IMS Collections
Contemporary Developments in Bayesian Analysis and Statistical Decision Theory: A Festschrift for William E. Strawderman
Vol. 8 (2012) 104–116

DOI: 10.1214/11-IMSCOLL808

Inadmissible estimators of normal quantiles and two-sample problems with additional information

Éric Marchand[1], **Mohammad Jafari Jozani**[*2] **and Yogesh Mani Tripathi** [3]

Université de Sherbrooke, University of Manitoba, Indian Institute of Technology

Abstract: We consider estimation problem of a normal quantile $\mu + \eta\sigma$. For the scale invariant squared error loss and unrestricted values of the population mean and standard deviation μ and σ, [13] established the inadmissibility of the MRE estimator for $\eta \neq 0$. In this paper, we explore: (i) the impact of the loss with the study of scale invariant absolute value loss, and (ii) situations where there is a parameter space restriction of a lower bounded mean μ. We establish

(i) the inadmissibility of the MRE estimator of $\mu + \eta\sigma$; $\eta \neq 0$; under scale invariant absolute value loss;

(ii) the inadmissibility of the Generalized Bayes estimator of $\mu + \eta\sigma$; $\eta > 0$; under scale invariant squared error loss, associated with the prior measure $1_{(0,\infty)}(\mu)1_{(0,\infty)}(\sigma)$ which represents the truncation of the usual non-informative prior measure onto the restricted parameter space.

Both of these results are obtained through a conditional risk analysis and may be viewed as extensions of [13]. Finally, we provide further applications to two-sample problems under the presence of the additional information of ordered means.

Contents

[*]Corresponding author. Tel: +1 204 272 1563; fax:+1 204 474 7621.
[1]Université de Sherbrooke, Département de mathématiques, Sherbrooke, QC, Canada, J1K 2R1, e-mail: eric.marchand@usherbrooke.ca
[2]University of Manitoba, Departement of Statistics, Winnipeg, MB, Canada, R3T 2N2, e-mail: m_jafari_jozani@umanitoba.ca
[3]Department of Mathematics, Indian Institute of Technology, Patna, Patna, India 800013, e-mail: yogesh@iitp.ac.in
AMS 2000 subject classifications: 62F10, 62F30, 62C15
Keywords and phrases: absolute value loss, additional information, complete class, conditional risk, estimation, inadmissibility, normal quantiles, restricted parameter space, squared error loss

1. Introduction

Bill Strawderman has contributed in deep and original ways to statistical decision theory, Bayesian analysis and their applications. He has shared with enthusiasm his sharp intuition and vast knowledge in statistics with many of the contributors of this volume, us included. The statistical community has benefited greatly from his contributions through his own research, his influence on others and his untiring devotion and service to the cause of the profession for over 40 years. We are pleased to join in this celebration in honor of Professor Strawderman.

Many practical situations and fields of study such as reliability, life testing, mortality data, insurance, economics, and education, require efficient statistical methods for drawing inference upon percentiles or quantiles. In this regard, [1] survey various interval estimation methods, while many papers have been dedicated to practical aspects of such problems, as described for instance in [4]. Consider a random sample from a normal population with unknown mean μ and standard deviation σ, and the problem of estimating a quantile $\mu + \eta\sigma$, with η known. A very interesting decision-theoretic result due to Zidek [13], which was expanded upon by Rukhin [11], is the inadmissibility of the benchmark minimum risk equivariant (MRE) and minimax estimator δ_0 of $\mu + \eta\sigma$, for $\eta \neq 0$, under scale and location invariant squared error loss $(\frac{\delta-\mu-\eta\sigma}{\sigma})^2$. This finding also represents a particular instance of an inadmissible Generalized Bayes estimator ($\eta \neq 0$), since δ_0 is indeed Bayes with respect to the non-informative prior measure

$$\pi(\mu,\sigma) = \frac{1}{\sigma} 1_{(-\infty,\infty)}(\mu)\, 1_{(0,\infty)}(\sigma)\,. \tag{1.1}$$

These findings contrasts with the case $\eta = 0$, that is the admissibility of the MRE estimator (i.e., the sample mean) for estimating the mean of a normal population, as well as the admissibility of the Bayes estimator for a known variance with respect the flat prior $1_{(0,\infty)}(\mu)$ ([3]). Observe as well that the admissibility of the MRE estimator for $\eta = 0$ holds for many symmetric losses, such as absolute value loss.

It is of intrinsic interest to revisit Zidek's result in [13], with the thought of **(i)** assessing the impact of the loss function and thus potentially gaining a more general understanding if similar results are found to be true, and **(ii)** investigating whether the inadmissibility result persists for other generalized Bayes estimators, such as for priors which incorporate parametric restrictions on (μ,σ). We will focus in **(i)** on scale invariant absolute value loss, which possesses attractive features of its own, and which has been recently considered for MRE estimators under various models in [5]. For **(ii)**, we investigate the case of a lower bounded mean. Key findings of this paper include:

- **(a)** the inadmissibility for $\eta \neq 0$ of the MRE (or Bayes with respect to the prior in (1.1)) estimator under scaled absolute value L_1 error loss $|\frac{\delta-\mu-\eta\sigma}{\sigma}|$;
- **(b)** the inadmissibility for $\eta > 0$ under scale invariant squared error L_2 loss of the generalized Bayes estimator δ_{π_0} with respect to the prior,

$$\pi_0(\mu,\sigma) = \frac{1}{\sigma} 1_{(0,\infty)}(\mu)\, 1_{(0,\infty)}(\sigma)\,, \tag{1.2}$$

which represents the truncation of the prior π in (1.1) onto the restricted parameter space. We expand further on the estimation context relative to **(b)** at the beginning of Section 4.

As in [12] or [13], the inadmissibility results stem actually from more general complete class results, and are based on the study of scale invariant estimators and a risk analysis conditional on the maximal invariant (Section 2). Various other technical results and arguments with L_1 loss and analytical properties of the Bayes estimator δ_{π_0}, are required (e.g., Lemma 2). The treatment is unified in **(a)** with [13]'s result for L_2 loss (Section 3), while our findings in **(b)** are cast amongst a scarcity of work concerning the estimation of quantiles under a parametric restriction (Section 4). Finally in Section 5, we expand upon further implications for two-sample problems with additional information on the means, by making use of a variant of the so-called "rotation" technique introduced in the late 60's by Blumenthal, Cohen and Sackrowitz.

