Robustness of a Truncated Estimator for the Smaller of Two Ordered Means

Yasuyuki Hamura* and Tatsuya Kubokawa†

August 10, 2022

Abstract

In this note, we consider the problem of estimating the smaller of two ordered means. Such problems frequently arise in applications where, for example, aggregated data are observed. In order to combine information from direct and indirect observations, we use the Stein-type truncated estimator. We show that it dominates the direct estimator for distributions with log-concave or log-convex densities.

Key words and phrases: conditional expectation, dominance, log-concave densities, order restriction.

1 Introduction

Suppose that observations X and Y are independently distributed with means μ_1 and μ_2 and suppose that it is known that $\mu_1 \leq \mu_2$. We consider the problem of estimating μ_1 using not only X but also Y together with the information that $\mu_1 \leq \mu_2$. Thus, μ_1 is the parameter of interest and μ_2 is a nuisance parameter, and X is the direct estimator of μ_1 and Y contains additional indirect information about μ_1. Such situations frequently arise in applications where we observe aggregated data. For example, suppose that X and $Y_1, \dots, Y_n$ are independent Poisson variables with unknown means μ_1 and $\mu_1, \dots, \mu_n$, respectively, and that we observe X and $Y = \sum_{i=1,\dots,n} Y_i$. Then we know that $E[X] \leq E[Y]$ and we want to use Y as well as X when we are interested in μ_1.

In order to combine information from X and Y to estimate μ_1, we consider using $\widehat{\mu}_1(X, Y) = \min\{X, (X + Y)/2\}$, because it is expected that $\mu_1 \leq (X + Y)/2$ since $E[(X + Y)/2] = (\mu_1 + \mu_2)/2 \geq \mu_1$. Stein (1964) first considered a shrinkage estimator of this form for the estimation of a normal variance in the presence of an unknown mean. Such estimators have been proved to be useful in the context of estimation under order restriction. For example, Lee (1981), Lee (1988), and Kelly (1989) considered the normal case. Hwang and Peddada (1994) considered elliptically contoured distributions. The Poisson case was considered by Kushary and Cohen (1991) and Chang and Shinozaki (2006). See also van Eeden (2006).

*Graduate School of Economics, Kyoto University, Yoshida-Honmachi, Sakyo-ku, Kyoto, 606-8501, JAPAN. E-Mail: yasu.stat@gmail.com

†Faculty of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, JAPAN. E-Mail: tatsuya@e.u-tokyo.ac.jp

In this paper, we show that the estimator $\widehat{\mu}_1(X,Y)$ dominates X under the squared error loss function for distributions whose densities are log-concave or log-convex. We assume that Y can be written as $Y = Z + V$, where Z is an independent copy of X while V is a nonnegative error that is either random or fixed. This assumption is satisfied, for example, when the distribution of X is Poisson or normal. Our approach is to write the risk function of $\widehat{\mu}_1(X,Y)$ as

$$E^V\Big[E^{(X,Z)}\Big[\Big(\min\Big\{X, \frac{X+Z+V}{2}\Big\} - \mu\Big)^2\Big]\Big] \tag{1.1}$$

and show that the inner expectation, which converges to the risk of the direct estimator X as $V \to \infty$, is a nondecreasing function of V. We obtain results for distributions that have not been fully considered in the literature, including the gamma and negative binomial distributions with unknown shape parameters. Moreover, since the error V can have any distribution on $[0,\infty)$, our results are relevant from a practical point of view.

There are available in the literature different forms of estimators of ordered parameters for simultaneous as well as componentwise estimation problems. These include maximum likelihood estimators (e.g., Lee, 1981; Misra, Iyer and Singh, 2004; Singh, Misra and Li, 2005; Jena and Tripathy, 2019), mixed estimators (e.g., Katz, 1963; Kumar and Sharma, 1988; Vijayshree and Singh, 1991; Patra and Kumar, 2017), Pitman-type estimators (e.g., Blumenthal and Cohen, 1968a, 1968b; Cohen and Sackrowitz, 1970; Kumar and Sharma, 1989; Kumar and Sharma, 1993; Kumar, Kumar and Tripathi, 2005), Stein-type shrinkage estimators (e.g., Kumar, Tripathi and Misra, 2005), and Brewster–Zidek-type shrinkage estimators (e.g., Vijayshree et al., 1995; Misra et al., 2002; Patra, Kumar and Petropoulos, 2021). However, specific distributions such as normal and exponential distributions are considered in most cases and location/scale families are considered in other cases. In contrast, we show the robustness of $\widehat{\mu}_1(X,Y) = \min\{X,(X+Y)/2\}$ based on the aforementioned distributional assumptions.

In Section 2, we derive conditions for the monotonicity of (1.1). In Section 3, we briefly discuss the estimation of μ_2 using a gamma model. In Section 4, simulation studies are performed. Some concluding remarks are given in Section 5. Proofs and details of an algorithm used in Section 4 are given in the Appendix.

2 Main results

We consider the continuous and discrete cases in Sections 2.1 and 2.2, respectively. In both the sections, we use Lemma 2.1.

Lemma 2.1 *Suppose that X and Z are independently and identically distributed with mean $\mu \in \mathbb{R}$ and with a finite second moment. Suppose that*

$$E[X+Z+v|X>Z+v] \geq 2\mu$$

for all $v \in (0,\infty)\setminus\mathbb{N}$. Then

$$R(v) = E\Big[\Big(\min\Big\{X, \frac{X+Z+v}{2}\Big\} - \mu\Big)^2\Big]$$

is a nondecreasing function of $v \geq 0$, and in particular

$$E\Big[\Big(\min\Big\{X, \frac{X+Z+v}{2}\Big\} - \mu\Big)^2\Big] \leq E[(X-\mu)^2]$$

for all $v \geq 0$.

Proof. We have

$$R'(v) = E\Big[\Big(\frac{X+Z+v}{2} - \mu\Big)1(X > Z+v)\Big]$$
$$= P(X > Z+v)(E[X+Z+v|X > Z+v] - 2\mu)/2$$

for all $v \in (0,\infty) \setminus \mathbb{N}$ even when X and Z are discrete variables, and this proves the desired result. □

2.1 The continuous case

Here, we consider the continuous case.

Proposition 2.1 *Let $c \in [-\infty,\infty)$. Suppose that X and Z are independently and identically distributed according to a probability density function $f\colon (c,\infty) \to (0,\infty)$ with mean $\mu \in (c,\infty)$. Then $E[X+Z+v|X>Z+v] \geq 2\mu$ for all $v > 0$ if (i) f is log-concave or (ii) f is log-convex and* $\lim_{z\to\infty} f(z) = 0$.

Proof. Let F denote the distribution function corresponding to f. Fix $v > 0$.

For part (i), suppose that f is log-concave. Then

$$E[X|X > Z+v] = \frac{\int_{c+v}^{\infty}\big\{xf(x)\int_c^{x-v} f(z)dz\big\}dx}{\int_{c+v}^{\infty}\big\{f(x)\int_c^{x-v} f(z)dz\big\}dx}$$
$$= \frac{\int_{c+v}^{\infty} xf(x)F(x-v)dx}{\int_{c+v}^{\infty} f(x)F(x-v)dx}$$
$$= \frac{\int_c^{\infty} xf(x)\{1(x > c+v)F(x-v)/F(x)\}F(x)dx}{\int_c^{\infty} f(x)\{1(x > c+v)F(x-v)/F(x)\}F(x)dx}.$$

Since the log-concavity of f implies the log-concavity of F (see, for examle, Bagnoli and Bergstrom (2005)), $1(x > c+v)F(x-v)/F(x)$ is a nondecreasing function of $x > c$. Therefore, by the covariance inequality,

$$E[X|X > Z+v] \geq \frac{\int_c^{\infty} xf(x)F(x)dx}{\int_c^{\infty} f(x)F(x)dx} = E[X|X > Z].$$

On the other hand,

$$E[Z+v|X > Z+v] = \frac{\int_c^{\infty}\big\{(z+v)f(z)\int_{z+v}^{\infty} f(x)dx\big\}dz}{\int_c^{\infty}\big\{f(z)\int_{z+v}^{\infty} f(x)dx\big\}dz}$$
$$= \frac{\int_c^{\infty}(z+v)f(z)\{1-F(z+v)\}dz}{\int_c^{\infty} f(z)\{1-F(z+v)\}dz}$$
$$= \frac{\int_c^{\infty} zf(z)\{1(z > c+v)f(z-v)/f(z)\}\{1-F(z)\}dz}{\int_c^{\infty} f(z)\{1(z > c+v)f(z-v)/f(z)\}\{1-F(z)\}dz}.$$

Since $1(z > c+v)f(z-v)/f(z)$ is a nondecreasing function of $z > c$ by the log-concavity of f, we have, by the covariance inequality,

$$E[Z+v|X > Z+v] \geq \frac{\int_c^{\infty} zf(z)\{1-F(z)\}dz}{\int_c^{\infty} f(z)\{1-F(z)\}dz} = E[Z|X > Z].$$

Thus, $E[X+Z+v|X>Z+v] \geq E[X+Z|X>Z] = 2\mu$.

