APPLICATIONES MATHEMATICAE
40,4 (2013), pp. 393–404

Jarosław Bartoszewicz
Piotr Bolesław Nowak (Opole)

MONOTONICITY OF BAYES ESTIMATORS

Abstract. Let $X = (X_1, \dots, X_n)$ be a sample from a distribution with density $f(x;\theta)$, $\theta \in \Theta \subset \mathbb{R}$. In this article the Bayesian estimation of the parameter θ is considered. We examine whether the Bayes estimators of θ are pointwise ordered when the prior distributions are partially ordered. Various cases of loss function are studied. A lower bound for the survival function of the normal distribution is obtained.

1. Preliminaries. Let X and Y be real valued random variables with distribution functions F and G, respectively, and density functions f and g, if they exist. Denote by F^{-1} the quantile function, by $\bar{F}$ the survival function, by S_X the support, and by l_X and u_X the left and right endpoints of the support of X, respectively.

We say that X is smaller than Y:

- in the (usual) *stochastic order* ($X \le_{\mathrm{st}} Y$) if $F(x) \ge G(x)$ for all $x \in \mathbb{R}$;
- in the *hazard rate order* ($X \le_{\mathrm{hr}} Y$) if $\bar{G}(x)/\bar{F}(x)$ is increasing in $x \in (-\infty, \max(u_X, u_Y))$;
- in the *reverse hazard rate order* ($X \le_{\mathrm{rh}} Y$) if $G(x)/F(x)$ is increasing in $x \in (\min(l_X, l_Y), \infty)$;
- in the *likelihood ratio order* ($X \le_{\mathrm{lr}} Y$) if $g(x)/f(x)$ is increasing in $x \in S_X \cup S_Y$.

We say that X and Y are *pointwise ordered* ($X \le Y$) if X and Y are defined on the same probability space $(\Omega, \mathcal{F}, P)$ and $X(\omega) \le Y(\omega)$ for all $\omega \in \Omega$.

2010 *Mathematics Subject Classification*: 60E15, 62F15.
Key words and phrases: Bayes estimation, maximum a posteriori estimation, exponential family, weighted distribution, stochastic ordering, pointwise domination, total positivity, unimodality, squared error loss, weighted squared error loss, uniform loss function, asymmetric loss function.

DOI: 10.4064/am40-4-1

The following relations among the above defined stochastic orders are well known:

$$\begin{array}{ccccc} X \leq_{\mathrm{lr}} Y & \Rightarrow & X \leq_{\mathrm{hr}} Y & & \\ \Downarrow & & \Downarrow & & \\ X \leq_{\mathrm{rh}} Y & \Rightarrow & X \leq_{\mathrm{st}} Y & \Leftarrow & X \leq Y \end{array}$$

We also write $F \leq_{\mathrm{S}} G$ instead of $X \leq_{\mathrm{S}} Y$, where S is some stochastic order.

Recall the following definition of a weighted distribution.

Definition 1.1. Let X be a random variable with distribution function F and let $w : \mathbb{R} \to \mathbb{R}_+$ be a function for which $0 < E[w(X)] < \infty$. The distribution function

$$F_w(x) = \frac{1}{E[w(X)]} \int_{-\infty}^{x} w(u)\, dF(u) \tag{1.1}$$

is called the *weighted distribution* related to F with weight function w. A random variable X_w with distribution F_w is called a *weighted version* of X.

Now we discuss the relationship between weighted distributions and Bayesian estimation.

Let X be a random variable with density $f(x;\theta)$, $\theta \in \Theta$, absolutely continuous with respect to a σ-finite measure μ. We assume that θ is a value of a random variable ϑ with distribution function G and density g absolutely continuous with respect to ν. This distribution is called the *prior distribution* of θ. Denote by $f(\theta \mid x)$ and $F(\theta \mid x)$ the density and distribution function of the conditional distribution of ϑ given $X = x$, respectively.

From the Bayes formula it follows that

$$f(\theta \mid x) = \frac{f(x;\theta)g(\theta)}{m(x)}, \tag{1.2}$$

where $m(x) = E[f(x;\vartheta)]$ is the density of the marginal distribution of X with respect to the measure μ.