2. Preliminaries and conditional risk analysis

Our results are derived for the canonical form:

$$X \sim N(\mu, \sigma^2),\ S^2 \sim \text{Gamma}(\frac{n-1}{2}, 2\sigma^2),\ \ n \geq 2, \tag{2.1}$$

(X, S^2) independent. The objective is to estimate the quantile $\mu + \eta\sigma$ ($\eta \neq 0$) with scale invariant loss

$$\rho\,(\frac{\delta - \mu - \eta\sigma}{\sigma})\,, \tag{2.2}$$

ρ being nonnegative, absolutely continuous, convex, $\rho(0) = 0$.

Remark 1. Results derived for the canonical form in (2.1) apply for independent observable $X_1, \ldots, X_n \sim N(\theta, \sigma^2)$, with sample mean $\bar{X}$, and for estimating a quantile $\theta + \beta\sigma$ by δ' under the loss $\rho^*(\frac{\delta' - \theta - \beta\sigma}{\sigma})$. This is achieved by setting $X = \sqrt{n}\,\bar{X}, S^2 = \sum_{i=1}^n (X_i - \bar{X})^2, \mu = \sqrt{n}\,\theta, \eta = \sqrt{n}\,\beta, \delta = \sqrt{n}\,\delta'$, and $\rho(z) = \rho^*(z/\sqrt{n})$.

Remark 2. If the parameter space is unrestricted (i.e., $\mu \in \Re, \sigma > 0$), we may assume without loss of generality that $\eta > 0$. Indeed in the model:

$$X' \sim N(\mu', \sigma^2),\ S^2 \sim \text{Gamma}(\frac{n-1}{2}, 2\sigma^2),$$

(X', S^2) independent, the loss $\rho_0(\frac{\delta'(X',S) - \mu' - \eta\sigma}{\sigma})$ with $\eta < 0$ matches the loss $\rho(\frac{\delta(X,S) - \mu - \eta\sigma}{\sigma})$ for the canonical form in (2.1) with $X = -X', \mu = -\mu', \eta = -\eta', \delta = -\delta'$, and $\rho(z) = \rho_0(-z)$ for all $z \in \Re$. However, for positive η and positivity constraint $\mu \geq 0$, the equivalent problem for negative η is as above but with a negativity constraint $\mu' \leq 0$.

Following [13], our conditional risk analysis proceeds as follows. In (2.1), set $Y = \frac{X}{S}, V = \frac{S}{\sigma}$, and $\lambda = \frac{\mu}{\sigma}$. Next, consider the class of scale invariant estimators of the form $\delta_\psi(X, S) = S\psi(Y)$, to which the MRE estimator belongs as $\delta_{\text{mre}}(X, S) = S(Y + c_{\rho,n}\eta)$ (see (3.2) for a representation of $c_{\rho,n}$). Then, decompose the risk of δ_ψ,

$$R((\mu, \sigma), \delta_\psi) = E[\rho(\frac{S\,\psi(Y)\ -\ \mu\ -\ \eta\sigma}{\sigma})] = E[r(\lambda, \psi(Y))], \tag{2.3}$$

with

$$r(\lambda, \psi(y)) = E[\rho(V\,\psi(y) \;-\; \lambda \;-\; \eta)\,|Y = y], \tag{2.4}$$

being the conditional risk of δ_ψ (given $Y = y$) and depending on the parameters (μ, σ) only through their ratio λ. Hence, dominance or complete class results are available by comparisons of the conditional risks $r(\lambda, \psi(y))$ only.

Lemma 1. **(a)** *For fixed λ, there exists an optimal choice $\psi^*_\lambda(y)$ which minimizes in $\psi(y)$ the conditional risk in (2.4), and it is found from*

$$E[\,\rho'(W\,\psi^*_\lambda(y) - \lambda - \eta)\,] \;=0\,, \tag{2.5}$$

with the distribution of W depending on λ and y, with density proportional to

$$w^n\, e^{-(\frac{1+y^2}{2})\,(w-\frac{\lambda y}{1+y^2})^2}\, 1_{(0,\infty)}(w)\,. \tag{2.6}$$

(b) *For fixed λ, δ_{ψ_1} dominates δ_{ψ_2} ($\delta_{\psi_1} \neq \delta_{\psi_2}$) if for all $y \in \Re$: $\delta_{\psi_2}(y) \geq \delta_{\psi_1}(y) \geq \delta_{\psi^*_\lambda}(y)$ or $\delta_{\psi_2}(y) \leq \delta_{\psi_1}(y) \leq \delta_{\psi^*_\lambda}(y)$, with strict inequality between $\psi_1(y)$ and $\psi_2(y)$ on a set of positive Lebesgue measure.*

Proof. **(a)** First, observe that $r(\lambda, \psi(y))$ is convex in $\psi(y)$ since ρ is convex. Therefore,

$$E[\,V\rho'(V\,\psi^*_\lambda(y) - \lambda - \eta)\,|Y = y] \;=0\,. \tag{2.7}$$

The conditional density $f_{V|Y=y}$ is obtained as in (2.6) with n replaced by $n-1$. Finally, defining W as a random variable with density proportional to $w f_{V|Y=y}(w)$ in (2.7), leads to (2.5) and (2.6).
(b) This is a consequence of part (a) and convexity of $r(\lambda, \psi(y))$ as a function of $\psi(y)$. ☐

As a function of λ, $\psi^*_\lambda(y)$ may be bounded, and this leads to complete class results as implied by part **(b)** of the previous lemma. Indeed, in the cases under study here, $\psi^*_\lambda(y)$ will be shown to be upper bounded (e.g., Theorems 1 and 2), and the inadmissibility results will concern estimators δ_ψ with large ψ.