For part (ii), suppose that f is log-convex. Then

$$\begin{aligned} E[X|X>Z+v] &= \frac{\int_{c+v}^{\infty} xf(x)F(x-v)dx}{\int_{c+v}^{\infty} f(x)F(x-v)dx} \\ &= \frac{\int_c^{\infty}(x+v)\{f(x+v)/f(x)\}f(x)F(x)dx}{\int_c^{\infty}\{f(x+v)/f(x)\}f(x)F(x)dx} \\ &\geq \frac{\int_c^{\infty} x\{f(x+v)/f(x)\}f(x)F(x)dx}{\int_c^{\infty}\{f(x+v)/f(x)\}f(x)F(x)dx}. \end{aligned}$$

Since $f(x+v)/f(x)$ is a nondecreasing function of $x>c$, by the covariance inequality,

$$E[X|X>Z+v] \geq \frac{\int_c^{\infty} xf(x)F(x)dx}{\int_c^{\infty} f(x)F(x)dx} = E[X|X>Z].$$

On the other hand,

$$\begin{aligned} E[Z+v|X>Z+v] &= \frac{\int_c^{\infty}(z+v)f(z)\{1-F(z+v)\}dz}{\int_c^{\infty} f(z)\{1-F(z+v)\}dz} \\ &= \frac{\int_c^{\infty}(z+v)[\{1-F(z+v)\}/\{1-F(z)\}]f(z)\{1-F(z)\}dz}{\int_c^{\infty}[\{1-F(z+v)\}/\{1-F(z)\}]f(z)\{1-F(z)\}dz} \\ &\geq \frac{\int_c^{\infty} z[\{1-F(z+v)\}/\{1-F(z)\}]f(z)\{1-F(z)\}dz}{\int_c^{\infty}[\{1-F(z+v)\}/\{1-F(z)\}]f(z)\{1-F(z)\}dz}. \end{aligned}$$

By Theorem 4 of Bagnoli and Bergstrom (2005), $1-F$ is log-convex on (c,∞), which implies that $\{1-F(z+v)\}/\{1-F(z)\}$ is a nondecreasing function of $z>c$. Therefore, by the covariance inequality,

$$E[Z+v|X>Z+v] \geq \frac{\int_c^{\infty} zf(z)\{1-F(z)\}dz}{\int_c^{\infty} f(z)\{1-F(z)\}dz} = E[Z|X>Z],$$

and thus $E[X+Z+v|X>Z+v] \geq E[X+Z|X>Z] = 2\mu$. This completes the proof. □

By Lemma 2.1 and Proposition 2.1, we have the following results.

Example 2.1 Let $g\colon \mathbb{R} \to (0,\infty)$ be a log-concave probability density function with mean 0. Suppose that X and Z are independently and identically distributed according to the probability density function $g(x-\mu)$, $x \in \mathbb{R}$, where $\mu \in \mathbb{R}$ is an unknown location parameter. Suppose that we observe $Y = Z+V$ for an independent error $V \geq 0$ which is random or fixed. Then we have $E[\{\min\{X,(X+Y)/2\}-\mu\}^2] \leq E[(X-\mu)^2]$. This is a known result when X and Y are normal (see, for example, Lee (1981)).

Example 2.2 If $X \sim \mathrm{Ga}(\alpha_1,\beta)$ and $Y \sim \mathrm{Ga}(\alpha_2,\beta)$ are independent with $\alpha_1 \leq \alpha_2$, then $\min\{X,(X+Y)/2\}$ dominates X in estimating $E[X] = \alpha_1/\beta$. (Here, we use the fact that Y can be decomposed as $Z+V$, where $Z \overset{\mathrm{d}}{=} X$ and $V \sim \mathrm{Ga}(\alpha_2-\alpha_1,\beta)$ are independent.) This result has not been reported in the literature, so far as we know.

When we consider scale familes, Proposition 2.1 is not directly applicable. In fact, we modify Lemma 2.1 for this case.

Proposition 2.2 *Suppose that X and Z are independently and identically distributed positive random variables having mean $\mu \in (0,\infty)$ and a finite second moment. Suppose that $xf'(x)/f(x)$ is a nonincreasing function of $x>0$. Then*

$$\widetilde{R}(v) = E\Big[\Big\{ \min\Big\{X, \frac{X+(1+v)Z}{2}\Big\} - \mu\Big\}^2\Big]$$

is a nondecreasing function of $v \geq 0$, and in particular

$$E\Big[\Big\{ \min\Big\{X, \frac{X+(1+v)Z}{2}\Big\} - \mu\Big\}^2\Big] \leq E[(X-\mu)^2]$$

for all $v \geq 0$.

Example 2.3 If $X/\gamma_1 \sim \mathrm{Ga}(\alpha,1)$ and $Y/\gamma_2 \sim \mathrm{Ga}(\alpha,1)$ are independent with $\gamma_1 \leq \gamma_2$, then $\min\{X,(X+Y)/2\}$ dominates X in estimating $E[X]=\alpha\gamma_1$, which is a known result (see Section 4 of Hwang and Peddada (1994)).

2.2 The discrete case

In the discrete case, the situation is more complicated since $P(X=Z)>0$. Let $\mathbb{N}_0 = \mathbb{N}\cup\{0\} = \{0,1,2,\dots\}$. Let $2\mathbb{N}_0 = \{2n | n \in \mathbb{N}_0\}$ and $2\mathbb{N}_0+1 = \{2n+1 | n \in \mathbb{N}_0\}$.

Proposition 2.3 *Suppose that X and Z are independently and identically distributed on $\mathbb{N}_0$ with mean $\mu \in (0,\infty)$ according to the probability mass function defined by $f\colon [0,\infty) \to (0,\infty)$. Then $E[X+Z+v|X>Z+v] \geq 2\mu$ for all $v \in (0,\infty)\setminus\mathbb{N}$ if one of the following two conditions are satisfied.*

(i) *f is log-concave, $\{\partial/(\partial x)\}\{\log f(x)\}$ is a convex function of $x \in [0,\infty)$, and $E[X+Z|X+Z \in 2\mathbb{N}_0] \leq E[X+Z|X+Z\in 2\mathbb{N}_0+1]$.*

(ii) *f is log-convex and $\lim_{z\to\infty} f(z)=0$.*

Proof. Let F denote the distribution function corresponding to f and fix $v \in (0,\infty)\setminus\mathbb{N}$. Let $[v]$ denote the integer satisfying $[v] \leq v < [v]+1$. Then

$$E[X|X>Z+v] = \frac{\sum_{x=0}^{\infty}\big\{xf(x)\sum_{z<x-v}f(z)\big\}}{\sum_{x=0}^{\infty}\big\{f(x)\sum_{z<x-v}f(z)\big\}} = \frac{\sum_{x=0}^{\infty}xf(x)F(x-[v]-1)}{\sum_{x=0}^{\infty}f(x)F(x-[v]-1)}$$

and

$$E[Z+v|X>Z+v] = \frac{\sum_{z=0}^{\infty}\big\{(z+v)f(z)\sum_{x>z+v}f(x)\big\}}{\sum_{z=0}^{\infty}\big\{f(z)\sum_{x>z+v}f(x)\big\}} \geq \frac{\sum_{z=0}^{\infty}(z+[v])f(z)\{1-F(z+[v])\}}{\sum_{z=0}^{\infty}f(z)\{1-F(z+[v])\}}.$$

Similarly,

$$E[X|X>Z] = \frac{\sum_{x=0}^{\infty}\big\{xf(x)\sum_{z<x}f(z)\big\}}{\sum_{x=0}^{\infty}\big\{f(x)\sum_{z<x}f(z)\big\}} = \frac{\sum_{x=0}^{\infty}xf(x)F(x-1)}{\sum_{x=0}^{\infty}f(x)F(x-1)}$$

and

$$E[Z|X > Z] = \frac{\sum_{z=0}^{\infty}\left\{zf(z)\sum_{x>z}f(x)\right\}}{\sum_{z=0}^{\infty}\left\{f(z)\sum_{x>z}f(x)\right\}} = \frac{\sum_{z=0}^{\infty}zf(z)\{1-F(z)\}}{\sum_{z=0}^{\infty}f(z)\{1-F(z)\}}.$$

