\pi_x^*}(e^{-a\theta}) = -\frac{1}{a} log(\omega e^{-a\delta_0(x)} + (1-\omega)E_{\pi_x}(e^{-a\theta})) = -\frac{1}{a} log(\omega e^{-a\delta_0(x)} + (1-\omega)e^{-a\delta_{0,\pi}(x)})$.*

The following development concerns intrinsic balanced loss functions, where ρ in (1) is derived in some automated manner from the model. For a general reference on intrinsic losses and additional details we refer to Robert (1996).

Example 6 *(Intrinsic balanced loss functions). For a model $X|\theta \sim f(x|\theta)$, the choice $\rho(\theta,\delta) = d(f(\cdot|\theta), f(\cdot|\delta))$ in (1); i.e., the distance between $f(\cdot|\theta)$ and $f(\cdot|\delta)$; where $d(\cdot,\cdot)$ is a suitable distance function, leads to intrinsic balanced loss functions of the form,*

$$\omega d\left(f(\cdot|\delta_0), f(\cdot|\delta)\right) + (1-\omega) d\left(f(\cdot|\theta), f(\cdot|\delta)\right). \tag{4}$$

Two candidates of interest for d in (4) are Hellinger and Kullback-Leibler distance leading to:

- *Kullback-Leibler Balanced Loss,*

$$L^{KL}_{\omega,\delta_0}(\theta,\delta) = \omega E_{\delta_0}[\log\frac{f(X|\delta_0)}{f(X|\delta)}] + (1-\omega)E_\theta[\log\frac{f(X|\theta)}{f(X|\delta)}]; \tag{5}$$

- *Hellinger Balanced Loss,*

$$L^{H}_{\omega,\delta_0}(\theta,\delta) = \omega\frac{1}{2}E_{\delta_0}[(\sqrt{\frac{f(X|\delta)}{f(X|\delta_0)}}-1)^2] + (1-\omega)\frac{1}{2}E_\theta[(\sqrt{\frac{f(X|\delta)}{f(X|\theta)}}-1)^2]; \tag{6}$$

or equivalently,

$$L^{H}_{\omega,\delta_0}(\theta,\delta) = 1 - \omega E_{\delta_0}[\sqrt{\frac{f(X|\delta)}{f(X|\delta_0)}}] - (1-\omega)E_\theta[\sqrt{\frac{f(X|\delta)}{f(X|\theta)}}]. \tag{7}$$

Hence, with known Bayesian representations in the unbalanced case ; i.e., $\omega = 0$; (e.g., Brown, 1986; Robert, 1996) and with Lemma 1, we can build a catalog of Bayesian representations for losses of the type L^{KL} and L^{H}. For instance, natural parameter exponential family of distributions with densities $f(x|\theta) = e^{\theta T(x) - \psi(\theta)} h(x)$ (with respect to a σ-finite measure ν on $\mathcal{X}$), and unknown natural parameter θ, lead to Kullback-Leibler balanced loss functions:

$$\omega\left[(\delta_0 - \delta)\psi'(\delta_0) + \psi(\delta) - \psi(\delta_0)\right] + (1-\omega)\left[(\theta - \delta)\psi'(\theta) + \psi(\delta) - \psi(\theta)\right]. \tag{8}$$

Furthermore, under L_0 (i.e., $\omega = 0$ in (8)), and for prior π, the Bayes estimator of θ is given as a solution of $\psi'(\delta_{0,\pi}(x)) = E_{\pi_x}[\psi'(\theta)]$ in $\delta_{0,\pi}(x)$. Hence, together with Lemma 1, we have that the Bayes estimator $\delta_{\omega,\pi}(x)$ admits the following representation

$$\psi'(\delta_{\omega,\pi}(x)) = \omega\psi'(\delta_0(x)) + (1-\omega)\psi'(\delta_{0,\pi}(x)),$$

with the special case of estimating $E_\theta[T(X)] = \psi'(\theta)$ (as for $X \sim N(\theta, 1)$, $T(X) = X$), yielding $\delta_{\omega,\pi}(x) = \omega\delta_0(x) + (1-\omega)\delta_{0,\pi}(x)$; i.e., a convex linear combination of δ_0 and the Bayes estimator $\delta_{0,\pi}(x)$. Also, it is easy to see that for natural exponential family of distributions as above, Hellinger balanced loss in (7) reduces to:

$$1 - \omega e^{\left\{\psi(\frac{\delta_0+\delta}{2}) - \frac{\psi(\delta_0)+\psi(\delta)}{2}\right\}} - (1-\omega) e^{\left\{\psi(\frac{\theta+\delta}{2}) - \frac{\psi(\theta)+\psi(\delta)}{2}\right\}}. \tag{9}$$

