ESTIMATING BOUNDED MEAN VECTOR IN MULTIVARIATE NORMAL: THE GEOMETRY OF HARTIGAN ESTIMATOR

Cristiane Alvarenga GAJO[1]
Leandro da Silva PEREIRA[1]
Lucas Monteiro CHAVES[1]
Devanil Jaques SOUZA[1]

- *ABSTRACT: The problem on estimating a multivariate normal mean $N_p(\theta, I)$ when the vector mean is bounded awaked interest practical and theoretical. Under such hypothesis it's possible to obtain estimators which dominate the sample mean estimator in relation to square loss. Generalizing previous results obtained, for univariate normal, J.A. Hartigan obtained, for multivariate normal with independent components, a Bayes estimator defined on a bounded closed convex set, with non-empty interior, which dominates the sample mean estimator. In this work, this result is presented in details for the case where the restriction set is a sphere centered at origin. A geometrical interpretation, useful to understand the phenomenon, is presented. Others estimators based on Gatsonis et. al. (1987) are proposed and the risks of all these estimators are compared through simulations, for the cases of dimensions $p = 1$ and $p = 2$.*

- *KEYWORDS: Multivariate normal; convex sets; uniform priors; Bayes estimator.*

1 Introduction

After Stein's (1956) result, namely, a shrinkage estimator for the mean vector of a p-variate normal, $\mathrm{N_p}(\theta, \mathrm{I})$, which dominates the usual estimator $\delta(\mathrm{X}) = \mathrm{X}$ with respect to squared risk, there was an intense search for other estimators which dominate $\delta(\mathrm{X})$. In this line of thought, it arises the problem of, supposing the mean vector $\boldsymbol{\theta}$ being restricted to a limited set $\mathrm{C} \subseteq \mathbb{R}^{\mathrm{p}}$, obtaining the estimator which dominate $\delta(\mathrm{X})$ when to the risk function is restricted to $\boldsymbol{\theta} \in \mathrm{C}$. Generalizing

[1]Federal University of Lavras, Department of Exact Sciences, CEP: 372000-000, Lavras, MG, Brazil. E-mail: *cristianegajo@yahoo.com.br; leandropereira@utfpr.edu.br; lucas@dex.ufla.br; devaniljaques@dex.ufla.br*

previous results from Casella-Strawderman (1981), Gatsonis et al. (1987) and Hartigan (2004) obtained a generalized Bayes estimator for a uniform priori in a bounded closed convex set, with non-empty interior C and smooth enough border that has squared risk smaller or equal to $\delta(\mathrm{X})$ for any $\boldsymbol{\theta} \in \mathrm{C}$. Since the result is stated for general convex sets it uses techniques as measure theory and also the paper being succinctly written makes its comprehension difficult, both for the mathematical aspects and for a intuitive interpretation of the result.

In this paper, Hartingan's result is rewritten in details where is assumed that C is a sphere centered at origin and emphasizing the geometrical aspect. Others estimators based on Gatsonis et al. (1987) are proposed and a computational simulation study related to risk reduction is performed. It is proved that for origin-centered hypercube, the Bayes estimator also dominates the maximum likelihood estimator. The used notation is similiar to that one used by Gatsonis close as possible to the of Gatsonis et al. (1987) and Hartigan (2004).

2 Bayes estimator related to uniform priori dominates mean estimator

Consider the random vector $\mathbf{X} = (X_1, \ldots, X_{\mathrm{p}}) \sim \mathrm{N_p}(\boldsymbol{\theta}, \mathbf{I})$, $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_{\mathrm{p}}) \in \mathrm{C} \subseteq \mathbb{R}^{\mathrm{p}}$ with C as a ball centered at origin and radius m. Probability density of $\mathbf{X}$ is the p-multivariate normal given by

$$f_{\mathbf{X}}(\mathbf{x};\boldsymbol{\theta}) = \frac{1}{(2\pi)^{\frac{\mathrm{p}}{2}}} e^{-\frac{1}{2}\|\mathbf{x}-\boldsymbol{\theta}\|^2} = \prod_{\mathrm{i}=1}^{\mathrm{p}} \frac{1}{(2\pi)^{\frac{\mathrm{p}}{2}}} e^{-\frac{1}{2}(x_{\mathrm{i}}-\theta_{\mathrm{i}})^2} = \prod_{\mathrm{i}=1}^{\mathrm{p}} \phi(x_{\mathrm{i}} - \theta_{\mathrm{i}})$$

in which $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$. Is used, with certain abuse of the notation, $\phi(\mathbf{x} - \boldsymbol{\theta}) = f_{\mathbf{X}}(x_1, \ldots x_{\mathrm{p}}; \boldsymbol{\theta}) = \prod_{\mathrm{i}=1}^{\mathrm{p}} \phi(x_{\mathrm{i}} - \theta_{\mathrm{i}})$.