From the general theory of Bayesian estimation it is known (see Ferguson [4]) that the Bayes estimator of θ under squared error loss is given by

$$\hat{\theta}(x) = \int_{\Theta} \theta f(\theta \mid x)\, d\nu(\theta). \tag{1.3}$$

Thus, this estimator is the mean of the posterior distribution. In what follows we also consider Bayesian estimation under weighted squared error loss, uniform loss and LINEX loss function.

We assume that all densities considered are absolutely continuous with respect to the counting measure or Lebesgue measure. Assume, without loss of generality, that X is a sufficient statistic for the family $\{P_\theta,\ \theta \in \Theta\}$. Suppose that G_1 and G_2 (the prior distribution functions of θ) have density functions g_1 and g_2, respectively. Denote by $\hat\theta_1$ and $\hat\theta_2$ the Bayes estimators of θ if the prior distributions are G_1 and G_2, respectively.

2. Main results. Our first aim is to examine the stochastic ordering of estimators of the form (1.3) when the prior distributions are also ordered.

We show that under some assumptions about the weight function and prior distributions the Bayes estimators are stochastically ordered. In fact, we show that they are pointwise ordered, i.e. $\hat\theta_1(x) \le \hat\theta_2(x)$ for all x (we always assume that x belongs to the set of values of X).

Assume now that we estimate under squared error loss.

THEOREM 2.1. *If $G_1 \le_{\rm lr} G_2$, then the corresponding Bayes estimators of θ are pointwise ordered.*

Proof. First we prove that $F_1(\cdot \mid x) \le_{\rm lr} F_2(\cdot \mid x)$ for all x. The corresponding densities of the posterior distributions are

$$f_i(\theta \mid x) = \frac{f(x;\theta)g_i(\theta)}{m_i(x)}, \qquad i = 1,2.$$

From the assumption it follows that $g_2(\theta)/g_1(\theta)$ is increasing in θ. Hence, $f_2(\theta \mid x)/f_1(\theta \mid x)$ is also increasing in θ for all x. Thus, we have shown that the posterior distributions are ordered with respect to the likelihood ratio order and hence they are also ordered with respect to the usual stochastic order.

We know that stochastically ordered distributions have ordered means, provided they exist. Hence $\hat\theta_1(x) \le \hat\theta_2(x)$ for all x. ∎

From Theorem 2.1 it follows that in order to obtain stochastic ordering of Bayes estimators it suffices to order posterior distributions with respect to the usual stochastic order. On the other hand, the assumption of that theorem is so strong that ordering of estimators holds for any weight function $f(x;\theta)$. Below, we relax the strong assumption that the prior distributions are ordered with respect to the likelihood ratio order. First we recall the following result, due to Bartoszewicz and Skolimowska [3].

LEMMA 2.2.

(a) *Let w be an increasing function. If $X \le_{\rm hr} Y$, then $X_w \le_{\rm hr} Y_w$.*
(b) *Let w be a decreasing function. If $X \le_{\rm rh} Y$, then $X_w \le_{\rm rh} Y_w$.*

The crucial observation in our study is the following corollary.

COROLLARY 2.3. *The posterior density* (1.2) *is the density of the weighted version of ϑ with weight function $f(x;\theta)$.*

Treating the problem of ordering of posterior distributions as a problem of ordering of weighted distributions one can formulate the following theorem which is a direct consequence of Lemma 2.2.

Theorem 2.4.

(a) *If $f(x;\theta)$ is increasing in θ for all x and $G_1 \le_{\mathrm{hr}} G_2$, then the corresponding Bayes estimators of θ are pointwise ordered.*
(b) *If $f(x;\theta)$ is decreasing in θ for all x and $G_1 \le_{\mathrm{rh}} G_2$, then the corresponding Bayes estimators of θ are pointwise ordered.*

Observe that the likelihood function as a weighted function has often a local maximum so it is not monotonic. Therefore Theorem 2.4 has limited applications.

Now we give two special cases when we can use Theorem 2.4, because the likelihood functions are in fact monotonic in θ.

Case (a): *Shifted exponential distribution with density* $f(x;\theta) = e^{-(x-\theta)}\mathbf{1}_{(\theta,\infty)}(x)$, $\theta \in \mathbb{R}$. Then the likelihood function L is

$$L(x_1,\dots,x_n;\theta) = \prod_{i=1}^{n} f(x_i;\theta) = \exp\Big(n\theta - \sum_{i=1}^{n} x_i\Big)\mathbf{1}\{\theta < x_{1:n}\}.$$

Thus it is increasing in $\theta \in (-\infty, x_{1:n})$.