For instance, a $N(\theta, 1)$ model leads to Hellinger balanced loss in (9) given by

$$1 - \omega e^{-\frac{(\delta_0-\delta)^2}{8}} - (1-\omega) e^{-\frac{(\theta-\delta)^2}{8}},$$

and which we can refer to as a Balanced bounded reflected normal loss (e.g., Spiring, 1993, for the case $\omega = 0$).

Remark 3 *In Example 2, loss (3) is equivalent to Kullback-Leibler balanced loss (8) for a Gamma(α, θ) model with known $\alpha > 0$, and unknown scale parameter θ (see also Parsian and Nematollahi, 1996; Robert, 1996).*

Example 7 *(Balanced absolute loss or L_1 loss). The choice $\rho(\theta, \delta) = |\delta - \theta|$ in (1) leads to balanced L_1 loss:*

$$\omega|\delta - \delta_0| + (1-\omega)|\delta - \theta|. \tag{10}$$

Since Bayesian solutions associated with L_1 losses are posterior medians, Lemma 1 tells us that the Bayes estimate of θ, $\delta_{\omega,\pi}(X)$, with respect to loss (10) and priors π is a median of

$\pi_x^*(\theta)$, *for all* x. *Working now with priors* π *and models* $f(\cdot|\theta)$ *leading to unique posterior medians, setting* $\alpha(x) = (1-\omega)P_{\pi_x}(\theta \leq \delta_0(x))$ *and* $F^{-1}_{\theta|x}$ *as the inverse posterior cdf, it is easy to see that for* $\omega \geq \frac{1}{2}$, $\delta_{\omega,\pi}(x) = \delta_0(x)$, *(i.e., the point mass of* $\delta_0(x)$ *has large enough probability to force the median to be* $\delta_0(x)$*), while for* $\omega < \frac{1}{2}$ *a straightforward evaluation leads to:*

$$\delta_{\omega,\pi}(x) = F^{-1}_{\theta|x}(\frac{1-2\omega}{2(1-\omega)})1_{[0,\frac{1}{2}-\omega)}(\alpha(x)) + \delta_0(x)1_{[\frac{1}{2}-\omega,\frac{1}{2})}(\alpha(x)) + F^{-1}_{\theta|x}(\frac{1}{2(1-\omega)})1_{[\frac{1}{2},1)}(\alpha(x)), \tag{11}$$

for all x. *Here are some interesting applications of (11).*

(A) (Case of a positive normal mean). Let $X|\theta \sim N(\theta,\sigma^2)$ *with* $\theta \geq 0$, σ^2 *known, and cdf* F_0. *Decision theoretic elements concerning the estimation of such constrained* θ*'s have a long history (e.g., see Marchand and Strawderman, 2004, for a review). Here, take* $\delta_0(x) = \delta_{mle}(x) = max(0,x)$ *and* $\omega < \frac{1}{2}$ *in (10), and consider the estimator* $\delta_{\omega,\pi}(x)$ *associated with the flat prior* $\pi(\theta) = 1_{(0,\infty)}(\theta)$. *In evaluating (11), we have:* **(i)** $P_{\pi_x}(\theta \geq y) = \frac{F_0(x-y)}{F_0(x)}$ *for* $y \geq 0$; **(ii)** $F^{-1}_{\theta|x}(z) = x - F_0^{-1}((1-z)F_0(x))$, *and* **(iii)** $\alpha(x) = (1-\omega)P_{\pi_x}(\theta \leq \delta_0(x)) = (1-\omega)(1-\frac{1}{2F_0(x)})1_{(0,\infty)}(x)$. *Now, since* $\alpha(x) < \frac{1}{2}$, *and* $x \leq F_0^{-1}(1-\omega)$, *we obtain*