Corollary 1. *Suppose there exists an upper envelope $\bar\psi(y)$ on a region D for $\psi^*_\lambda(y)$ such, that for all $y \in D$, $\bar\psi(y) \geq \psi^*_\lambda(y)$ for all $(\mu, \sigma) \in \Theta$. Let δ_ψ be a scale invariant estimator and let C be a subset of D such that the Lebesgue measure of C is positive, where $C = \{y \in D : \psi(y) > \bar\psi(y)\}$. Then δ_ψ is inadmissible for estimating $\mu + \eta\sigma$, being dominated by $\delta_{\psi'}$ with $\psi'(y) = \bar\psi(y)\, I_C(y) + \psi(y)\, I_{C'}(y)$, $y \in \Re$.*

Proof. Follows from part **(b)** of Lemma 1 with $\psi_2 \equiv \psi$ and $\psi_1 \equiv \psi'$, since we have for $y \in C:\;\; \delta_\psi(y) > \delta_{\psi'}(y) = \delta_{\bar\psi}(y) \geq \delta_{\psi^*_\lambda}(y)$ for all $(\mu, \sigma) \in \Theta$. ☐

Remark 3. In applying Corollary 1 with continuous ψ and $\bar\psi$, it suffices to determine a singleton y_0 such that $\psi(y_0) > \bar\psi(y_0)$, in which case C can be taken as a small neighborhood of y_0.

We next collect some useful properties of the distribution of W, as defined in (2.6).

Lemma 2. *Let $a = \frac{\lambda y}{\sqrt{1+y^2}}$ and $c(\lambda, y, n) = \frac{1}{\sqrt{1+y^2}}\,(\frac{a}{2} + \sqrt{\frac{a^2}{4} + n})$. Then*

(a) $E(W)$ is a strictly increasing function of n;

(b) $c(\lambda, y, n) < E(W) < c(\lambda, y, n+1)$;
(c) *Median*$(W) \geq c(\lambda, y, n)$.

Proof. See Appendix. □

Remark 4. It is easy to verify that the density of W is unimodal with the mode at $c(\lambda, y, n)$. Consequently, the above inequalities may also be interpreted as mean-mode-median inequalities. The inequality $E(W) > c(\lambda, y, n)$ is due to [13].

3. Inadmissibility of the MRE estimator under absolute value invariant error loss

We now turn to a useful representation of the MRE estimator under location-scale changes for the scale invariant absolute value error loss.

Lemma 3. **(a)** *For the model (2.1), the MRE estimator of* $\mu + \eta\sigma$ $(\eta \neq 0)$ *under the loss* $|\frac{d-\mu-\eta\sigma}{\sigma}|$ *is given by* $\delta_{mre}(X, S) = X + c_{1,n}\eta S$, *where* $c_{1,n}$ *is uniquely defined by the equation*

$$E_{0,1}[S\{1 - 2\Phi(\eta(1 - c_{1,n}S)\}] = 0\,, \tag{3.1}$$

and Φ *is the standard normal cdf.*
(b) *For all* $\eta > 0$, $n \geq 2$, $c_{1,n} > \frac{1}{\sqrt{n}}$.

Proof. **(a)** It is well understood that equivariant estimators here are of the form $X + c\eta S$ and have constant risk for losses as in (2.2) (e.g., [11]). It follows that the optimal choice of c minimizes $E_{0,1}[\rho(X + c\eta S - \eta)]$ in c, and is uniquely given by

$$E_{0,1}[S\rho'(X + c\eta S - \eta)] = 0\,, \tag{3.2}$$

when ρ is convex. For $\rho(z) = |z|$ in (2.2), (3.1) becomes $E_{0,1}[SE_{0,1}[\,\mathrm{sgn}(X + c\eta S - \eta)|S\,]\,] = 0$, which yields (3.1) given the independence of X and S.
(b) Since $\Phi(\cdot)$ is an increasing function and $\eta > 0$, we have from (3.1), that

$$\begin{aligned} c_{1,n} > \frac{1}{\sqrt{n}} \quad &\Leftrightarrow \quad E_{0,1}[\frac{S}{\sqrt{n}}\{1 - 2\Phi(\eta(1 - \frac{S}{\sqrt{n}})\}] < 0 \\ &\Leftrightarrow \quad \int_0^\infty u^{n-1}\,\{1 - 2\Phi(\eta(1-u))\}\,e^{-\frac{nu^2}{2}}\,du \,<\, 0 \\ &\Leftrightarrow \quad h(\eta) > 1/2 \text{ for all } \eta > 0. \end{aligned}$$

Here $h(\eta) = E[\Phi(\eta(1-U))]$, and U has density proportional to $u^{n-1}e^{-\frac{nu^2}{2}}\,1_{(0,\infty)}(u)$. We next show that $h(\cdot)$ is strictly increasing on $(0, \infty)$, which will suffice since $h(0) = 1/2$. A direct computation yields

$$h^{'}(\eta) \propto \int_0^\infty (1-u)\,e^{-\frac{\eta^2(1-u)^2}{2}}\,u^{n-1}\;e^{-\frac{nu^2}{2}}\,du,$$

so that $h^{'}(\eta) > 0$ if and only if

$$\frac{\int_0^\infty u^n \;\; e^{\frac{-(n+\eta^2)}{2}\,(u-\frac{\eta^2}{n+\eta^2})^2}\,du}{\int_0^\infty u^{n-1}\,e^{\frac{-(n+\eta^2)}{2}\,(u-\frac{\eta^2}{n+\eta^2})^2}\,du} < 1\,. \tag{3.3}$$

Finally, a re-parametrization of (2.6) and part (b) of Lemma 2 show that the ratio in (3.3) is bounded above by $c(\frac{\eta^2}{\sqrt{n+\eta^2-1}}, \sqrt{n+\eta^2-1}, n) = 1$, which establishes the result. □

Theorem 1. *In model (2.1) to estimate the quantile $\mu + \eta\sigma$ ($\eta > 0$) under the loss $|\frac{d-\mu-\eta\sigma}{\sigma}|$, the estimator δ_{mre} is inadmissible. It is dominated by $\delta_{\psi'}(X,S) = \psi'(\frac{X}{S})S$ with $\psi'(y) = \{y + \min(\frac{1}{y} + \frac{\eta^2 y}{4n}, c_{1,n}\eta)\}1_{(0,\infty)}(y) + \{y + c_{1,n}\eta\}\, 1_{(-\infty,0]}(y)$.*