By an argument similar to the proof of Proposition 2.1, it follows from parts (i) and (ii) of Lemma 2.2 below that $E[X|X > Z + v] \geq E[X|X > Z]$ and that $E[Z + v|X > Z + v] \geq E[Z|X > Z]$. Furthermore, by part (iii) of Lemma 2.2 given below, $E[X + Z|X \geq Z + 1] \geq E[X + Z]$. Thus, $E[X+Z+v|X > Z+v] \geq E[X+Z|X \geq Z+1] \geq E[X+Z] = 2\mu$. This proves the proposition. □

Lemma 2.2 *Suppose that $f\colon [0,\infty) \to (0,\infty)$ defines a probability mass function on $\mathbb{N}_0 = \{0, 1, 2, \dots\}$. Let F denote the distribution function corresponding to f. Let X and Z be independently distributed according to F. Let $k \in \mathbb{N}_0$.*

(i) *If f is log-concave, then for any $x_1, x_2 \in \mathbb{N}_0$ satisfying $x_1 \leq x_2$, we have $F(x_1)/F(x_1+k) \leq F(x_2)/F(x_2 + k)$.*

(ii) *If f is log-convex and $\lim_{z\to\infty} f(z) = 0$, then for any $x_1, x_2 \in \mathbb{N}_0$ satisfying $x_1 \leq x_2$, we have $\{1 - F(x_1 + k)\}/\{1 - F(x_1)\} \leq \{1 - F(x_2 + k)\}/\{1 - F(x_2)\}$.*

(iii) *If either $\{\partial/(\partial x)\}\{\log f(x)\}$ is a convex function of $x \in [0,\infty)$ and $E[X + Z|X + Z \in 2\mathbb{N}_0] \leq E[X+Z|X+Z \in 2\mathbb{N}_0+1]$ or f is log-convex, then $E[X+Z|X \geq Z+1] \geq E[X+Z]$.*

By Lemma 2.1 and Proposition 2.3, we have the following results.

Example 2.4 Note that $W \sim \mathrm{Po}(\mu)$ implies $P(W \in 2\mathbb{N}_0) = (1 + e^{-2\mu})/2$ and $E[W|W \in 2\mathbb{N}_0 + 1] = \mu(1 + e^{-2\mu})/(1 - e^{-2\mu}) \geq E[W] \geq E[W|W \in 2\mathbb{N}_0]$. If $X \sim \mathrm{Po}(\mu_1)$ and $Y \sim \mathrm{Po}(\mu_2)$ are independent with $\mu_1 \leq \mu_2$, then $\min\{X, (X + Y)/2\}$ dominates X, which is a known result (see, for example, Kushary and Cohen (1991)). Meanwhile, the unbalanced case needs to be studied further (Chang and Shinozaki (2006)). If $X \sim \mathrm{Po}(m_1\mu_1)$ and $Y \sim \mathrm{Po}(m_2\mu_2)$ with m_1 and m_2 known and if $m_1 \leq m_2$, we see that $\min\{X/m_1, (X+Y)/(2m_1)\}$ dominates X_1/m_1 for the estimation of μ_1. However, it has not been determined whether $\min\{X/m_1, (X+Y)/(m_1+m_2)\}$ dominates X_1/m_1.

Example 2.5 Note that $W|\lambda \sim \mathrm{Po}(\lambda)$ with $\lambda \sim \mathrm{Ga}(r, p/(1 - p))$ implies $W \sim \mathrm{NB}(r, p)$ and $E[W|W \in 2\mathbb{N}_0 + 1] = E[E[W1(W \in 2\mathbb{N}_0 + 1)|\lambda]]/E[E[1(W \in 2\mathbb{N}_0 + 1)|\lambda]] \geq E[\lambda P(W \in 2\mathbb{N}_0 + 1|\lambda)]/E[P(W \in 2\mathbb{N}_0 + 1|\lambda)] \geq E[\lambda] = E[W]$ by Example 2.4 and by the covariance inequality. If $X \sim \mathrm{NB}(r_1, p)$ and $Y \sim \mathrm{NB}(r_2, p)$ are independent with $r_1 \leq r_2$, then $\min\{X, (X + Y)/2\}$ dominates X in estimating $E[X] = r_1(1 - p)/p$.

Example 2.6 If X and Z are independently distributed according to the probability mass function $\int_0^1 \mathrm{Geo}(x|p)d\Pi(p)$, $x \in \mathbb{N}_0$, where Π is a probability measure on $(0, 1)$, then since the marginal distribution is log-convex, we have $E[(\min\{X, (X + Z + V)/2\} - \mu)^2] \leq E[(X - \mu)^2]$ for any nonnegative error V that is either random or fixed.

Remark 2.1 When the sample space is finite, Lemma 2.1 may not be useful. For instance, if X and Z are independently distributed according to $\mathrm{Bin}(n, p)$, then $E[X + Z|X - Z = t] - E[X + Z] \to -t$ as $p \to 1$ for any $t \in \mathbb{N}$, and therefore $E[X + Z|X \geq Z + 1] < E[X + Z]$ for some $p \in (0, 1)$. (If $p = 1/2$, the condition of Lemma 2.1 can be verified by noting that $E[1(X > Z + v)|X + Z = w]$ is a function of $w \in \{0, 1, \dots, 2n\}$ which is symmetric around n).

3 Estimation of the Larger of Two Ordered Gamma Shape Parameters

In the previous section, we showed that in many cases, $\min\{X, (X+Y)/2\}$ is a good estimator of $\mu_1 = E[X]$ when $\mu_1 \le \mu_2 = E[Y]$. Then, at first glance, it may seem natural that $\max\{Y, (X+Y)/2\} = X+Y-\min\{X, (X+Y)/2\}$ should be a good estimator of μ_2. However, by considering a simple gamma model, we see that this is not always true.

Let X and Y be independent observations distributed according to $\mathrm{Ga}(\alpha_1, 1)$ and $\mathrm{Ga}(\alpha_2, 1)$, respectively.

Proposition 3.1 *Consider the estimation of* α_2 *under the squared error loss function.*

(i) $\max\{Y, (X+Y)/2\}$ *does not dominate* Y *under the restriction* $\alpha_1 \le \alpha_2$.

(ii) $\max\{Y, (X+Y)/2\}$ *dominates* Y *under the restriction* $\alpha_1 + 1 < \alpha_2$.

We note that the above result depends on the loss function.

Proposition 3.2

(i) *Consider the estimation of* α_1 *under Stein's loss function. Then* $\min\{X, (X+Y)/2\}$ *dominates* X *under the restriction* $\alpha_1 \le \alpha_2$.

(ii) *Consider the estimation of* α_2 *under Stein's loss function. Then* $\max\{Y, (X+Y)/2\}$ *dominates* Y *under the restriction* $\alpha_1 \le \alpha_2$.

4 Numerical Studies

In this section, we investigate through simulation the numerical performance of several estimators of $\mu_1 = E[X]$ under the squared error loss function and under the restriction that $\mu_1 \le \mu_2 = E[Y]$, where Y is independent of X. We consider the following estimators:

U: the unbiased estimator X,

T: the truncated estimator $\min\{X, (X+Y)/2\}$,

B: the Bayesian estimator against the prior $(\mu_1, \mu_2) \sim \pi^{\mathrm{J}}(\mu_1, \mu_2) 1(\mu_1 \le \mu_2)$, where $\pi^{\mathrm{J}}(\mu_1, \mu_2)$ is the Jeffreys prior in the unrestricted case.

The percentage relative improvement in average loss (PRIAL) of an estimator $\delta(X, Y)$ over X is defined by

$$\mathrm{PRIAL} = 100[E[(X-\mu_1)^2] - E[\{\delta(X,Y) - \mu_1\}^2]]/E[(X-\mu_1)^2].$$

We consider the problems of estimating smaller normal location, gamma scale, Poisson rate, gamma shape, and negative binomial shape parameters in Sections 4.1, 4.2, 4.3, 4.4, and 4.5, respectively. We note that the Bayesian estimator is not a robust estimator since its form depends on the model we consider. For simplicity, we assume that all nuisance parameters other than μ_2 are known.