$$\begin{aligned} \delta_{\omega,\pi}(x) &= F^{-1}_{\theta|x}(1-\frac{1}{2(1-\omega)})1_{(-\infty,F_0^{-1}(1-\omega))}(x) + \delta_0(x)1_{[F_0^{-1}(1-\omega),\infty)}(x) \\ &= \{x - F_0^{-1}(\frac{F_0(x)}{2(1-\omega)})\}1_{(-\infty,F_0^{-1}(1-\omega))}(x) + x1_{[F_0^{-1}(1-\omega),\infty)}(x) \\ &= x - \{F_0^{-1}(\frac{F_0(x)}{2(1-\omega)})1_{(-\infty,F_0^{-1}(1-\omega))}(x)\}. \end{aligned}$$

We point out that the above expressions hold for the unbalanced case (i.e., $\omega = 0$*), as well as for location models* $X|\theta \sim f_0(x-\theta)$, *with* $\theta \geq 0$, *and known, symmetric, and absolutely continuous density* f_0.

(B) Let $X|\theta \sim N(\theta,\sigma^2)$; $\theta \in \Re$, σ^2 *known; with cdf* $\Phi(\frac{x-\theta}{\sigma})$. *Using conjugate prior* $\theta \sim \pi(\theta) = N(\mu,\tau^2)$; *known* μ *and* τ^2; *and balanced* L_1 *loss in (10) with* $\delta_0(X) = X$, $\omega < \frac{1}{2}$, *and since the posterior* $\theta|x \sim N(\frac{\tau^2}{\tau^2+\sigma^2}x + \frac{\sigma^2}{\tau^2+\sigma^2}\mu, \frac{\tau^2\sigma^2}{\tau^2+\sigma^2})$, *we have* $\alpha(x) = (1-\omega)P_{\pi_x}(\theta \leq x) = (1-\omega)\Phi(d(x-\mu))$; *with* $d = \frac{\sigma}{\tau\sqrt{\tau^2+\sigma^2}}$. *With* $F^{-1}_{\theta|x}(\alpha) = \frac{\tau^2}{\tau^2+\sigma^2}x +$

$\frac{\sigma^2}{\tau^2+\sigma^2}\mu + \sqrt{\frac{\tau^2\sigma^2}{\tau^2+\sigma^2}}\,\Phi^{-1}(\alpha)$, *an evaluation of (11) leads to the piecewise linear estimators*

$$\delta_{\omega,\pi}(x) = \begin{cases} \frac{x\tau^2+\mu\sigma^2}{\tau^2+\sigma^2} + \sqrt{\frac{\tau^2\sigma^2}{\tau^2+\sigma^2}}\Phi^{-1}(\frac{1-2\omega}{2(1-\omega)}) & \textit{if } x \leq \mu + \frac{1}{d}\Phi^{-1}(\frac{1-2\omega}{2(1-\omega)}) \\ x & \textit{if } x \in (\mu + \frac{1}{d}\Phi^{-1}(\frac{1-2\omega}{2(1-\omega)}), \mu + \frac{1}{d}\Phi^{-1}(\frac{1}{2(1-\omega)})) \\ \frac{x\tau^2+\mu\sigma^2}{\tau^2+\sigma^2} + \sqrt{\frac{\tau^2\sigma^2}{\tau^2+\sigma^2}}\Phi^{-1}(\frac{1}{2(1-\omega)}) & \textit{if } x \geq \mu + \frac{1}{d}\Phi^{-1}(\frac{1}{2(1-\omega)}). \end{cases}$$

3 Robust Bayesian analysis under Balanced loss

In Bayesian analysis, prior knowledge is often vague and its precise elicitation can be quite challenging. An interesting approach to tackle this problem is to acknowledge such prior uncertainty in considering a class Γ of plausible prior distributions and studying the corresponding range of Bayesian solutions. We refer to Berger (1994) and Insua and Ruggeri (2000), among others, for useful references on robust Bayesian analysis. One may also attempt to determine an optimal estimator δ minimizing some measure such as: **(i)** maximal posterior regret (Posterior regret Γ-minimax or PRGM; Zen and DasGupta, 1993; Insua, Ruggeri and Vidakovic, 1995; etc); **(ii)** maximal posterior risk (Conditional Γ-minimax or CGM; DasGupta and Studden, 1991; Betro and Ruggeri, 1991; Boratyńska, 1997; etc) or **(iii)** the posterior risk range (Most stable or MS; Meczarski and Zielinski, 1991; Meczarski, 1993; Boratyńska, 1997; etc), evaluated as π varies over Γ.