Proof. We apply Corollary 1 with $\psi(y) = y + c_{1,n}\,\eta$, $D = (0,\infty)$, and $\bar{\psi}(y) = (y + \frac{1}{y} + \frac{\eta^2 y}{4n})\, 1_{(0,\infty)}(y)$, in which case $C = \{\, y > 0 : \psi(y) > \bar{\psi}(y)\} = \{\, y : |y - \frac{2nc_{1,n}}{\eta}| < \frac{2n}{\eta}\sqrt{c_{1,n}^2 - \frac{1}{n}}\,\} \neq \emptyset$ by part (b) of Lemma 3. It remains to establish that $\sup_{(\mu,\sigma)\in\Theta}\{\,\psi_\lambda^*(y)\,\} \leq \bar{\psi}(y)$ for all $y > 0$. From Lemmas 1 and 2, for all $y > 0$:

$$\psi_\lambda^*(y) = \frac{\lambda+\eta}{\text{Median}(W)} \leq \frac{\lambda+\eta}{c(\lambda,y,n)}\,. \tag{3.4}$$

Now write $\frac{1}{c(\lambda,y,n)} = \frac{y}{\sqrt{1+y^2}}(\frac{y+\frac{1}{y}}{n})\,(\sqrt{\frac{a^2}{4}+n} - \frac{a}{2})$, with $a = \frac{\lambda y}{\sqrt{1+y^2}}$ (as in Lemma 2). Setting $b = \frac{\eta y}{\sqrt{1+y^2}}$ and $a_0(b) = \frac{1}{2b}(4n - b^2)$, we obtain (following [13]) for all $y > 0$

$$\begin{aligned}
\sup_{(\mu,\sigma)\in\Theta} \psi_\lambda^*(y) &\leq \sup_{\lambda\in\Re} \frac{\lambda+\eta}{c(\lambda,y,n)} \\
&= \sup_{\lambda\geq-\eta} \frac{\lambda+\eta}{c(\lambda,y,n)} \\
&= \frac{1}{n}\,(y+\frac{1}{y}) \sup_{a\geq -b}\{(a+b)\,(\sqrt{\frac{a^2}{4}+n} - \frac{a}{2})\} \\
&= \frac{1}{n}\,(y+\frac{1}{y})\{(a_0(b)+b)\,(\sqrt{\frac{a_0^2(b)}{4}+n} - \frac{a_0(b)}{2})\} \\
&= \bar{\psi}(y)\,.
\end{aligned}$$

$\square$

Much of our motivation for Theorem 1 rests not only with the technical challenges and required intermediate results, but also with the common features with [13]'s result for the squared error invariant loss. We point out that the MRE estimator under the squared error invariant loss $(\frac{d-\mu-\eta\sigma}{\sigma})^2$, which may be obtained from (3.2), is given by

$$X + c_{2,n}\eta S, \quad c_{2,n} = \frac{E_{0,1}(S)}{E_{0,1}(S^2)} = \frac{\Gamma(\frac{n}{2})}{\sqrt{2}\,\Gamma(\frac{n+1}{2})}. \tag{3.5}$$

Furthermore, one can verify that $c_{2,n}c_{2,n+1} = \frac{1}{n}$, that $c_{2,n}$ decreases in n, so that $\frac{1}{\sqrt{n}} < c_{2,n} < \frac{1}{\sqrt{n-1}}$ for all $n \geq 2$. Consequently, the given lower bounds for $c_{1,n}$ and $c_{2,n}$ coincide even though $c_{1,n}$ depends on η and $c_{2,n}$ does not. As established by [13], Theorem 1 holds if we replace the L_1 loss by the L_2 loss and $c_{1,n}$ by $c_{2,n}$. Hence, our proof for L_1 loss parallels that of the L_2 loss.

4. Estimating a quantile in presence of a lower bounded mean

4.1. Inadmissibility of a generalized Bayes estimator

Consider our quantile estimation problem as defined in (2.1) and (2.2), with a lower bound constraint on μ, without loss of generality $\mu \geq 0$. We establish the inadmissibility under the scale invariant squared error loss of the generalized Bayes estimator

δ_{π_0} of $\mu+\eta\sigma$ with $\eta > 0$, where the prior measure is $\pi_0(\mu,\sigma) = \frac{1}{\sigma} 1_{(0,\infty)}(\mu)\, 1_{(0,\infty)}(\sigma)$. As the truncation of the non-informative prior $\pi(\mu,\sigma) = \frac{1}{\sigma} 1_{(-\infty,\infty)}(\mu)\, 1_{(0,\infty)}(\sigma)$, this choice is interesting. On one hand, the Bayes estimator δ_π coincides with δ_{mre} and is thus inadmissible under both losses L_1 and L_2 following the results of Section 3 and those of [13]. On the other hand, truncations such as π_0 have been studied before. For instance, [3] considered estimating a non negative normal mean μ with a known variance under the loss $(d-\mu)^2$, for which the Bayes estimator is both minimax and admissible. Further more general minimax results for location models and location invariant losses were obtained by [2] and [9], among others. Bayesian HPD credible intervals based on such truncations of non-informative priors have satisfactory frequentist coverage properties (e.g., [8]).

The next lemma, whose proof is relegated to the Appendix, pertains to the Bayes estimator δ_{π_0}.

Lemma 4. *Let $\beta_\rho(w,z) = E[V \int_{-\infty}^{Vw} \rho'(u-\eta+V(z-w))\phi(u)du]$ for $z,w \in \Re$, where ϕ stands for the standard normal pdf and $V = S/\sigma$ in (2.1).*

(a) For model (2.1) and loss (2.2), the Bayes estimator δ_{π_0} is scale invariant, $\delta_{\pi_0}(X,S) = S\psi_{\pi_0}(\frac{X}{S})$, with $\beta_\rho(y,\psi_{\pi_0}(y)) = 0$ for all $y \in \Re$;

(b) For the scale invariant squared error loss, δ_{π_0} may be expressed as in part (a) with

$$\psi_{\pi_0}(y) = y + \frac{A_n(y) + \eta c_{2,n+1} B_n(y)}{B_{n+1}(y)}, \tag{4.1}$$

where $A_n(y) = \frac{1}{n}(1+y^2)^{-\frac{n}{2}}$, $B_n(y) = \int_{-\infty}^{y}(1+x^2)^{-(\frac{n+1}{2})}dx$; $y \in \Re, n \geq 2$; and $c_{2,n+1} = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{2}\Gamma(\frac{n+2}{2})}$.