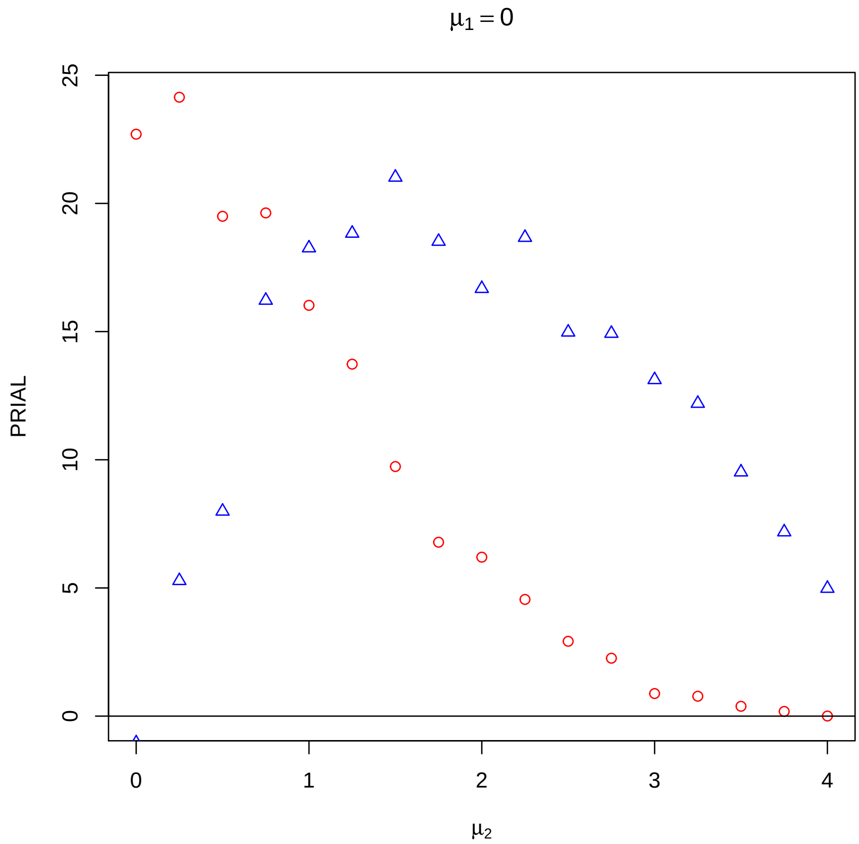

Figure 1: Values of PRIAL of T and B over U. The circles and triangles correspond to T and B, respectively.

4.1 Estimation of a smaller normal location parameter

Suppose that X and Y are normal with unit variance. In this case, the Bayesian estimator (B) is

$$\frac{\int_{-\infty}^{\infty} \mu_1 g(\mu_1 - X)\{1 - G(\mu_1 - Y)\} d\mu_1}{\int_{-\infty}^{\infty} g(\mu_1 - X)\{1 - G(\mu_1 - Y)\} d\mu_1},$$

where g and G denote the probability density and distribution functions of the standard normal distribution, respectively. The integrals are calculated via the Monte Carlo simulation with $10,000$ replications. We obtain approximated values of the risk function by simulation with $1,000$ replications.

Results are shown in Figure 1. As expected, the truncated estimator (T) is always better than the unbiased estimator (U). Also, T is better than B when $\mu_2 - \mu_1$ is small.

4.2 Estimation of a smaller gamma scale parameter

Suppose that X/γ_1 and Y/γ_2 are distributed as Ga$(5,1)$ with $\mu_i = 5\gamma_i$ for $i = 1,2$. In this case, B is given by

$$5X \frac{\int_0^\infty (1/u_1) g(u_1) G((Y/X)u_1) du_1}{\int_0^\infty g(u_1) G((Y/X)u_1) du_1},$$

where g and G denote the probability density and distribution functions of Ga$(5,1)$, respectively. As in Section 4.1, the integrals are calculated via the Monte Carlo simulation with $10,000$ replications and we obtain approximated values of the risk function by simulation with $1,000$ replications.

Results are shown in Figure 2. It can be seen that T dominates U, while the risk values of B are larger than those of U when $\gamma_2 - \gamma_1$ is large.

In the present case, it is well known that X is uniformly dominated by the estimator $\{5/(5+1)\}X$, which is improved upon by $\min\{X/(5+1), (X+Y)/(2\times 5+1)\}$ and

$$5X \frac{\int_0^\infty u_1 g(u_1) G((Y/X)u_1) du_1}{\int_0^\infty {u_1}^2 g(u_1) G((Y/X)u_1) du_1}$$

(see, for example, Kubokawa and Saleh (1994)). However, we numerically confirmed that T is not necessarily dominated by these improved estimators especially when $\gamma_2 - \gamma_1$ is small.

4.3 Estimation of a smaller Poisson rate parameter

Suppose that X and Y are Poisson variables. Then B is given by

$$\frac{\int_0^\infty \mu_1 \mathrm{Ga}(\mu_1 | X+1/2, 1)\{1 - G(\mu_1; Y)\} d\mu_1}{\int_0^\infty \mathrm{Ga}(\mu_1 | X+1/2, 1)\{1 - G(\mu_1; Y)\} d\mu_1},$$

where

$$G(\mu_1; Y) = \int_0^{\mu_1} \mathrm{Ga}(\mu_2 | Y + 1/2, 1) d\mu_2$$

for $\mu_1 \in (0, \infty)$. We also consider another smooth estimator, denoted by B†, given by

$$\frac{\int_0^\infty \mathrm{Ga}(\mu_1 | X+1/2, 1)\{1 - G(\mu_1; Y)\} d\mu_1}{\int_0^\infty (1/\mu_1) \mathrm{Ga}(\mu_1 | X+1/2, 1)\{1 - G(\mu_1; Y)\} d\mu_1},$$

which is generalized Bayes under the standardized squared error loss. The integrals and expectations are calculated as in Sections 4.1 and 4.2.

Results are shown in Figure 3. The robustness of T is confirmed. Although B† can be efficient in many cases, it seems that T is not uniformly dominated by B†.

4.4 Estimation of a smaller gamma shape parameter

Suppose that $X \sim \mathrm{Ga}(\mu_1, 1)$ and $Y \sim \mathrm{Ga}(\mu_2, 1)$. Then B is given by

$$\int_{(0,1)^2} \{\omega_1/(1-\omega_1)\} 1(\omega_1 \le \omega_2) \frac{\sqrt{\psi_1(\omega_1/(1-\omega_1))}}{(1-\omega_1)^2} \frac{X^{\omega_1/(1-\omega_1)}}{\Gamma(\omega_1/(1-\omega_1))} \frac{\sqrt{\psi_1(\omega_2/(1-\omega_2))}}{(1-\omega_2)^2} \frac{Y^{\omega_2/(1-\omega_2)}}{\Gamma(\omega_2/(1-\omega_2))} d(\omega_1, \omega_2)$$

$$/ \int_{(0,1)^2} 1(\omega_1 \le \omega_2) \frac{\sqrt{\psi_1(\omega_1/(1-\omega_1))}}{(1-\omega_1)^2} \frac{X^{\omega_1/(1-\omega_1)}}{\Gamma(\omega_1/(1-\omega_1))} \frac{\sqrt{\psi_1(\omega_2/(1-\omega_2))}}{(1-\omega_2)^2} \frac{Y^{\omega_2/(1-\omega_2)}}{\Gamma(\omega_2/(1-\omega_2))} d(\omega_1, \omega_2),$$