We now study these criteria under balanced loss function L_{ρ,ω,δ_0} as in (1), and establish general relationships (e.g., Corollary 1) between optimal actions for the cases $\omega > 0$ and $\omega = 0$. For an observed value x, a prior distribution π, corresponding posterior distribution π_x, loss L, and estimate δ, let $r(\pi_x, \delta)$ represent the posterior risk given by $E_{\pi_x}[L(\theta, \delta(x))]$. For balanced loss L_{ρ,ω,δ_0}, we denote this posterior risk as $r_{\rho,\omega,\delta_0}(\pi_x, \delta)$ for $\omega > 0$, and $r_0(\pi_x, \delta)$ for $\omega = 0$. Here are definitions of the criteria and optimal actions alluded to above, and applicable to a given class Γ of priors.

I. *Posterior Regret* Γ*-Minimax:* δ^{PRGM} is the Posterior Regret Γ-minimax estimator of θ under loss L if, for all x,

$$\sup_{\pi\in\Gamma} \psi(\pi_x, \delta^{PRGM}(x)) = \inf_{\delta} \sup_{\pi\in\Gamma} \psi(\pi_x, \delta), \tag{12}$$

where $\psi(\pi_x, \delta) = r(\pi_x, \delta) - \inf_\delta r(\pi_x, \delta)$ is the posterior regret measuring the loss entailed in choosing the action $\delta(x)$ instead of the optimal Bayes action $\delta_\pi(x)$, (for prior π, under loss L).

II. *Conditional* Γ*-minimax:* δ^{CGM} is the Conditional Γ-Minimax estimator of θ under loss L if, for all x,

$$\sup_{\pi\in\Gamma} r(\pi_x, \delta^{CGM}(x)) = \inf_{\delta} \sup_{\pi\in\Gamma} r(\pi_x, \delta)\,. \tag{13}$$

III. *Most stable:* δ^{MS} is the Most stable estimator of θ under loss L if, for all x,

$$\sup_{\pi\in\Gamma} r(\pi_x, \delta^{MS}(x)) - \inf_{\pi\in\Gamma} r(\pi_x, \delta^{MS}(x)) = \inf_{\delta} \{\sup_{\pi\in\Gamma} r(\pi_x, \delta) - \inf_{\pi\in\Gamma} r(\pi_x, \delta)\}\,. \tag{14}$$

The following result, already implicitly present within the proof of Lemma 1, describes a key posterior risk connection.

Lemma 2 *For estimating* θ *under loss* $L_{\rho,\omega,\delta_0}(\theta, \delta)$*, as in* (1)*, we have* $r_{\rho,\omega,\delta_0}(\pi_x, \delta) = r_0(\pi^*_x, \delta)$*, with* π^*_x *the mixture of a point mass at* $\delta_0(x)$ *and posterior* π_x *given in* (2)*.*

Proof. We have for all x, δ, ω, ρ, π, and by definition of π^*_x:

$$r_{\rho,\omega,\delta_0}(\pi_x, \delta) = E_{\pi_x}[L_{\rho,\omega,\delta_0}(\theta, \delta)] = E_{\pi^*_x}[L_0(\theta, \delta)] = r_0(\pi^*_x, \delta). \qquad \square$$

To pursue, we define for an observed x, and a given class Γ of priors:

$$\Gamma_x = \{\pi_x : \; \pi \in \Gamma\},$$

and

$$\Gamma_{x,\omega,\delta_0} = \{\pi^*_x : \; \pi^*_x(\theta) = \omega I_{\{\delta_0(x)\}}(\theta) + (1-\omega)\pi_x(\theta), \, \pi_x \in \Gamma_x\}\,;$$

i.e., $\Gamma_{x,\omega,\delta_0}$ represents the class of mixtures (weights ω, $1-\omega$) of a point mass at $\delta_0(x)$ with the posterior π_x. Also observe that the supremum and infimum terms (with respect to $\pi \in \Gamma$) of (12), (13), and (14) can be taken as well over $\pi_x \in \Gamma_x$. We now obtain the following as a consequence of Lemma 2.