Case (b): *Uniform distribution on* $(0,\theta)$, $\theta > 0$. Then the likelihood function L is

$$L(x_1,\dots,x_n;\theta) = \prod_{i=1}^{n} f(x_i;\theta) = \frac{1}{\theta^n}\prod_{i=1}^{n}\mathbf{1}_{(x_i,\infty)}(\theta) = \frac{1}{\theta^n}\mathbf{1}\{\theta > x_{n:n}\}.$$

Thus it is decreasing in $\theta \in (x_{n:n}, \infty)$.

Using this method, based on weighted distributions, we cannot obtain ordering of Bayes estimators if we only assume that the prior distributions are ordered in the usual stochastic order. This situation is described in the following example.

Example 2.5. Let X be a random variable with uniform distribution $U(0;\theta)$, where $\theta \in \{1,2,3\}$. Assume that ϑ has the following distribution:

$$P_\lambda(\vartheta = 1) = p_1 = \tfrac{1}{3} - \tfrac{1}{3}\lambda,\quad P_\lambda(\vartheta = 2) = p_2 = \tfrac{1}{3} + \tfrac{1}{3}\lambda,\quad P_\lambda(\vartheta = 3) = p_3 = \tfrac{1}{3},$$

where $\lambda \in (0,1)$. It is easy to see that the family of distributions $\{P_\lambda, \lambda \in (0,1)\}$ is stochastically increasing in θ but is not ordered with respect to the hazard rate order. Indeed, for example

$$\frac{\bar F(0;\lambda_2)}{\bar F(0;\lambda_1)} = 1, \qquad \frac{\bar F(1;\lambda_2)}{\bar F(1;\lambda_1)} = \frac{\frac{2}{3} + \frac{1}{3}\lambda_2}{\frac{2}{3} + \frac{1}{3}\lambda_1}, \qquad \frac{\bar F(2;\lambda_2)}{\bar F(2;\lambda_1)} = 1.$$

Thus, the function $\bar F(t;\lambda_2)/\bar F(t;\lambda_1)$ is not monotonic for $t \in (-\infty, 3)$ when $\lambda_1 < \lambda_2$.

First we determine the marginal distribution of the random variable X:

$$m(x) = p_1\mathbf{1}_{(0,1)}(x) + \tfrac{1}{2}p_2\mathbf{1}_{(0,2)}(x) + \tfrac{1}{3}p_3\mathbf{1}_{(0,3)}(x)$$
$$= \begin{cases} p_1 + \frac{1}{2}p_2 + \frac{1}{3}p_3 & \text{when } x \in (0,1), \\ \frac{1}{2}p_2 + \frac{1}{3}p_3 & \text{when } x \in (1,2), \\ \frac{1}{3}p_3 & \text{when } x \in (2,3). \end{cases}$$

The posterior distribution is

$$P(\vartheta = i \mid X = x) = \frac{\frac{1}{i}p_i\mathbf{1}_{(0,i)}(x)}{m(x)}, \qquad i \in \{1,2,3\}.$$

Thus, the Bayes estimator of θ under squared error loss is

$$\hat{\theta}(x) = \frac{1}{m(x)}\big(p_1\mathbf{1}_{(0,1)}(x) + p_2\mathbf{1}_{(0,2)}(x) + p_3\mathbf{1}_{(0,3)}(x)\big)$$
$$= \begin{cases} \dfrac{18}{11 - 3\lambda} & \text{when } x \in (0,1), \\ \dfrac{12 + 6\lambda}{5 + 3\lambda} & \text{when } x \in (1,2), \\ 3 & \text{when } x \in (2,3). \end{cases}$$

Now let $\lambda_1 < \lambda_2$, $\lambda_1, \lambda_2 \in (0,1)$. Thus the inequality $\hat{\theta}_1(x) \le \hat{\theta}_2(x)$ does not hold for all $x \in (0,3)$.