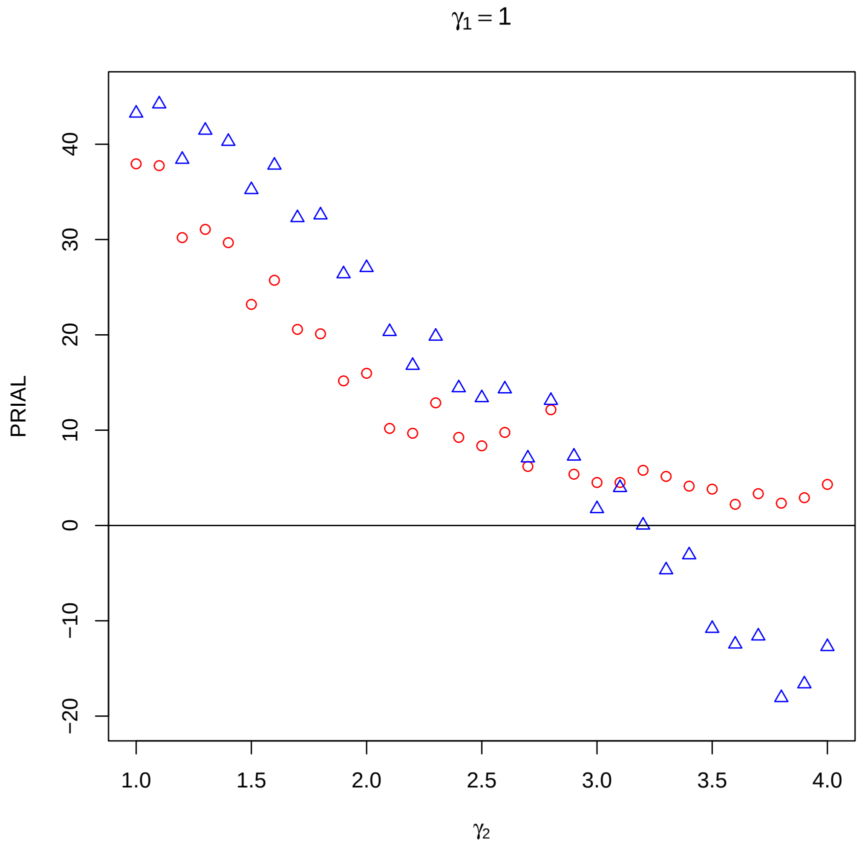

Figure 2: Values of PRIAL of T and B over U. The circles and triangles correspond to T and B, respectively.

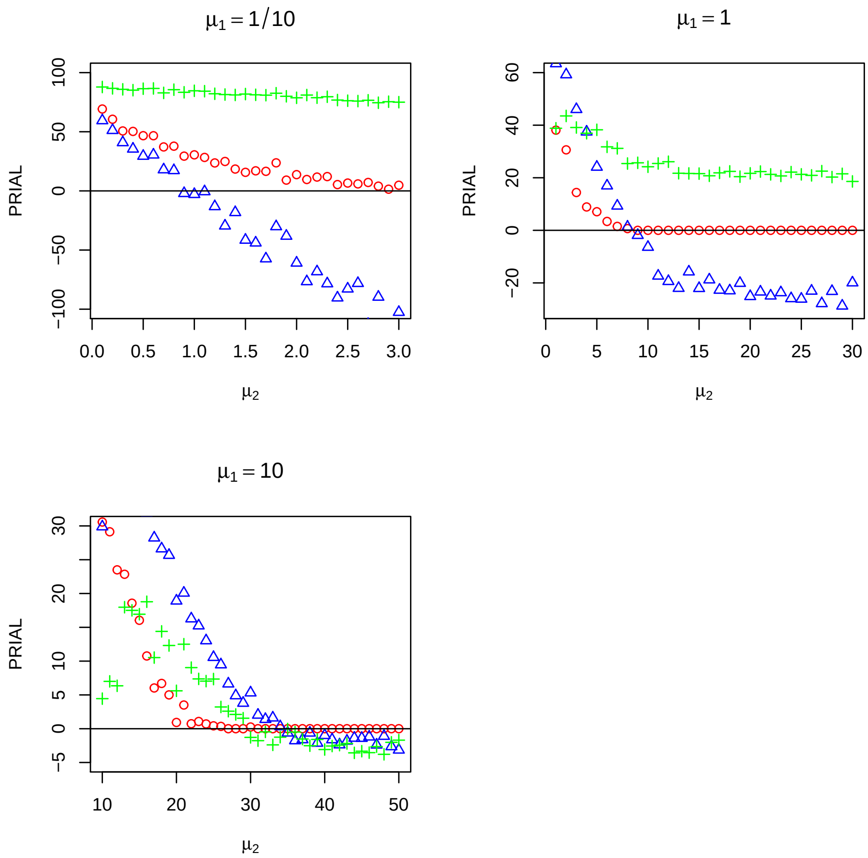

Figure 3: Values of PRIAL of T, B, and $B^{\dagger}$ over U. The circles, triangles, and pluses correspond to T, B, and $B^{\dagger}$, respectively.

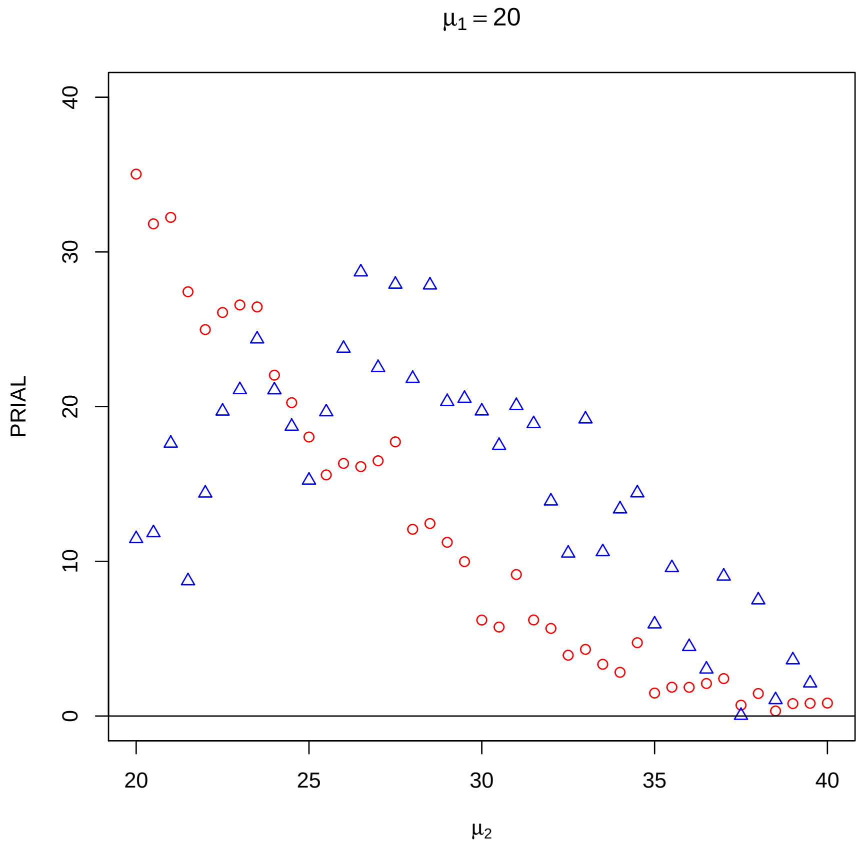

Figure 4: Values of PRIAL of T and B over U. The circles and triangles correspond to T and B, respectively.

where ψ_1 denotes the trigamma function. The integrals are calculated via the Monte Carlo simulation with $10,000$ replications and we obtain approximated values of the risk function by simulation with $1,000$ replications.

Results are shown in Figure 4 and Table 1 and it is numerically confirmed that T dominates U. Also, from Table 1, we see that B is not always better than U.

Table 1: Values of PRIAL of T and B over U. Here, we set $\mu_1 = \mu_2$.

(μ_1, μ_2)	T	B
$(5, 5)$	39.05	39.06
$(10, 10)$	34.65	27.86
$(15, 15)$	32.98	19.12
$(20, 20)$	31.67	4.40
$(25, 25)$	33.63	-4.73
$(30, 30)$	32.04	-34.23

4.5 Estimation of a smaller negative binomial shape parameter

Suppose that $X \sim \mathrm{NB}(\mu_1, 1/2)$ and $Y \sim \mathrm{NB}(\mu_2, 1/2)$. Then B is given by

$$\int_{(0,\infty)^2} \mu_1 1(\mu_1 \le \mu_2) \pi^{\mathrm{J}}(\mu_1) \binom{\mu_1 + X - 1}{X} \Big(\frac{1}{2}\Big)^{\mu_1} \pi^{\mathrm{J}}(\mu_2) \binom{\mu_2 + Y + 1}{Y} \Big(\frac{1}{2}\Big)^{\mu_2} d(\mu_1, \mu_2)$$
$$/ \int_{(0,\infty)^2} 1(\mu_1 \le \mu_2) \pi^{\mathrm{J}}(\mu_1) \binom{\mu_1 + X - 1}{X} \Big(\frac{1}{2}\Big)^{\mu_1} \pi^{\mathrm{J}}(\mu_2) \binom{\mu_2 + Y + 1}{Y} \Big(\frac{1}{2}\Big)^{\mu_2} d(\mu_1, \mu_2),$$

where

$$\pi^{\mathrm{J}}(\mu) = \sqrt{\sum_{k=0}^{\infty} \mathrm{NB}(k|\mu, 1/2)\{\psi_1(\mu) - \psi_1(\mu + k)\}}$$

for $\mu \in (0, \infty)$. In order to calculate the integrals, we use the Metropolis–Hastings algorithm to sample from the density proportional to $\pi^{\mathrm{J}}(\mu)(1/2)^{\mu}$, $\mu \in (0, \infty)$; see the Appendix for details. We generate $15,000$ samples after discarding $5,000$ samples. We obtain approximated values of the risk function by simulation with $1,000$ replications. To reduce computation time, we consider only cases where $\mu_1 = \mu_2$.