(c) Furthermore, $2\, c_{2,n+1} B_n(y) \geq (c_{2,n} + c_{2,n+1})\, B_{n+1}(y)$ for all $n \geq 2, y > 0$.

We now establish the inadmissibility of the Bayes estimator δ_{π_0} by making use of the above properties and Corollary 1.

Theorem 2. *To estimate the quantile $\mu+\eta\sigma$ $(\eta > 0)$ under the constraint $\mu \geq 0$ and the loss $\frac{(d-\mu-\eta\sigma)^2}{\sigma^2}$ in model (2.1), the Bayes estimator δ_{π_0} from Lemma 4 is inadmissible and dominated by $\delta_{\psi''}(X,S) = \psi''(\frac{X}{S})S$, with $\psi''(y) = \psi_{\pi_0}(y)\, I_{(-\infty,0]}(y) + \min(\psi_{\pi_0}(y), \bar{\psi}(y))\, I_{(0,\infty)}(y)$, with $\bar{\psi}(y) = y + \frac{1}{y} + \frac{\eta^2 y}{4n}; y > 0$.*

Proof. Since $\sup_{\mu\geq 0,\sigma>0} \psi^*_\lambda(y) \leq \bar{\psi}(y)$ for all $y > 0$ (i.e., [13]), we can follow Corollary 1 and Remark 3, where it suffices to show that the set $\{y > 0 : \psi_{\pi_0}(y) > \bar{\psi}(y)\}$ has positive Lebesgue measure, or that there exists a positive y_0 such that $\psi_{\pi_0}(y_0) > \bar{\psi}(y_0)$. By putting $y_0 = \frac{2nc_{2,n}}{\eta}$, $\bar{\psi}(y_0) = y_0 + y_0^{-1} + \eta^2 y_0/4n = y_0 + \frac{\eta}{2}(c_{2,n} + c_{2,n+1})$, since $nc_{2,n}c_{2,n+1} = 1$. Hence, we obtain from (4.1),

$$\begin{aligned}
\psi_{\pi_0}(y_0) - \bar{\psi}(y_0) &= \left[y_0 + \frac{A_n(y_0) + \eta c_{2,n+1} B_n(y_0)}{B_{n+1}(y_0)}\right] - \left[y_0 + \frac{\eta}{2}(c_{2,n} + c_{2,n+1})\right] \\
&= \frac{1}{B_{n+1}(y_0)}\Big[A_n(y_0) + \eta c_{2,n+1} B_n(y_0) \\
&\qquad -\frac{\eta}{2}(c_{2,n} + c_{2,n+1})\, B_{n+1}(y_0)\Big] \\
&> 0\,,
\end{aligned}$$

by virtue of part (c) of Lemma 4 and since $A_n(\cdot)$ and $B_{n+1}(\cdot)$ are positive valued functions. $\square$

Remark 5. The inadmissibility of the estimators δ_{π_0} (Theorem 2) and δ_{mre} ([13]), are obtained as an application of Corollary 1 by showing the estimators expand "too much". Since δ_{π_0} is Bayes with respect to π_0, one might anticipate that δ_{π_0} expands further on δ_{mre}, which would provide an easy route to establishing Theorem 2. However, this is not necessarily the case. Indeed, for the scale invariant squared error loss, it follows from (4.1) and (3.5) that $sgn(\psi_{\pi_0}(y) - \psi_{\text{mre}}(y)) = sgn(h(y))$, with $h(y) = A_n(y) + \eta c_{2,n+1} B_n(y) - \eta c_{2,n} B_{n+1}(y)$, and

$$A_n'(y) = -yB_{n+1}'(y) = -\frac{y}{\sqrt{1+y^2}} B_n'(y).$$

We infer that

$$h'(y) = \sqrt{1+y^2}\, B_{n+1}'(y)\{\eta c_{2,n+1} - \frac{\eta c_{2,n}}{\sqrt{1+y^2}} - \frac{y}{\sqrt{1+y^2}}\}.$$

Finally, for $\eta > \frac{1}{c_{2,n+1}}$, we see that $h'(y)$ is positive for large enough y. Indeed $B_{n+1}'(\cdot)$ is positive and $\lim_{y\to\infty}\{\eta c_{2,n+1} - \frac{\eta c_{2,n}}{\sqrt{1+y^2}} - \frac{y}{\sqrt{1+y^2}}\} = \eta c_{2,n+1} - 1 > 0$, which implies that $\psi_{\pi_0}(y) < \psi_{\text{mre}}(y)$ for large y given that $\lim_{y\to\infty} h(y) = 0$. However, it can be verified that $\psi_{\pi_0}(y) \geq \psi_{\text{mre}}(y)$ for all $y \in \Re$ whenever $\eta \leq \frac{1}{c_{2,n+1}}$.

4.2. Further remarks and numerical evaluations

When $\mu \geq 0$, a minimum risk equivariant estimator $\delta_{\text{mre}}(X,S) = X + \eta c_\rho S$ of $\mu + \eta\sigma$ is clearly inefficient. It is improved upon by truncating onto $[0,\infty)$ for any loss (2.2) since $\mu + \eta\sigma \geq 0$, and $P_{\mu,\sigma}(\delta_{\text{mre}}(X,S) < 0) > 0$ for all $\mu \geq 0$, $\sigma > 0$. However, as shown recently by [7], δ_{mre} remains minimax even if $\mu \geq 0$, as its constant risk equals the minimax risk for general ρ in (2.2) subject to risk finiteness. Therefore such minimum risk equivariant estimators remain useful benchmarks, and determination of dominating estimators, which remains to be studied, must produce minimax estimators. Another motivation for the search of efficient estimators in the presence of a lower bound on the mean resides in applications for two sample problems presented in the next section.

We do not know if δ_{π_0} is a minimax estimator under scale invariant squared error loss ρ, despite being inadmissible itself for $\eta > 0$, except for $\eta = 0$ where [6] obtained a class of dominating (minimax) estimators which includes δ_{π_0}. The plausible conjecture of minimaxity is supported by some numerical evidence, part of which is illustrated in Figure 1. Figure 1 represents the risk functions of the δ_{mre}, the generalized Bayes estimator δ_{π_0} and Lemma 2's estimator $\delta_{\psi''}$ for $n = 10, \eta = 1$, as a function of $\lambda = \frac{\mu}{\sigma} \geq 0$. With other choices of $n, \eta > 0$, leading to similar results, one sees the important gains in risk provided by δ_{π_0} in comparison to δ_{mre}, and the minuscule gains in risk for $\delta_{\psi''}$ as opposed to δ_{π_0}.