Now we consider the ordering of the Bayes estimators in a particular case, when the density f belongs to a one-parameter exponential family, i.e. $f(x;\theta) = c(\theta)h(x)\exp(\theta x)$. Thus the Bayes estimator of θ under squared error loss is (see Lehmann and Casella [6])

$$\hat{\theta}(x) = \frac{\partial}{\partial x}\log m(x) - \frac{\partial}{\partial x}\log h(x). \tag{2.1}$$

Hence, the inequality $\hat{\theta}_1(x) \le \hat{\theta}_2(x)$ for all x is equivalent to $m_2(x)/m_1(x)$ being increasing in x, i.e. $m_i(x)$, $i \in \{1,2\}$, is TP_2 in (i,x) (for definition of TP_2 see Karlin [5]).

Corollary 2.6. *If* $m_i(x) = \int_\Theta f(x;\theta)g_i(\theta)\,d\theta$ *is* TP_2, $i \in \{1,2\}$, *then* $\hat{\theta}_1(x) \le \hat{\theta}_2(x)$ *for all* x, *where the Bayes estimator of* θ *is given by* (2.1).

In particular, when $G_1 \le_{\mathrm{lr}} G_2$, then $m_i(x)$ is TP_2, $i \in \{1,2\}$. It follows directly from the Basic Composition Formula (see also Karlin [5]) since both $f(x;\theta)$ and $g_i(\theta)$ are TP_2.

Until now we have only considered stochastic ordering of the Bayes estimators under squared error loss. The following theorem is known (see Ferguson [4]).

THEOREM 2.7. *If the loss function is*

$$L(\theta, a) = \begin{cases} (\theta - a)k_0 & \text{if } \theta - a \geq 0, \\ (a - \theta)k_1 & \text{if } \theta - a < 0, \end{cases} \tag{2.2}$$

where k_0 and k_1 are positive constants, then the Bayes estimator of θ is the quantile of order $k_0/(k_0 + k_1)$ of the posterior distribution.

It is well known that if $X \leq_{\mathrm{st}} Y$, then $F_X^{-1}(p) \leq F_Y^{-1}(p)$ for all $p \in (0, 1)$. Hence, we have the following conclusion.

COROLLARY 2.8. *Theorems* 2.1 *and* 2.4 *remain valid when $\hat{\theta}$ is the Bayes estimator under the loss function* (2.2).

Now we consider the following generalization of the squared error loss:

$$L(a, \theta) = \chi(\theta)[\gamma(\theta) - a]^2, \qquad \chi(\theta) > 0. \tag{2.3}$$

It is well known (see e.g. Ferguson [4]) that the Bayes estimator under weighted squared error loss is

$$\hat{\theta}(x) = \frac{E[\chi(\theta)\gamma(\theta) \mid x]}{E[\chi(\theta) \mid x]} \tag{2.4}$$

provided that all posterior expectations in (2.4) exist and $E[\chi(\theta) \mid x] \neq 0$.

Now we prove an auxiliary lemma.

LEMMA 2.9. *Let f_1, f_2, g_1, g_2 be nonnegative functions such that f_2/f_1 and g_2/g_1 are increasing functions. Then*

$$\frac{\int f_1(x)g_1(x)\,dx}{\int f_2(x)g_1(x)\,dx} \geq \frac{\int f_1(x)g_2(x)\,dx}{\int f_2(x)g_2(x)\,dx},$$

provided all integrals exist.

Proof. Monotonicity of f_2/f_1 is equivalent to the kernel $f_i(x)$ being TP_2 in $(i, x) \in \{1, 2\} \times \mathbb{R}$. Similarly, $g_j(x)$ is TP_2 in $(j, x) \in \{1, 2\} \times \mathbb{R}$. By the Basic Composition Formula the integral $\int f_i(x)g_j(x)\,dx$ is TP_2 in $(i, j) \in \{1, 2\} \times \{1, 2\}$, which is equivalent to the assertion of the lemma. ∎

Now we give a result on ordering of Bayes estimators under the loss function (2.3).

THEOREM 2.10. *If $G_1 \leq_{\mathrm{lr}} G_2$ and the function γ in* (2.3) *is increasing and nonnegative on Θ, then the corresponding Bayes estimators of θ under the loss function* (2.3) *are pointwise ordered.*

Proof. From the assumptions it follows that the posterior distributions are ordered with respect to the likelihood ratio order. Hence the conclusion follows immediately from Lemma 2.9 and formula (2.4). ∎

Before we consider the comparison of Bayes estimators under the uniform loss function we recall the definition of unimodality. In further considerations we will work only with distributions having densities with respect to Lebesgue measure.