Corollary 1 *Given a class* Γ *of priors, and its corresponding class* Γ_x *of posteriors, the (PRGM), (CGM), (MS) actions under loss* L_{ρ,ω,δ_0} *are (respectively) equivalent to the (PRGM), (CGM), (MS) actions under loss* L_0 *for the corresponding class* $\Gamma_{x,\omega,\delta_0}$*;*

Proof. With $\psi_{\rho,\omega,\delta_0}(\pi_x, \delta) = r_{\rho,\omega,\delta_0}(\pi_x, \delta) - \inf_\delta r_{\rho,\omega,\delta_0}(\pi_x, \delta)$, we obtain by Lemma 2

$$\psi_{\rho,\omega,\delta_0}(\pi_x, \delta) = r_0(\pi^*_x, \delta) - \inf_{\delta} r_0(\pi^*_x, \delta) = \psi_0(\pi^*_x, \delta).$$

Hence, for all x, it is the case that $\sup_{\pi\in\Gamma}\{\psi_{\rho,\omega,\delta_0}(\pi_x,\delta)\} = \sup_{\pi\in\Gamma}\{\psi_0(\pi_x^*,\delta)\}$, or equivalently

$$\sup_{\pi_x\in\Gamma_x}\psi_{\rho,\omega,\delta_0}(\pi_x,\delta) = \sup_{\pi_x^*\in\Gamma_{x,\omega,\delta_0}}\psi_0(\pi_x^*,\delta)\,,$$

which implies the result. The Conditional Γ-minimax and most stable equivalencies are established along the same lines. $\square$

For the PRGM solution, say $\delta_{\omega,\delta_0}^{PRGM}$, we obtain a further simplification in the case of a squared error ρ.

Corollary 2 *Given a class Γ of priors, with $\rho(y) = y^2$ in (1), and supposing that a PRGM solution δ_0^{PRGM} exists under L_0, then*

$$\delta_{\omega,\delta_0}^{PRGM}(x) = \omega\delta_0(x) + (1-\omega)\delta_0^{PRGM}(x)\,.$$

Proof. We make use of the well known PRGM solution under squared-error loss given by $\frac{1}{2}\{\inf_{\pi_x\in\Gamma_x}E_{\pi_x}(\theta) + \sup_{\pi_x\in\Gamma_x}E_{\pi_x}(\theta)\}$, and part (a) of Corollary 1, which together tell us that the PRGM action under loss L_{ρ,ω,δ_0} with $\rho(y) = y^2$ is given by

$$\delta_{\omega,\delta_0}^{PRGM}(x) = \frac{1}{2}\{\inf_{\pi_x^*\in\Gamma_{x,\omega,\delta_0}}E_{\pi_x^*}(\theta) + \sup_{\pi_x^*\in\Gamma_{x,\omega,\delta_0}}E_{\pi_x^*}(\theta)\}\,.$$

The result follows by observing that $E_{\pi_x^*}(\theta) = \omega\delta_0(x) + (1-\omega)E_{\pi_x}(\theta)$ (as the Bayesian solution in Example 1).

Example 8 *For a normal model ($X \sim N(\theta;\sigma^2)$) with $|\theta| \le m$ and known σ^2, Insua, Ruggeri and Vidakovic (1995) obtain, under squared-error loss, the posterior regret Γ-minimax solutions :*

(a) $\delta^{PRGM}(x) = \frac{m}{2}\tanh(\frac{mx}{\sigma^2})$, *for the class Γ of symmetric priors on $[-m,m]$;*

(b) $\delta^{PRGM}(x) = \frac{1}{2}\delta_U(x)$, *for the class Γ of unimodal, symmetric, absolutely continuous priors on $[-m,m]$, where δ_U is the Bayes estimator with respect to a uniform prior on $[-m,m]$ given by $\delta_u(x) = x + \sigma\frac{\phi(\frac{x+m}{\sigma})-\phi(\frac{x-m}{\sigma})}{\Phi(\frac{x+m}{\sigma})-\Phi(\frac{x-m}{\sigma})}$; with ϕ and Φ representing the pdf and cdf of a standard normal.*