5. Estimation of quantiles in the presence of additional information on the means

We describe here a correspondence between: **(i)** a two-sample problem with additional information present on the ordering of the means, and **(ii)** the quantile estimation settings of Section 4 with a lower bounded mean. We show that any dominance or admissibility result under the squared error loss in **(ii)** translates into a companion result in **(i)** and vice-versa.

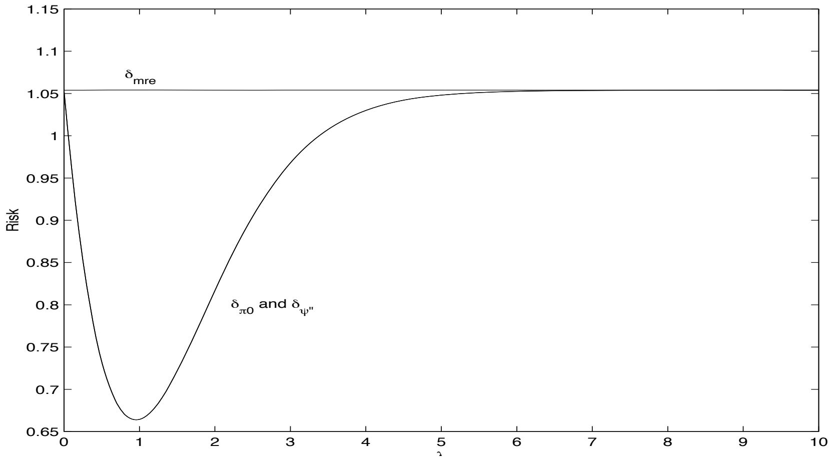

FIG 1. *Risks as functions of* $\lambda = \mu/\sigma$, $n = 10, \eta = 1.0$.

Start with a canonical form as in (2.1) for independent X_1 and S_1^2:

$$X_1 \sim N(\mu_1, \sigma^2),\ S_1^2 \sim \text{ Gamma}(\frac{n-1}{2}, 2\sigma^2),\ \ n \geq 2,$$

with the objective of estimating $\mu_1 + \eta\sigma$, $\eta > 0$, under the loss $(\frac{d-\mu_1-\eta\sigma}{\sigma})^2$, and where we already know that δ_{mre} is an inadmissible estimator. Suppose now that a second, independently generated, sample is available with independent

$$X_2 \sim N(\mu_2, \sigma^2),\ S_2^2 \sim \text{ Gamma}(\frac{m-1}{2}, 2\sigma^2),\ \ m \geq 2,$$

and suppose further that the means μ_1 and μ_2 are ordered in such a way that

$$\mu_1 \geq \mu_2 \quad \text{(additional information)}. \tag{5.1}$$

Clearly, given the homogeneity of the variances, more degrees of freedom are available and $X_1 + \eta c_{2,m+n}\sqrt{S_1^2 + S_2^2}$ seems preferable to $X_1 + c_{2,n}S_1$. Indeed, the former dominates the latter as the risk of δ_{mre} in model (2.1), given by $1 + \eta^2(1 - \frac{(\Gamma(\frac{n}{2}))^2}{\Gamma(\frac{n-1}{2})\,\Gamma(\frac{n+1}{2})})$, decreases in n.

One can also use a version of the so-called rotation technique, introduced by Blumenthal, Cohen, Sackrowitz in the late 60's, revisited by van Eeden and Zidek in a series of more recent papers, and further described by [10] (see references there). The key feature of the technique is a decomposition of the problem in **(i)** into two separate and additive subproblems, one of which corresponds to the problem in **(ii)**. To pursue, set

$$Y_1 = \frac{X_1 - X_2}{2},\ Y_2 = \frac{X_1 + X_2}{2},\ W = \frac{S_1^2 + S_2^2}{2},$$

$$\theta_1 = \frac{\mu_1 - \mu_2}{2},\ \theta_2 = \frac{\mu_1 + \mu_2}{2},\ \text{ and }\ \tau = \frac{\sigma}{\sqrt{2}}.$$

Observe that Y_1, Y_2, and W are independent with $Y_1 \sim N(\theta_1, \tau^2)$, $Y_2 \sim N(\theta_2, \tau^2)$, and $W \sim \text{Gamma}(\frac{n+m-2}{2}, 2\tau^2)$.

Lemma 5. *Consider the above data* (Y_1, Y_2, W) *for estimating the quantile* $\mu_1 + \eta\sigma$ *when* $\mu_1 \geq \mu_2$*, under the loss* $(\frac{d-\mu_1-\eta\sigma}{\sigma})^2$*. The risk of an estimator of the form* $\delta_\phi(Y_1, Y_2, W) = Y_2 + \phi(Y_1, W)$ *is given by*

$$R((\mu_1, \mu_2, \sigma), \delta_\phi) = \frac{1}{\sigma^2}\{\tau^2 + E[(\phi(Y_1, W) - \theta_1 - \sqrt{2}\,\eta\,\tau)^2]\}.$$

Also $\delta_{\phi_1}(Y_1, Y_2, W)$ *dominates* $\delta_{\phi_2}(Y_1, Y_2, W)$ *if and only if* $\phi_1(Y_1, W)$ *dominates* $\phi_2(Y_1, W)$ *as an estimator of the quantile* $\theta_1 + \eta^*\tau$*, with* $\eta^* = \sqrt{2}\,\eta$ *under the constraint* $\theta_1 \geq 0$*, based on the data* (Y_1, W) *as in (2.1).*