DEFINITION 2.11. A function $f : [a,b] \to \mathbb{R}$ is *unimodal* with *mode* $M \in [a,b]$ if f is increasing on $[a,M]$ and decreasing on $[M,b]$. We say that f is *strictly unimodal* if it has a single mode.

DEFINITION 2.12. We say that a random variable X is (*strictly*) *unimodal* if its density is (strictly) unimodal on the support S_X.

The following lemma gives a relation between the modes of two strictly unimodal random variables which are ordered in the likelihood ratio order.

LEMMA 2.13. *Let X and Y be strictly unimodal random variables with densities f and g, respectively. Let M_1 and M_2 be the modes of X and Y, respectively. If $X \le_{\mathrm{lr}} Y$, then $M_1 \le M_2$.*

Proof. Suppose, on the contrary, that $M_1 > M_2$. Observe that

$$g(M_1)f(M_2) \ge g(M_2)f(M_1) > g(M_2)f(M_2).$$

The first inequality follows since $X \le_{\mathrm{lr}} Y$. The second inequality follows from $f(M_1) > f(M_2)$, since M_1 is the mode of X. We infer that $g(M_1) > g(M_2)$, which is impossible, since M_2 is the mode of Y. ∎

EXAMPLE 2.14. The condition $X \le_{\mathrm{lr}} Y$ in Lemma 2.13 cannot be relaxed. Let X be a random variable with distribution function

$$G_1(\theta) = \begin{cases} 1 - e^{-\theta^2}, & \theta > 0, \\ 0, & \theta \le 0, \end{cases} \tag{2.5}$$

and let Y be a random variable with distribution function

$$G_2(\theta) = \begin{cases} 1 - \frac{1}{2}e^{-3\theta}(9\theta^2 + 6\theta + 2), & \theta > 0, \\ 0, & \theta \le 0. \end{cases} \tag{2.6}$$

One can verify that $X \le_{\mathrm{hr}} Y$ and $X \le_{\mathrm{rh}} Y$, but $X \not\le_{\mathrm{lr}} Y$. We omit the tedious and elementary proof. Moreover, the mode of X equals $1/\sqrt{2}$, whereas the mode of Y is $2/3$.

Now we consider the Bayesian estimation of θ under the uniform loss function, i.e.

$$L_\varepsilon(\theta, a) = \begin{cases} 0, & |\theta - a| \le \varepsilon, \\ 1, & |\theta - a| > \varepsilon, \end{cases} \tag{2.7}$$

where $\varepsilon > 0$.

It is easy to check that in this case

$$\hat{\theta}(x) = \arg\min_{a\in\Theta} \int_{a-\varepsilon}^{a+\varepsilon} f(\theta \mid x)\, d\theta \tag{2.8}$$

gives the Bayes estimator for θ. When the posterior density $f(\theta \mid x)$ is unimodal and symmetric about M, then $\hat{\theta}(x) = M$. In this simple case we immediately have the following lemma.

Lemma 2.15. *If $G_1 \le_{\mathrm{lr}} G_2$ and the corresponding posterior distributions are symmetric and unimodal, then the corresponding Bayes estimators of θ under the loss function* (2.7) *are pointwise ordered.*

Using Exercise 4.10 of Shao [7] we have a generalization of Lemma 2.15.

Corollary 2.16. *Under the assumptions of Lemma* 2.15 *the inequality $\hat{\theta}_1(x) \le \hat{\theta}_2(x)$ holds for all x if the loss function is of the form $L(|\theta - a|)$, where L is an increasing function on* $[0, \infty]$.

Note that for small ε,

$$\frac{1}{2\varepsilon} \int_{a-\varepsilon}^{a+\varepsilon} f(\theta \mid x)\, d\theta \approx f(a \mid x).$$

Thus, in the limit $\varepsilon \to 0$ the Bayes estimator under the uniform loss function (2.7) is

$$\hat{\theta}(x) = \arg\max_{a\in\Theta} f(a \mid x). \tag{2.9}$$

This estimator is also called the MAP (*maximum a posteriori*) estimator.

Now we investigate the case when the posterior density is a unimodal function. Then the problem of comparison of MAP estimators is equivalent to comparison of modes of the posterior distributions. If we combine previous facts with Lemma 2.13 we obtain the following result about comparison of MAP estimators.