For balanced squared error loss, the corresponding solutions are thus available directly from Corollary 2. For instance, if the target estimator is chosen as the maximum likelihood estimator of θ, *i.e.,* $\delta_0(x) = \delta_{mle}(x) = (m \wedge x)sgn(x)$, *we obtain the optimal PRGM action* $\delta^{PRGM}_{\omega,\delta_0}(x) = \omega\,(m \wedge x)\,sgn(x) + (1-\omega)\,\frac{m}{2}\,\tanh(\frac{mx}{\sigma^2})$ *for the class of symmetric priors on* $[-m, m]$.

4 Concluding Remarks

Our findings here do provide general connections for Bayesian point estimation (section 2) and posterior risk inference (section 3). These are achieved through Lemma 1 and Lemma 2's representations of the Bayes estimator $\delta_{\omega,\pi}$ and risk $r_{\rho,\omega,\delta_0}(\pi_x, \delta)$ associated with prior π, loss L_{ρ,ω,δ_0}, target estimator δ_0, as the Bayes solution δ^* and risk $r_0(\pi^*_x, \delta)$ associated with π^*_x and loss L_0. Although not pursued here as far as illustrations, we do point out that Lemma 1 holds in a multivariate setting (see, Dey, Ghosh and Strawderman, 1999, for balanced squared error loss). As well, similar developments are available for prediction problems (see Ahmadi et al., 2008b)

With regards to the key issue of admissibility one can infer, of course, the admissibility of such $\delta_{\omega,\pi}$'s when they are unique, the prior is proper and the Bayes risk is finite (see Berger, 1985). However, more general frequentist connections between losses L_{ρ,ω,δ_0} and L_0 with regards to dominance, admissibility and minimaxity, such as those established by Jafari Jozani, Marchand and Parsian (2006), remain to be investigated. However, when the estimator of interest and the target estimator coincide, we have the following.

Theorem 1 *For any* ρ *and* ω *in (1),*
a) if δ_0 *is Bayes under loss* L_0 *and prior* π, *then* δ_0 *is Bayes under loss* L_{ρ,ω,δ_0};
b) if δ_0 *is admissible under loss* L_0, *then* δ_0 *is admissible under loss* L_{ρ,ω,δ_0};
c) If δ_0 *is minimax under loss* L_0 *with minimax risk* M, *then* δ_0 *is minimax under loss* L_{ρ,ω,δ_0}, ρ, *with minimax risk* $(1-\omega)M$;

Proof. We prove part (c) only, with (a) and (b) following analogously by rewriting definitions in the context of loss L_{ρ,ω,δ_0}. By assumption, since $\delta_0(X)$ is minimax under loss L_0, we have:

$$\sup_\theta E_\theta[\rho(\theta, \delta(X))] \geq \sup_\theta E_\theta[\rho(\theta, \delta_0(X))] = M, \ \textit{for all } \delta(X),$$

which implies, given the assumption $\rho \geq 0$ and $\rho(\delta_0, \delta_0) = 0$:

$$\begin{aligned}\sup_\theta E_\theta[L_{\rho,\omega,\delta_0}(\theta, \delta(X))] &= \sup_\theta \{\omega E_\theta[\rho(\delta_0(X), \delta(X))] + (1-\omega)E_\theta[\rho(\theta, \delta(X))]\} \\ &\geq \sup_\theta \{\omega E_\theta[\rho(\delta_0(X), \delta_0(X))] + (1-\omega)E_\theta[\rho(\theta, \delta_0(X))]\} \\ &= (1-\omega)\sup_\theta E_\theta[\rho(\theta, \delta_0(X))] = (1-\omega)M.\end{aligned}$$

Finally, we do wish to point out (following a suggestion by Alexandre Leblanc) that most of the general results in Jafari Jozani, Marchand and Parsian (2006) (i.e, Lemma 1, 2; Corollary 1,2; Theorem 1) hold under the following modification of squared error loss: $\omega q(\delta_0)(\delta - \delta_0)^2 + (1-\omega)q(\delta_0)(\delta - \theta)^2$.