Considering a uniform priori on C, given by $\pi(\boldsymbol{\theta}) = \frac{1}{vol(\mathrm{C})}$, then the posteriori distribution is given by
$\pi(\boldsymbol{\theta}|\,\mathbf{x}) = \frac{\phi(\mathbf{x}-\boldsymbol{\theta})\frac{1}{vol(\mathrm{C})}}{\int_{\mathrm{C}} \phi(\mathbf{x}-\boldsymbol{\theta})\frac{1}{vol(\mathrm{C})}\mathrm{d}\boldsymbol{\theta}} = \frac{\phi(\mathbf{x}-\boldsymbol{\theta})}{\int_{\mathrm{C}} \phi(\mathbf{x}-\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta}}$.

Bayes estimator $\mathbf{T}(\mathbf{x}) = (\mathrm{T}_1(\mathbf{x}), \ldots, \mathrm{T_p}(\mathbf{x}))$is given by the posteriori mean, e.g.

$$T_j(\mathbf{x}) = \frac{\int_C \theta_j \phi(\mathbf{x} - \boldsymbol{\theta})\, d\boldsymbol{\theta}}{\int_C \phi(\mathbf{x} - \boldsymbol{\theta})\, d\boldsymbol{\theta}}.$$

For simplicity, it is used the notation $\mathbf{T}(\mathbf{x}) = \frac{\int_{\mathrm{C}} \boldsymbol{\theta}\, \phi(\mathbf{x} - \boldsymbol{\theta})\, \mathrm{d}\boldsymbol{\theta}}{\int_{\mathrm{C}} \phi(\mathbf{x} - \boldsymbol{\theta})\, \mathrm{d}\boldsymbol{\theta}}$. Squared risk of this estimator is given by

$$\mathrm{R}_{\mathbf{T}}\left(\boldsymbol{\theta}\right)=\mathrm{E}_{\theta}\left[\sum_{\mathrm{i}=1}^{\mathrm{p}}\left(\mathrm{T}_{\mathrm{i}}\left(\mathbf{x}\right)-\theta_{\mathrm{i}}\right)^{2}\right]=\sum_{i=1}^{p}\int\limits_{\mathbb{R}^{p}}\left(T_{i}(\mathbf{x})-\theta_{\mathbf{i}}\right)^{2}\phi\left(\mathbf{x}-\boldsymbol{\theta}\right)d\mathbf{x}.$$

For the estimator $\delta\left(\mathrm{X}\right)=\mathrm{X}$, not considering any restriction about mean vector $\boldsymbol{\theta}$, the risk is

$$\mathrm{R}_{\delta}\left(\boldsymbol{\theta}\right)=\mathrm{E}_{\theta}\left[\left\|\mathrm{X}-\boldsymbol{\theta}\right\|^{2}\right]=\mathrm{E}_{\theta}\left[\sum_{\mathrm{i}=1}^{\mathrm{p}}\left(\mathrm{X}_{\mathrm{i}}-\theta_{\mathrm{i}}\right)^{2}\right]=\sum_{\mathrm{i}=1}^{\mathrm{p}}\mathrm{E}_{\theta}\left[\left(\mathrm{X}_{\mathrm{i}}-\theta_{\mathrm{i}}\right)^{2}\right]=\sum_{\mathrm{i}=1}^{\mathrm{p}}1=\mathrm{p}.$$

The boundary $\partial\mathrm{C}$ of the p-dimensional ball C is a $(p-1)$-dimensional sphere. The unit normal vector to $\partial\mathrm{C}$ in a point $\boldsymbol{\theta}\in\partial\mathrm{C}$ is given by $\frac{\boldsymbol{\theta}}{\|\boldsymbol{\theta}\|}=\eta\left(\boldsymbol{\theta}\right)$. See Figure 1.

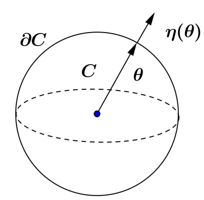

Figure 1 - Sphere unit normal field.