Theorem 2.17. *If $G_1 \le_{\mathrm{lr}} G_2$ and $f_i(\theta \mid x)$, $i = 1, 2$, are strictly unimodal, then the corresponding MAP estimators of θ are pointwise ordered.*

Now we recall some well known sufficient conditions for unimodality.

Remark 2.18. Every logconcave density f is unimodal (see e.g. Barlow and Proschan [2]) and the product of logconcave functions is also logconcave.

Finally we consider estimation under the asymmetric loss function LINEX defined by

$$L(\theta, \delta) = b\{e^{a(\theta-\delta)} - a(\theta - \delta) - 1\}, \qquad a \ne 0,\ b > 0.$$

The Bayes estimator of θ under the LINEX loss function is (see Zellner [8])

$$\delta(x) = (1/a) \log E(e^{a\theta} \mid x). \tag{2.10}$$

From (2.10) it follows that in order to compare Bayes estimators under the LINEX loss function it suffices to compare the moment generating functions of the posterior distributions. In the following theorem we give a simple sufficient condition for ordering of Bayes estimators under the LINEX loss function.

THEOREM 2.19. *If $G_1 \le_{\mathrm{lr}} G_2$, then the corresponding Bayes estimators of θ under the* LINEX *loss function are pointwise ordered.*

3. Examples of applications. Now we illustrate our results by examples.

EXAMPLE 3.1. Let $\mathbf{X} = (X_1, \ldots, X_n)$ be a random sample of binary random variables with $P(X_1 = 1) = \theta \in (0,1)$. By the factorization theorem, $T = \sum_{i=1}^n X_i$ is a sufficient statistic for θ.

We consider stochastic ordering of the Bayes estimators of θ assuming that the prior has noncentral beta distribution with density

$$g(\theta;\delta,\alpha,\beta) = \sum_{j=0}^{\infty} \frac{e^{-\delta}\delta^j}{j!}\,\frac{\theta^{\alpha+j-1}(1-\theta)^{\beta-1}}{B(\alpha+j,\beta)}, \qquad \delta,\alpha,\beta > 0,\ \theta \in (0,1). \tag{3.1}$$

Let α and β be fixed and let $0 < \delta_1 < \delta_2$. First we show that $G_1 \le_{\mathrm{lr}} G_2$. Note that this is equivalent to (3.1) being TP_2 in (θ,δ). Notice that (3.1) can be expressed as

$$g(\theta;\delta,\alpha,\beta) = \sum_{j=0}^{\infty} k(\delta,j)\,l(j,\theta;\alpha,\beta), \tag{3.2}$$

where

$$k(\delta,j) = \frac{e^{-\delta}\delta^j}{j!}, \qquad \delta > 0,\ j = 0,1,\ldots,$$
$$l(j,\theta;\alpha,\beta) = \frac{\theta^{\alpha+j-1}(1-\theta)^{\beta-1}}{B(\alpha+j,\beta)}, \qquad \alpha,\beta > 0,\ \theta \in (0,1).$$

It is obvious that $k(\delta,j)$ is TP_2 and $l(j,\theta;\alpha,\beta)$ is TP_2 in (j,θ). Therefore, by the Basic Composition Formula, $g(\theta;\delta,\alpha,\beta)$ is TP_2 in (θ,δ). Using Theorem 2.1 we have $\hat\theta_1(t) \le \hat\theta_2(t)$ for all t, where t is a value of T.

Now we derive the Bayes estimator of θ. Firstly we find the marginal distribution of T:

$$m(t) = \int_0^1 f(t,\theta)g(\theta)\,d\theta = \binom{n}{t}\sum_{k=0}^{\infty} \frac{e^{-\delta}\delta^k}{k!}\,\frac{B(\alpha+t+k,\beta+n-t)}{B(\alpha+k,\beta)}.$$