Results are shown in Table 2. T performs worse than B when the mean is small but better when the mean is large.

Table 2: Values of PRIAL of T and B over U. Here, we set $\mu_1 = \mu_2$.

(μ_1, μ_2)	T	B
$(2, 2)$	45.46313	72.84997
$(4, 4)$	37.12357	54.97274
$(6, 6)$	37.14836	21.45862
$(8, 8)$	31.87264	−38.44798

5 Discussion

In this paper, we focused on the estimator $\min\{X, (X + Y)/2\}$ for the estimation of $\mu_1 = E[X] \le \mu_2 = E[Y]$. One important extension would be to obtain similar results under weaker assumptions such as the assumption that $P(X \ge Y) \le P(X \le Y)$. In general, we would have to consider a general class of truncated estimators to reflect distributional assumptions. For instance, if $X \sim \mathrm{N}(\mu_1, \sigma_1^2)$ and $Y \sim \mathrm{N}(\mu_2, \sigma_2^2)$ are independent with $\mu_1 \le \mu_2$, where μ_1 and μ_2 are unknown while σ_1^2 and σ_2^2 are known, then $\min\{X, (X+Y)/2\}$ does not necessarily dominate X since $X \to \mu_1$ but $\min\{X, (X + Y)/2\} \nrightarrow \mu_1$ as $\sigma_1^2 \to 0$ with σ_2^2 fixed. Lee (1981) showed that $\min\{X, (X/\sigma_1^2 + Y/\sigma_2^2)/(1/\sigma_1^2 + 1/\sigma_2^2)\}$ dominates X.

Although Kubokawa and Saleh (1994) and recently Garg and Misra (2021) investigated smooth estimators under order restrictions, further study is needed in this direction. For example, in the Poisson case, we can check that the direct estimator X is minimax for the standardized squared error loss. An important question is whether we can construct an admissible (generalized) Bayes estimator which dominates X. A candidate prior is the uniform prior on $(\Lambda, \theta) = (\mu_1 + \mu_2, \mu_1/(\mu_1 + \mu_2)) \in (0, \infty) \times (0, 1/2)$. However, it remains to be proved that the associated estimator dominates X.

6 Appendix

6.1 Proofs

Here we prove Propositions 2.2, 3.1, and 3.2 and Lemma 2.2.

Proof of Proposition 2.2. We have

$$\begin{aligned}
&\frac{2\widetilde{R}'(v)}{E[Z1(X > (1+v)Z)]} = 2E\Big[\Big\{\frac{X + (1+v)Z}{2} - \mu\Big\} Z1(X > (1+v)Z)\Big] \Big/ E[Z1(X > (1+v)Z)] \\
&= \frac{\int_0^\infty x f(x) \{\int_0^{x/(1+v)} z f(z) dz\} dx}{\int_0^\infty f(x) \{\int_0^{x/(1+v)} z f(z) dz\} dx} + \frac{\int_0^\infty (1+v) z^2 f(z) \{\int_{(1+v)z}^\infty f(x) dx\} dz}{\int_0^\infty z f(z) \{\int_{(1+v)z}^\infty f(x) dx\} dz} - 2\mu \\
&= \frac{\int_0^\infty x f(x) \{\int_0^x z f(z/(1+v)) dz\} dx}{\int_0^\infty f(x) \{\int_0^x z f(z/(1+v)) dz\} dx} + \frac{\int_0^\infty z^2 f(z/(1+v)) \{\int_z^\infty f(x) dx\} dz}{\int_0^\infty z f(z/(1+v)) \{\int_z^\infty f(x) dx\} dz} \\
&\quad - \frac{\int_0^\infty x f(x) \{\int_0^x f(z) dz\} dx}{\int_0^\infty f(x) \{\int_0^x f(z) dz\} dx} - \frac{\int_0^\infty z f(z) \{\int_z^\infty f(x) dx\} dz}{\int_0^\infty f(z) \{\int_z^\infty f(x) dx\} dz}. \qquad (6.1)
\end{aligned}$$

Now, by assumption,

$$z \frac{\partial}{\partial z} \log \frac{z f(z/(1+v))}{f(z)} = 1 + \frac{z}{1+v} \frac{f'(z/(1+v))}{f(z/(1+v))} - z \frac{f'(z)}{f(z)} \geq 0.$$

Therefore, $z f(z/(1+v))/f(z)$ is a nondecreasing function of $z > 0$, which implies that $\int_0^x z f(z/(1+v)) dz / \int_0^x f(z) dz$ is a nondecreasing function of $x > 0$. Thus, by the covariance inequality, the right-hand side of (6.1) is nonnegative. The result follows. □

Proof of Lemma 2.2. We prove parts (i) and (ii) based on the proofs of Lemmas 3 and 4 of Bagnoli and Bergstrom (2005), which are for the continuous case. Without loss of generality, we can assume that $k = 1$ and $x_2 = x_1 + 1$. Let $F(-1) = f(-1) = 0$.

For part (i), note that $f(x-1)/f(x) \geq f(z-1)/f(z)$ for all $x, z \in \mathbb{N}_0$ with $x \geq z$ by the log-concavity of f. Then, for all $x \in \mathbb{N}_0$,

$$\frac{f(x) - f(x-1)}{f(x)} F(x) = \frac{f(x) - f(x-1)}{f(x)} \sum_{z=0}^{x} f(z) \leq \sum_{z=0}^{x} \frac{f(z) - f(z-1)}{f(z)} f(z) = f(x),$$

which implies that

$$f(x-1) \geq f(x) F(x-1)/F(x).$$

Therefore,

$$\begin{aligned}
\frac{F(x_2)}{F(x_2+k)} - \frac{F(x_1)}{F(x_1+k)} &= \frac{F(x_1+1)}{F(x_1+2)} - \frac{F(x_1)}{F(x_1+1)} = \frac{f(x_1+1)}{F(x_1+1)} - \frac{f(x_1+2)}{F(x_1+2)} \\
&\geq \frac{f(x_1+2)}{F(x_1+2)} - \frac{f(x_1+2)}{F(x_1+2)} = 0.
\end{aligned}$$

For part (ii), note that $f(x)/f(x+1) \geq f(z)/f(z+1)$ for all $x, z \in \mathbb{N}_0$ with $x \leq z$ by the log-convexity of f. Then, for all $x \in \mathbb{N}_0$,

$$\frac{f(x+1)-f(x)}{f(x+1)}\{1-F(x)\} = \frac{f(x+1)-f(x)}{f(x+1)}\sum_{z=x}^{\infty} f(z+1) \leq \sum_{z=x}^{\infty}\frac{f(z+1)-f(z)}{f(z+1)}f(z+1) = -f(x),$$

where the last equality follows from the assumption that $\lim_{z\to\infty} f(z) = 0$, and thus we have

$$f(x) \geq f(x+1)\{1-F(x-1)\}/\{1-F(x)\}.$$

Hence,

$$\begin{aligned}\frac{1-F(x_2+k)}{1-F(x_2)} - \frac{1-F(x_1+k)}{1-F(x_1)} &= \frac{1-F(x_1+2)}{1-F(x_1+1)} - \frac{1-F(x_1+1)}{1-F(x_1)} = \frac{f(x_1+1)}{1-F(x_1)} - \frac{f(x_1+2)}{1-F(x_1+1)} \\ &\geq \frac{f(x_1+2)}{1-F(x_1+1)} - \frac{f(x_1+2)}{1-F(x_1+1)} = 0.\end{aligned}$$