Proof. The dominance result is a direct consequence of the representation of the risk of δ_ϕ. To obtain $R((\mu_1, \mu_2, \sigma), \delta_\phi)$, note that

$$\begin{aligned} \sigma^2 R((\mu_1, \mu_2, \sigma), \delta_\phi) &= E[(Y_2 + \phi(Y_1, W) - \mu_1 - \eta\sigma)^2] \\ &= E[\{(Y_2 - \theta_2)^2 + (\phi(Y_1, W) - \theta_1 - \eta\sqrt{2}\,\tau)\}^2] \\ &= \tau^2 + E[(\phi(Y_1, W) - \theta_1 - \eta\sqrt{2}\,\tau)^2], \end{aligned}$$

given the independence of Y_2 and (Y_1, W), and since $E(Y_2 - \theta_2) = 0$ and $E[Y_2 - \theta_2)^2] = \tau^2$. □

Example 1. The MRE estimator of $\mu_1 + \eta\sigma$ is given by $X_1 + \eta c_{2,m+n}\sqrt{2W} = \delta_{\phi\mathrm{mre}}(Y_1, Y_2, W)$ with $\phi_{\mathrm{mre}}(Y_1, W) = Y_1 + c_{2,m+n}\eta^*\sqrt{W}$. Any dominating estimator $\phi_1(Y, W)$ of $\phi_{\mathrm{mre}}(Y_1, W)$ for estimating the quantile $\theta_1 + \eta^*\tau$, such as its truncation $\max(0, \phi_{\mathrm{mre}}(Y_1, W))$ for $\eta \geq 0$, leads to a corresponding dominating estimator $\delta_{\phi_1}(Y_1, Y_2, W)$, such as $Y_2 + \max(0, \phi_{\mathrm{mre}}(Y_1, W))$ for $\eta \geq 0$. Similarly, the estimator $Y_2 + \delta_{\pi_0}(Y_1, W)$, where δ_{π_0} is the Bayes estimator of the quantile of order η^*, is inadmissible for $\eta > 0$ and can be improved upon by making use of Lemma 5 and dominating estimators of δ_{π_0}.

Example 2. The above decomposition also applies for the case of the median (or mean) with $\eta = 0$. In this case X_1 is the MRE estimator of μ_1 (under the scale invariant squared error loss), and is admissible in absence of the second sample. With the additional information $\mu_1 \geq \mu_2$, X_1 is inadmissible. By virtue of [6]'s finding the class of dominating estimators includes $Y_2 + \delta_{\pi_0}(Y_1, W)$, where δ_{π_0} is given in (4.1).

6. Appendix

Proof of Lemma 2

(a) The result is immediate since the family of densities in (2.6), with parameter n, possesses monotone likelihood ratio in W.

(b) Set $Z = \sqrt{1 + y^2}\,W$, so that Z has density proportional to $z^n e^{-\frac{z^2}{2} + az}\, 1_{(0,\infty)}(z)$. Now write $E_n(Z) = I_{n+1}(a)/I_n(a)$, with $I_n(a) = \int_0^\infty z^n e^{-\frac{z^2}{2} + az} dz$, and integrate by parts to obtain the recurrence,

$$I_n(a) = \frac{1}{n+1} I_{n+2}(a) - \frac{a}{n+1} I_{n+1}(a),$$

or, for all $n \geq 0$

$$E_{n+1}(Z) = \frac{n+1}{E_n(Z)} + a\,. \tag{6.1}$$

Applying twice the result in (a) yields the inequalities: $E_n^2(Z) - aE_n(Z) - (n+1) < 0$, and $E_{n+1}^2(Z) - aE_{n+1}(Z) - (n+1) > 0$. Finally, since $E_n(Z) > a/2$,

$$\frac{a}{2} + \sqrt{\frac{a^2}{4} + n} < E_n(Z) < \frac{a}{2} + \sqrt{\frac{a^2}{4} + (n+1)}\,,$$

which is equivalent to, and establishes part (b).

(c) As in Remark 4, the density f_Z of Z has the mode at $M = \frac{a}{2} + \sqrt{\frac{a^2}{4} + n}$. We seek to establish that Median$(Z) \geq M$, which holds if for all $z \in [0, M]$

$$r(z) = \frac{f_Z(M-z)}{f_Z(M+z)} \leq 1. \tag{6.2}$$

Setting $T(z) = \log r(z) = n \log \frac{M-z}{M+z} + 2(M-a)z$, it is easy to verify that $T(\cdot)$ is concave on $(0, m)$ with $T(0) = 0$, and $T'(0^+) = \frac{-2n}{M} + 2(M-a) = 0$. Thus $T(z) \leq 0$ for all $z \in [0, m)$, which is equivalent to (6.2). □

Proof of Lemma 4

Proof. **(a)** Under the prior π_0 and the loss $\rho(\frac{d-\mu-\eta\sigma}{\sigma})$, $\delta_{\pi_0}(x, s)$ minimizes in δ for all (x, s) the posterior expected loss,

$$E[\rho(\frac{\delta-\mu-\eta\sigma}{\sigma})|(X,S)=(x,s)] = \int_0^\infty \int_0^\infty \rho(\frac{\delta-\mu-\eta\sigma}{\sigma})\frac{1}{\sigma}\phi(\frac{x-\mu}{\sigma})\frac{1}{\sigma}h(\frac{s}{\sigma})\frac{d\mu d\sigma}{\sigma}.$$

Here h is the density of V. With the change of variables $(\mu, \sigma) \to (u = \frac{x-\mu}{\sigma}, v = \frac{s}{\sigma})$, the above becomes proportional to

$$\int_0^\infty \int_{-\infty}^{v\frac{x}{s}} \rho(v(\frac{\delta - x}{s}) + u - \eta)\, \phi(u)\, h(v)\, du\, dv\,. \tag{6.3}$$

Observe that $\frac{1}{s}(\delta(x,s) - x)$ depends on (x, s) only through $y = \frac{x}{s}$, so that δ_{π_0} is indeed scale invariant. The result follows by differentiation of (6.3) in δ and convexity of ρ.

(b) Solving $\beta_\rho(y, \psi_{\pi_0}(y)) = 0$ for $\rho(y) = y^2$ yields

$$\psi_{\pi_0}(y) = y + \frac{E[V\phi(Vy)] + \eta E[V\Phi(Vy)]}{E[V^2\Phi(Vy)]},$$

since $\int_{-\infty}^t u\phi(u)du = -\phi(t)$ for all $t \in \Re$. The result follows by making use of identities for the terms $E[V\phi(Vy)]$ and $E[V^k\Phi(Vy)]$, given and proven below in Lemma 6, as well as the definitions of $A_n(\cdot)$, $B_n(\cdot)$, and $c_{2,n+1}$.