Hence the Bayes estimator of θ is

$$\hat{\theta}(t) = \int_0^1 \theta \frac{f(t,\theta)g(\theta)}{m(t)}\, d\theta = \frac{\sum_{k=0}^{\infty} \frac{\delta^k}{k!} \frac{B(\alpha+t+k+1, \beta+n-t)}{B(\alpha+k,\beta)}}{\sum_{k=0}^{\infty} \frac{\delta^k}{k!} \frac{B(\alpha+t+k, \beta+n-t)}{B(\alpha+k,\beta)}} = \frac{\alpha+t}{\alpha+\beta+n} \frac{{}_2F_2(\alpha+\beta, \alpha+t+1; \alpha, \alpha+\beta+n+1; \delta)}{{}_2F_2(\alpha+\beta, \alpha+t; \alpha, \alpha+\beta+n; \delta)}, \tag{3.3}$$

where ${}_2F_2$ denotes the generalized hypergeometric function (see Andrews et al. [1]). Observe that if $\delta \to 0$, then (3.3) tends to $(\alpha+t)/(\alpha+\beta+n)$. On the other hand, if $\delta \to \infty$, then (3.3) tends monotonically to 1.

Now we apply Theorem 2.4 to obtain a lower bound for a survival function of normal distribution.

Example 3.2. Let X be a random variable having uniform distribution $U(0,\theta)$, $\theta > 0$. Consider the estimation of θ under squared error loss. Assume that ϑ_1 and ϑ_2 are random variables with distribution functions given in (2.5) and (2.6), respectively. Recall that $G_1 \le_{\mathrm{rh}} G_2$ and $G_1 \not\le_{\mathrm{lr}} G_2$.

The Bayes estimator of the parameter θ with respect to the prior having distribution function G_1 is given by

$$\hat{\theta}_1(x) = \frac{\int_x^\infty \theta e^{-\theta^2}\, d\theta}{\int_x^\infty e^{-\theta^2}\, d\theta} = \frac{\frac{1}{2} e^{-x^2}}{\int_x^\infty e^{-\theta^2}\, d\theta}.$$

In turn, the Bayes estimator of the parameter θ with respect to the prior having distribution function G_2 is given by

$$\hat{\theta}_2(x) = \frac{\int_x^\infty \theta^2 e^{-3\theta}\, d\theta}{\int_x^\infty \theta e^{-3\theta}\, d\theta} = \frac{9x^2+6x+2}{3(3x+1)}.$$

By Theorem 2.4(b) we have $\hat{\theta}_1(x) \le \hat{\theta}_2(x)$, $x > 0$. This inequality can be equivalently written as

$$\int_x^\infty e^{-\theta^2}\, d\theta \ge \frac{3}{2} e^{-x^2} \frac{3x+1}{9x^2+6x+2}, \qquad x > 0.$$

Making the substitution $x := \sqrt{2}\, x$, after easy modifications we have

$$1 - \Phi(x) \ge \frac{3}{2\sqrt{\pi}} e^{-x^2/2} \frac{3\sqrt{2}x+2}{9x^2+6\sqrt{2}x+4}, \qquad x > 0, \tag{3.4}$$

where Φ is the distribution function of the standard normal distribution. In fact, the inequality in (3.4) is sharp.

Let us denote the right side of inequality (3.4) by $\tilde{\Phi}(x)$. By L'Hopital's rule, we can show that $\tilde{\Phi}$ and Φ are asymptotically equivalent, i.e. $\lim_{x\to\infty} \tilde{\Phi}(x)/\Phi(x) = 1$.

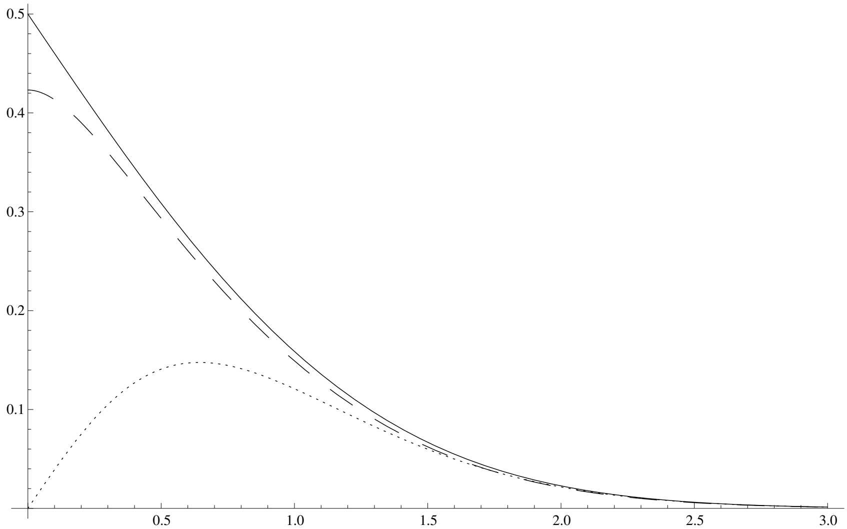

Fig. 1. Solid line: survival function of the normal distribution; dotted line: classical lower bound; dashed line: lower bound given in (3.4).