Acknowledgements

Nous remercions Alexandre Leblanc pour des suggestions et remarques forts utiles.

References

Ahmadi, J., Jafari Jozani, M., Marchand, É., and Parsian, A. (2008a). *Bayes estimation based on k-record data from a general class of distributions under balanced type loss functions.* To appear in Journal of Statistical Planning and Inference.

Ahmadi, J., Jafari Jozani, M., Marchand, É., and Parsian, A. (2008b). *Prediction of k-records from a general class of distributions under balanced type loss functions.* To appear in Metrika.

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

Berger, J.O. (1994). *An overview of robust Bayesian analysis.* Test, **3**, 5-124.

Betro, B., and Ruggeri, F. (1992). *Conditional* Γ*-minimax actions under convex losses.* Communications in Statistics: Theory and Methods, **21**, 1051-1066.

Boratyńska, A. (1997). *Stability of Bayesian inference in exponential families.* Statistics & Probability Letters, **36**, 173-178.

Brown, L.D. (1986). *Foundations of Exponential Families.* IMS Lecture Notes, Monograph Series **9**, Hayward, California.

DasGupta, A., and Studden, W. (1991). *Robust Bayesian experimental designs in normal linear models.* Annals of Statistics, **19**, 1244-1256.

Dey, D., Ghosh, M., and Strawderman, W.E. (1999). *On estimation with balanced loss functions.* Statistics & Probability Letters, **45**, 97-101.

Gómez-Déniz, E. (2008). *A generalization of the credibility theory obtained by using the weighted balanced loss function.* Insurance: Mathematics and Economics, **42**, 850-854.

Insua, R.D., and Ruggeri, F. (2000). *Bayesian sensitivity analysis.* Springer-Verlag, New York.

Inusa, R.D., Ruggeri, F., and Vidakovic, B. (1995). *Some results on posterior regret* Γ*-minimax estimation.* Statistics & Decisions, **13**, 315-351.

Jafari Jozani, M., Marchand, É., and Parsian, A. (2006). *On estimation with weighted balanced-type loss function.* Statistics & Probability Letters, **76**, 773-780.

Marchand, É., and Strawderman, W.E. (2004). *Estimation in restricted parameter spaces: A review.* A Festschrift for Herman Rubin, IMS Lecture Notes-Monograph Series 45, Institute of Mathematical Statistics, Hayward, California, 21-44.

Meczarski, M. (1993). *Stability and conditional* Γ*-minimaxity in Bayesian inference.* Applicationes Mathematicae, **22**, 117-122.

Meczarski, M., and Zielinski, R. (1991). *Stability of Bayesian estimator of the Poisson mean under the inexactly specified Gamma prior.* Statistics & Probability letters, **12**, 329-333.

Parsian, A., and Kirmani, S.N.U.A. (2002). *Estimation under LINEX loss function.* Handbook of applied econometrics and statistical inference, Statistics Textbooks Monographs, 53-76.

Parsian, A., and Nematollahi, N. (1996). *Estimation of scale parameter under entropy loss function.* Journal of Statistical Planning and Inference, **52**, 77-91.

Robert, C.P. (1996). *Intrinsic loss functions.* Theory and Decision, **40**, 192-214.

Rodrigues, J., and Zellner, A. (1994). *Weighted balanced loss function and estimation of the mean time to failure.* Communications in Statistics: Theory and Methods, **23**, 3609-3616.

Spiring, F.A. (1993). *The reflected normal loss function.* Canadian Journal of Statistics, **21**, 321-330.

Zellner, A. (1986). *Bayesian estimation and prediction using asymmetric loss functions.* Journal of the American Statistical Association, **81**, 446-451.

Zellner, A. (1994). *Bayesian and Non-Bayesian estimation using balanced loss functions.* Statistical Decision Theory and Methods *V*, (J.O. Berger and S.S. Gupta Eds). New York: Springer-Verlag, 337-390.

Zen, M., and DasGupta, A. (1993). *Estimating a binomial parameter: Is robust Bayes real Bayes?* Statistics & Decisions, **11**, 37-60.