For part (iii), let $W = X + Z$ and $T = X - Z$. Let $S = 1(W \in 2\mathbb{N}_0 + 1)$. Suppose first that $\{\partial/(\partial x)\}\{\log f(x)\}$ is a convex function of $x \in [0, \infty)$ and that $E[W|S = 0] \leq E[W|S = 1]$. Then

$$E[W|T \geq 1] = E[W||T| \geq 1] = \frac{\sum_{w=0}^{\infty} w \sum_{\substack{-w\leq t\leq w \\ w-t\in 2\mathbb{N}_0 \\ |t|\geq 1}} f\Big(\frac{w+t}{2}\Big) f\Big(\frac{w-t}{2}\Big)}{\sum_{w=0}^{\infty} \sum_{\substack{-w\leq t\leq w \\ w-t\in 2\mathbb{N}_0 \\ |t|\geq 1}} f\Big(\frac{w+t}{2}\Big) f\Big(\frac{w-t}{2}\Big)}$$

and

$$E[W] = \frac{\sum_{w=0}^{\infty} w \sum_{\substack{-w\leq t\leq w \\ w-t\in 2\mathbb{N}_0}} f\Big(\frac{w+t}{2}\Big) f\Big(\frac{w-t}{2}\Big)}{\sum_{w=0}^{\infty} \sum_{\substack{-w\leq t\leq w \\ w-t\in 2\mathbb{N}_0}} f\Big(\frac{w+t}{2}\Big) f\Big(\frac{w-t}{2}\Big)}.$$

Note that for $w \in \mathbb{N}_0$,

$$\sum_{\substack{-w\leq t\leq w \\ w-t\in 2\mathbb{N}_0 \\ |t|\geq 1}} f\Big(\frac{w+t}{2}\Big) f\Big(\frac{w-t}{2}\Big) \Big/ \sum_{\substack{-w\leq t\leq w \\ w-t\in 2\mathbb{N}_0}} f\Big(\frac{w+t}{2}\Big) f\Big(\frac{w-t}{2}\Big) = \begin{cases} \rho(w), & \text{if } w \in 2\mathbb{N}_0, \\ 1, & \text{if } w \in 2\mathbb{N}_0 + 1, \end{cases}$$

where

$$\begin{aligned}\rho(w) &= \sum_{\substack{-w\leq t\leq w \\ w-t\in 2\mathbb{N}_0 \\ |t|\geq 2}} f\Big(\frac{w+t}{2}\Big) f\Big(\frac{w-t}{2}\Big) \Big/ \sum_{\substack{-w\leq t\leq w \\ w-t\in 2\mathbb{N}_0}} f\Big(\frac{w+t}{2}\Big) f\Big(\frac{w-t}{2}\Big) \\ &= 1 - 1\Big/ \sum_{\substack{-w\leq t\leq w \\ w-t\in 2\mathbb{N}_0}} f\Big(\frac{w+t}{2}\Big) f\Big(\frac{w-t}{2}\Big) \Big/ \Big\{f\Big(\frac{w}{2}\Big)\Big\}^2 \in [0, 1]\end{aligned}$$

is a nondecreasing function of $w \in 2\mathbb{N}_0$ by assumption. Then it follows from Theorem 2.1 of Bhattacharya (1984) that

$$E[W|T \geq 1] = \frac{E[W\{(1-S)\rho(W)+S\}]}{E[(1-S)\rho(W)+S]} \geq E[W],$$

since $E[W|S]$ and $E[(1-S)\rho(W)+S|S]$ are nondecreasing functions of S by assumption and since W and $(1-S)\rho(W)+S$ are nondecreasing functions of W. Next, suppose instead that f is log-convex. Then

$$\begin{aligned}
E[X+Z|T \geq 1] &= E[E[X+Z|T]|T \geq 1] \\
&= E\Big[\frac{\sum_{z=0}^{\infty}(2z+T)f(z+T)f(z)}{\sum_{z=0}^{\infty} f(z+T)f(z)}\Big|T \geq 1\Big] \\
&\geq E\Big[\frac{\sum_{z=0}^{\infty}(2z+T)\{f(z)\}^2}{\sum_{z=0}^{\infty}\{f(z)\}^2}\Big|T \geq 1\Big] \\
&\geq E\Big[\frac{\sum_{z=0}^{\infty}2z\{f(z)\}^2}{\sum_{z=0}^{\infty}\{f(z)\}^2}\Big|T \geq 1\Big] = E[X+Z|T=0],
\end{aligned}$$

where the first inequality follows from the covariance inequality since $f(z+T)/f(z)$ is a nondecreasing function of $z \in \mathbb{N}_0$ by assumption. Thus,

$$\begin{aligned}
E[W] &= P(T=0)E[W|T=0] + P(T \geq 1)E[W|T \geq 1] + P(T \leq -1)E[W|T \leq -1] \\
&= P(T=0)E[W|T=0] + \{1-P(T=0)\}E[W|T \geq 1] \leq E[W|T \geq 1].
\end{aligned}$$

This completes the proof. □

Proof of Proposition 3.1. Let $W = X+Y$ and $R = X/W$. Let $\Delta = E[(\max\{Y, (X+Y)/2\} - \alpha_2)^2] - E[(Y-\alpha_2)^2]$. Note that $W \sim \mathrm{Ga}(\alpha_1+\alpha_2, 1)$ and $R \sim \mathrm{Beta}(\alpha_1, \alpha_2)$ and these are mutually independent.

For part (i), suppose that $\alpha_1 = \alpha_2$. Then it can be seen that

$$\begin{aligned}
\Delta &= E\Big[\Big[W^2\Big\{\Big(\frac{1}{2}\Big)^2 - (1-R)^2\Big\} - 2\alpha_2 W\Big\{\frac{1}{2} - (1-R)\Big\}\Big]1\Big(1-R<\frac{1}{2}\Big)\Big] \\
&= E\Big[\Big[\{2\alpha_2 + (2\alpha_2)^2\}\Big\{\Big(\frac{1}{2}\Big)^2 - (1-R)^2\Big\} - 2\alpha_2(2\alpha_2)\Big\{\frac{1}{2} - (1-R)\Big\}\Big]1\Big(1-R<\frac{1}{2}\Big)\Big] \\
&\sim \alpha_2/4 > 0
\end{aligned}$$

as $\alpha_2 \to 0$.

For part (ii), let $\phi(R) = \max\{Y, (X+Y)/2\}/Y = \max\{1, 1/\{2(1-R)\}\}$. Then

$$\begin{aligned}
\Delta &= E[\{\phi(R)-1\}(1-R)[\{\phi(R)+1\}(1-R)W^2 - 2\alpha_2 W]] \\
&= E[\{\phi(R)-1\}(1-R)[\{\phi(R)+1\}(1-R)(\alpha_1+\alpha_2+1) - 2\alpha_2]](\alpha_1+\alpha_2).
\end{aligned}$$

Since $\{\phi(R)+1\}(1-R) \leq 1$ when $\phi(R)-1 \neq 0$ and since $\alpha_1+\alpha_2+1 < 2\alpha_2$ by assumption, it follows that $\Delta < 0$. □

Proof of Proposition 3.2. Let $W = X+Y$ and $R = X/W$. For part (i), let $\phi_1(R) = \min\{X, (X+Y)/2\}/X = \min\{1, 1/(2R)\}$. Then we have

$$\begin{aligned}
&E[X\phi_1(R)/\alpha_1 - 1 - \log\{X\phi_1(R)/\alpha_1\}] - E[X/\alpha_1 - 1 - \log(X/\alpha_1)] \\
&= E[-WR\{1-\phi_1(R)\}/\alpha_1 - \log\phi_1(R)] = E[-R\{1-\phi_1(R)\}(\alpha_1+\alpha_2)/\alpha_1 - \log\phi_1(R)],
\end{aligned}$$

which is negative since $(\alpha_1+\alpha_2)/\alpha_1 \geq 2$ by assumption and since $2R\{1-\phi_1(R)\}+\log\phi_1(R) > 0$ when $\phi_1(R) \neq 1$. For part (ii), let $\phi_2(R) = \max\{1, 1/\{2(1-R)\}\}$. Then

$$
\begin{aligned}
&E[Y\phi_2(R)/\alpha_2 - 1 - \log\{Y\phi_2(R)/\alpha_2\}] - E[Y/\alpha_2 - 1 - \log(Y/\alpha_2)] \\
&= E[W(1-R)\{\phi_2(R)-1\}/\alpha_2 - \log\phi_2(R)] = E[(1-R)\{\phi_2(R)-1\}(\alpha_1+\alpha_2)/\alpha_2 - \log\phi_2(R)],
\end{aligned}
$$

which is negative since $(\alpha_1+\alpha_2)/\alpha_1 \leq 2$ and since $2(1-R)\{\phi_2(R)-1\} - \log\phi_2(R) < 0$ when $\phi_2(R) \neq 1$. □