The lower bound obtained in (3.4) is better than the classical lower bound $\frac{1}{\sqrt{2\pi}}\frac{x}{x^2+1}e^{-x^2/2}$ for $0 < x < \frac{1}{12}(5\sqrt{2}+\sqrt{194}) = 1.749\ldots$ (see Figure 1).

Example 3.3. Let ϑ_1 be a random variable having normal distribution $N(0,1)$ and let ϑ_2 be a random variable with density $e^{-\theta}\exp\{-e^{-\theta}\}$, $\theta \in \mathbb{R}$. We check directly that $\vartheta_1 \le_{\mathrm{lr}} \vartheta_2$. Now let us consider a random variable X having normal distribution $N(\theta,1)$, $\theta \in \mathbb{R}$ and MAP estimator of θ. The posterior density of ϑ_1 is

$$f_1(\theta \,|\, x) = \frac{1}{\sqrt{\pi}}\exp\{-(\theta - x/2)^2\}$$

and the posterior density of ϑ_2 is

$$f_2(\theta \,|\, x) \propto f(x \,|\, \theta)g_2(\theta) \propto \exp\{(x-\theta)^2/2 - \theta - e^{-\theta}\}.$$

Note also that both densities are unimodal since they are products of logconcave functions.

It is clear that $\hat{\theta}_1(x) = x/2$ and $\hat{\theta}_2(x) = a$, where a is the solution of the equation $e^{-a} - a = 1 - x$. As a corollary we obtain the inequality $a \ge x/2$, which holds for any $x \in \mathbb{R}$. Of course, this inequality can be easily proved by methods of elementary calculus.

Example 3.4. Consider the classical example of estimating the mean of the normal distribution $N(\theta,\sigma^2)$, where σ is known. Assume that the prior is normal $N(\mu,\eta^2)$, $\mu \in \mathbb{R}$, $\eta > 0$. Then the posterior distribution is also normal $N(m(x), \rho^2)$, where

$$m(x) = \frac{\sigma^2}{\eta^2 + \sigma^2}\mu + \frac{\eta^2}{\eta^2 + \sigma^2}x \quad \text{and} \quad \rho^2 = \frac{\eta^2\sigma^2}{\eta^2 + \sigma^2},$$

hence it is strictly unimodal and symmetric. It is also obvious that the family $\{g_\mu, \mu \in \mathbb{R}\}$, where g_μ is the density of the normal distribution with location parameter μ and a known scale parameter η, has monotone likelihood ratio. Hence, the corresponding Bayes estimators obtained under squared error loss, uniform loss function and LINEX loss function are increasing in μ.

References

[1] G. E. Andrews, R. Askey and R. Roy, *Special Functions*, Encyclopedia Math. Appl. 17, Cambridge Univ. Press, London, 1999.

[2] R. E. Barlow and F. Proschan, *The Statistical Theory of Reliability*, Holt, Rinehart and Winston, New York, 1975.

[3] J. Bartoszewicz and M. Skolimowska, *Preservation of classes of life distributions and stochastic orders under weighting*, Statist. Probab. Lett. 76 (2006), 587–596.

[4] T. S. Ferguson, *Mathematical Statistics. A Decision Theoretic Approach*, Academic Press, New York, 1967.

[5] S. Karlin, *Total Positivity*, Stanford Univ. Press, Stanford, CA, 1968.

[6] E. L. Lehmann and G. Casella, *Theory of Point Estimation*, 2nd ed., Springer, New York, 1998.

[7] J. Shao, *Mathematical Statistics*, 2nd ed., Springer, New York, 2003.

[8] A. Zellner, *Bayesian estimation and prediction using asymmetric loss functions*, J. Amer. Statist. Assoc. 81, (1986), 446–451.

Jarosław Bartoszewicz

Piotr Bolesław Nowak
Institute of Mathematics and Informatics
University of Opole
Oleska 48
45-052 Opole, Poland
E-mail: pnowak@math.uni.opole.pl

Received on 16.8.2013 (2191)