(c) First, we have $B_n(y) = B_n(0) + \int_0^y (1+x^2)^{-(\frac{n+1}{2})}\, dx \geq B_n(0) + \int_0^y (1+x^2)^{-(\frac{n+2}{2})}\, dx$, with $B_n(0) = \int_{-\infty}^0 (1+x^2)^{-(\frac{n+1}{2})}\, dx = \sqrt{\frac{\pi}{2}}\, c_{2,n}$. From this, we obtain

$$B_n(y) - B_{n+1}(y) \geq B_n(0) - B_{n+1}(0) = \sqrt{\frac{\pi}{2}}\,(c_{2,n} - c_{2,n+1})\,. \tag{6.4}$$

As well, notice that $B_{n+1}(y)$ increases in y for $y > 0$, and

$$B_{n+1}(y) \leq \int_{-\infty}^\infty (1+x^2)^{-(\frac{n+2}{2})}\, dx \;=\; 2\sqrt{\frac{\pi}{2}}\, c_{2,n+1}. \tag{6.5}$$

Finally, from (6.4) and (6.5), we obtain for all $n \geq 2, y > 0$: $2\,c_{2,n+1}(B_n(y) - B_{n+1}(y)) \geq (c_{2,n} - c_{2,n+1})\, B_{n+1}(y)$ yielding the result. □

Lemma 6. *Let ϕ and Φ represent the pdf and cdf (resp.) of a standard normal distribution and let $V^2 \sim Gamma(\frac{n-1}{2}, 2)$; $n \geq 2$. For all $k \geq 0$, $t \in \Re$,*

(a) $E[V^k \phi(Vt)] = \frac{2^{\frac{k-1}{2}}\,\Gamma(\frac{n+k-1}{2})}{\sqrt{\pi}\,\Gamma(\frac{n-1}{2})}\,(1+t^2)^{-(\frac{n+k-1}{2})}$;

(b) $E[V^k \Phi(Vt)] = \frac{\Gamma(\frac{n+k}{2})\,2^{\frac{k}{2}}}{\sqrt{\pi}\,\Gamma(\frac{n-1}{2})}\int_{-\infty}^{t}(1+x^2)^{-(\frac{n+k}{2})}\,dx,$ *for all* $k \geq 0$.

Proof. The part (a) follows directly with the identity

$$\int_0^\infty v^\alpha\, e^{-(v^2/2\beta)}\, dv \;=\; \Gamma(\frac{\alpha+1}{2})\, 2^{\frac{\alpha-1}{2}}\, \beta^{\frac{\alpha+1}{2}}\; ;\, \alpha \geq 0, \beta > 0.$$

For the part (b), observe that $\frac{\partial}{\partial t}E[V^k\Phi(Vt)] = E[V^{k+1}\phi(Vt)]$, which implies that $E[V^k\Phi(Vt)] = \int_{-\infty}^{t} E[V^{k+1}\phi(Vx)]\,dx + c$. Since $E(V^k\Phi(Vt)) \to 0$ as $t \to -\infty$, we obtain $c = 0$ and the stated result. □

Acknowledgments

The research programs of Éric Marchand and Mohammad Jafari Jozani are partially supported by NSERC of Canada. A substantial part of this work occurred while Jafari Jozani and Tripathi were postdoctoral fellows at the Université de Sherbrooke, and we gratefully acknowledge the additional financial support of François Perron (for Jafari Jozani; NSERC grant) and the Université de Sherbrooke. The authors are grateful to Bill Strawderman for stimulating discussions especially in Section 5. Finally, thanks to Claude Beauchamp for numerical evaluations and to Andrew Rukhin for a diligent reading of an earlier version of this manuscript and constructive comments and suggestions.

References

[1] Chakraborti, S.S. and Li, J.J. (2007). Confidence interval estimation of a normal percentile. *Amer. Statist.* **61** 331–336.

[2] Farrell, R.H. (1964). Estimators of a location parameter in the absolutely continuous case. *Ann. Math. Statist.* **35** 949–998.

[3] Katz, M. (1961). Admissible and minimax estimates of parameters in truncated spaces. *Ann. Math. Statist.* **32** 136–142.

[4] Keating, J.P. and Tripathi, R.C. (1985). Percentiles, estimation of, *Encyclopedia of statistical sciences*, vol. VI, Wiley, pp. 668–674.

[5] Keating, J.P., Mason, R.L., and Balakrishnan, N. (2010). Percentile estimators in location-scale parameter families under absolute loss. *Metrika* **72** 351–367.

[6] Kubokawa, T. (2004). Minimaxity in estimation of restricted parameters. *J. Japan Statist. Soc.* **34** 1-19.

[7] Marchand, É., and Strawderman, W. E. (2010). A unified minimax result for restricted parameter spaces. *Bernoulli*, To appear. *DOI:10.3150/10-BEJ336*.

[8] Marchand, É., and Strawderman, W.E. (2006). On the behaviour of Bayesian credible intervals for some restricted parameter space problems. In *Recent Developments in Nonparametric Inference and Probability*. Institute of Mathematical Statistics Lecture Notes—Monograph Series, **50** 112–126. Inst. Math. Statist., Beachwood, OH.

[9] Marchand, É., and Strawderman, W.E. (2005). On improving on the minimum risk equivariant estimator of a location parameter which is constrained to an interval or a half-interval. *Ann. Inst. Statist. Math.* **57** 129–143.

[10] Marchand, É., and Strawderman, W.E. (2004). Estimation in restricted parameter spaces: A review. In *A Festschrift for Herman Rubin*. Institute of Mathematical Statistics Lecture Notes—Monograph Series **45** 21–44. Inst. Math. Statist., Beachwood, OH.

[11] Rukhin, A. (1983). A class of minimax estimators of a normal quantile. *Statist. Probab. Lett.* **1** 217–221.

[12] Stein, C. (1964). Inadmissibility of the usual estimator for the variance of a normal distribution with unknown mean. *Ann. Inst. Statist. Math.* **16** 155–160.

[13] Zidek, J. (1971). Inadmissibility of a class of estimators of a normal quantile. *Ann. Math. Statist.* **42** 1444–1447.