6.2 Details of the approximation algorithm used in Section 4.5

First, the density proportional to

$$
\Big(\frac{1}{2}\Big)^{\mu} \sqrt{\sum_{k=0}^{\infty} \mathrm{NB}(k|\mu, 1/2)\{\psi_1(\mu) - \psi_1(\mu+k)\}}, \quad \mu \in (0,\infty),
$$

is approximated by $\mathrm{Ga}(1/2, \log 2)$ because for all $k \in \mathbb{N}_0$ and all $\mu \in (0,\infty)$,

$$
\Big\{\psi_1(\mu) - \psi_1(\mu+k), \frac{1}{\mu}\frac{k}{\mu+k-1}\Big\} \subset \Big[\frac{1}{\mu}\frac{k}{\mu+k}, \frac{1}{\mu-1}\frac{k}{\mu+k-1}\Big],
$$

provided that $\mu > 1$, and because

$$
\sqrt{\sum_{k=0}^{\infty} \mathrm{NB}(k|\mu, 1/2)\frac{1}{\mu}\frac{k}{\mu+k-1}} = \sqrt{\frac{1}{\mu}\frac{1}{2}} \propto \mu^{1/2-1}
$$

for all $\mu \in (0,\infty)$ (see Hudson (1978) for the inequality). Next, note that for all $K \in \mathbb{N}$ and all $\mu \in (0,\infty)$,

$$
\begin{aligned}
&\sqrt{\sum_{k=0}^{K} \mathrm{NB}(k|\mu, 1/2)\{\psi_1(\mu) - \psi_1(\mu+k)\}} \leq \pi^{\mathrm{J}}(\mu) \\
&\leq \sqrt{\psi_1(\mu) + \sum_{k=0}^{K} \mathrm{NB}(k|\mu, 1/2)\{-\psi_1(\mu+k)\} + P(\widetilde{X} > K)\{-\psi_1(\mu + E[\widetilde{X}|\widetilde{X} > K])\}}
\end{aligned}
$$

by Jensen's inequality, where $\widetilde{X} \sim \mathrm{NB}(\mu, 1/2)$ and where $E[\widetilde{X}|\widetilde{X} > K]$ is expressible in closed form (see Shonkwiler (2016)). Then, based on these inequalities, the acceptance ratio can be bounded below and above with arbitrary accuracy in MH steps (such an approach is used, for example, by Polson, Scott and Windle (2013)).

Acknowledgments

We would like to thank the editor, the associate editor, and the two reviewers for many valuable comments and helpful suggestions which improved the paper. Research of the authors was supported in part by Grant-in-Aid for Scientific Research (20J10427, 18K11188) from Japan Society for the Promotion of Science.

References

[1] Bagnoli, M. and Bergstrom, T. (2005). Log-concave probability and its applications. *Economic Theory*, **26**, 445–469.

[2] Bhattacharya, C.G. (1984). Two inequalities with an application. *Annals of the Institute of Statistical Mathematics*, **36**, 129–134.

[3] Blumenthal, S. and Cohen, A. (1968a). Estimation of the larger translation parameter. *The Annals of Mathematical Statistics*, **39**, 502–516.

[4] Blumenthal, S. and Cohen, A. (1968b). Estimation of two ordered translation parameters. *The Annals of Mathematical Statistics*, **39**, 517–530.

[5] Chang, Y.-T. and Shinozaki, N. (2006). Estimation of ordered means of two poisson distributions. *Communications in Statistics - Theory and Methods*, **35**, 1993–2003.

[6] Cohen, A. and Sackrowitz, H.B. (1970). Estimation of the last mean of a monotone sequence. *The Annals of Mathematical Statistics*, **41**, 2021–2034.

[7] Garg, N. and Misra, N. (2021). Componentwise equivariant estimation of order restricted location and scale parameters in bivariate models: A unified study. *arXiv preprint arXiv:2109.14997*.

[8] Hudson, H.M. (1978). A natural identity for exponential families with applications in multiparameter estimation. Ann. Statist., 6, 473–484.

[9] Hwang, J.T. and Peddada, S.D. (1994). Confidence interval estimation subject to order restrictions. *The Annals of Statistics*, **22**, 67–93.

[10] Jena, A.K. and Tripathy, M.R. (2019). Estimating ordered quantiles of two exponential populations with a common minimum guarantee time. *Communications in Statistics - Theory and Methods*, **48**, 3570–3585.

[11] Katz, M.W. (1963). Estimating ordered probabilities. *The Annals of Mathematical Statistics*, **34**, 967–972.

[12] Kelly, R.E. (1989). Stochastic reduction of loss in estimating normal means by isotonic regression. *The Annals of Statistics*, **17**, 937–940.

[13] Kubokawa, T. and Saleh, A.K.Md.E. (1994). Estimation of location and scale parameters under order restrictions. *Journal of Statistical Research*, **28**, 41–51.

[14] Kumar, S., Kumar, A. and Tripathi, Y.M. (2005). A Note on the Pitman estimator of ordered normal means when the variances are unequal. *Communications in Statistics - Theory and Methods*, **34**, 2115–2122.

[15] Kumar, S. and Sharma, D. (1988). Simultaneous estimation of ordered parameters. *Communications in Statistics - Theory and Methods*, **17**, 4315–4336.

[16] Kumar, S. and Sharma, D. (1989). On the pitman estimator op ordered normal means. *Communications in Statistics - Theory and Methods*, **18**, 4163–4175.

[17] Kumar, S. and Sharma, D. (1993). Minimaxity of the Pitman estimator of ordered normal means when variances are unequal. *Journal of the Indian Society of Agricultural Statistics*, **45**, 230–234.

[18] Kumar, S., Tripathi, Y.M. and Misra, N. (2005). James–Stein type estimators for ordered normal means. *Journal of Statistical Computation and Simulation*, **75**, 501–511.

[19] Kushary, D. and Cohen, A. (1991). Estimation of ordered Poisson parameters. *Sankhyā: The Indian Journal of Statistics, Series A*, **53**, 334–356.

[20] Lee, C.I.C. (1981). The quadratic loss of isotonic regression under normality. *The Annals of Statistics*, **9**, 686–688.

[21] Lee, C.I.C. (1988). Quadratic Loss of order restricted estimators for treatment means with a control. *The Annals of Statistics*, **16**, 751–758.

[22] Misra, N., Choudhary, P.K., Dhariyal, I.D. and Kundu, D. (2002). Smooth estimators for estimating order restricted scale parameters of two gamma distributions. *Metrika*, **56**, 143–161.

[23] Misra, N., Iyer, S.K. and Singh, H. (2004). The linex risk of maximum likelihood estimators of parameters of normal populations having order restricted means. *Sankhyā: The Indian Journal of Statistics*, **66**, 652–677.

[24] Patra, L.K. and Kumar, S. (2017). Estimating ordered means of a bivariate normal distribution. *American Journal of Mathematical and Management Sciences*, **36**, 118–136.

[25] Patra, K.L., Kumar, S. and Petropoulos, C. (2021). Componentwise estimation of ordered scale parameters of two exponential distributions under a general class of loss function. *Statistics*, **55**, 595–617.

[26] Polson, N.G., Scott, J.G. and Windle, J. (2013). Bayesian inference for logistic models using pólya-gamma latent variables. *Journal of the American Statistical Association*, **108**, 1339–1349.

[27] Shonkwiler, J.S. (2016). Variance of the truncated negative binomial distribution. *Journal of Econometrics*, **195**, 209–210.

[28] Singh, H., Misra, N. and Li, S. (2005). Estimation of order restricted concentration parameters of von Mises distributions. *Communications in Statistics - Simulation and Computation*, **34**, 21–40.

[29] Stein, C. (1964). Inadmissibility of the usual estimator for the variance of a normal distribution with unknown mean. *Annals of the Institute of Statistical Mathematics*, **16**, 155–160.

[30] van Eeden, C. (2006). Restricted parameter space problems - Admissibility and minimaxity properties. Lecture Notes in Statistics, 188. New York: Springer.

[31] Vijayasree, G., Misra, N. and Singh, H. (1995). Componentwise estimation of ordered parameters of $k(\geq 2)$ exponential populations. *Annals of the Institute of Statistical Mathematics*, **47**, 287–307.

[32] Vijayasree, G. and Singh, H. (1991). Simultaneous estimation of two ordered exponential parameters. *Communications in Statistics - Theory and Methods*, **20**, 2559–2576.