If $\mathrm{p}(\mathbf{x})=\mathrm{p}\left(\mathrm{x}_{1},\ldots\mathrm{x}_{\mathrm{p}}\right)=\int\limits_{\mathrm{C}}\phi\left(\mathbf{x}-\boldsymbol{\theta}\right)\mathrm{d}\boldsymbol{\theta}$, then

$$\begin{aligned}\frac{\partial\mathrm{p}}{\partial x_{\mathrm{i}}}\left(\mathbf{x}\right) &= \frac{\partial}{\partial x_{\mathrm{i}}}\int\limits_{\mathrm{C}}\phi\left(\mathbf{x}-\boldsymbol{\theta}\right)\mathrm{d}\boldsymbol{\theta}=\int\limits_{\mathrm{C}}\frac{\partial}{\partial x_{\mathrm{i}}}\phi\left(\mathbf{x}-\boldsymbol{\theta}\right)\mathrm{d}\boldsymbol{\theta}\\ &= \int\limits_{\mathrm{C}}\frac{\partial}{\partial x_{\mathrm{i}}}\frac{1}{\left(2\pi\right)^{\frac{\mathrm{p}}{2}}}e^{-\frac{1}{2}\sum\left(x_{\mathrm{i}}-\theta_{\mathrm{i}}\right)^{2}}\mathrm{d}\boldsymbol{\theta}=\int\limits_{\mathrm{C}}-\phi\left(\mathbf{x}-\theta\right)\left(x_{\mathrm{i}}-\boldsymbol{\theta}_{\mathrm{i}}\right)\mathrm{d}\boldsymbol{\theta}\\ &= -x_{\mathrm{i}}\int\limits_{\mathrm{C}}\phi\left(\mathbf{x}-\boldsymbol{\theta}\right)\mathrm{d}\boldsymbol{\theta}+\int\limits_{\mathrm{C}}\theta_{\mathrm{i}}\phi\left(\mathbf{x}-\boldsymbol{\theta}\right)\mathrm{d}\boldsymbol{\theta}=-x_{\mathrm{i}}\,\mathrm{p}\left(\mathbf{x}\right)+\mathrm{T}_{\mathrm{i}}\left(\mathbf{x}\right)\mathrm{p}\left(\mathbf{x}\right),\end{aligned}$$

and therefore $\mathrm{T}_{\mathrm{i}}\left(\mathbf{x}\right)=x_{\mathrm{i}}+\frac{1}{\mathrm{p}(\mathbf{x})}\frac{\partial\mathrm{p}}{\partial x_{\mathrm{i}}}\left(\mathbf{x}\right).$

Comparing risks for a vector $\boldsymbol{\xi}$ in the interior of C

$$
\begin{aligned}
\mathrm{R}_{\mathbf{T}}(\boldsymbol{\xi}) - \mathrm{R}_{\delta}(\boldsymbol{\xi}) &= \mathrm{R}_{\mathbf{T}}(\boldsymbol{\xi}) - p \\
&= \mathrm{E}_{\boldsymbol{\xi}}\left[\|\mathbf{T}-\boldsymbol{\xi}\|^2\right] - \mathrm{E}_{\boldsymbol{\xi}}\left[\|\mathbf{X}-\boldsymbol{\xi}\|^2\right] \\
&= \mathrm{E}_{\boldsymbol{\xi}}\left[\|\mathbf{T}-\boldsymbol{\xi}\|^2 - \|\mathrm{X}-\boldsymbol{\xi}\|^2\right] \\
&= \int\limits_{\mathrm{p}} \left\{\sum_{\mathrm{i}=1}^{\mathrm{p}}\left[\left(\mathrm{x_i} + \frac{1}{\mathrm{p}(\mathbf{x})}\frac{\partial \mathrm{p}}{\partial x_{\mathrm{i}}}(\mathbf{x}) - \xi_{\mathrm{i}}\right)^2 - (\mathrm{x_i}-\xi_{\mathrm{i}})^2\right]\phi(\mathbf{x}-\xi)\right\}\mathrm{d}\mathbf{x} \\
&= \int\limits_{\mathbb{R}^{\mathrm{p}}} \sum_{\mathrm{i}=1}^{\mathrm{p}}\left[\left(\frac{1}{\mathrm{p}(\mathbf{x})}\frac{\partial \mathrm{p}}{\partial x_{\mathrm{i}}}(\mathbf{x})\right)^2 + 2(\mathrm{x_i}-\xi_{\mathrm{i}})\frac{1}{\mathrm{p}(\mathbf{x})}\frac{\partial \mathrm{p}}{\partial x_{\mathrm{i}}}(\mathbf{x})\right]\phi(\mathbf{x}-\boldsymbol{\xi})\,\mathrm{d}\mathbf{x} \\
&= \sum_{\mathrm{i}=1}^{\mathrm{p}}\left[\int\limits_{\mathbb{R}^{\mathrm{p}}}\left(\frac{1}{\mathrm{p}(\mathbf{x})}\frac{\partial \mathrm{p}}{\partial x_{\mathrm{i}}}(\mathbf{x})\right)^2\phi(\mathbf{x}-\xi)\,dx \right.\\
&\qquad \left. +2\int\limits_{\mathbb{R}^{\mathrm{p}}}(\mathrm{x_i}-\xi_{\mathrm{i}})\frac{1}{\mathrm{p}(\mathbf{x})}\frac{\partial \mathrm{p}}{\partial x_{\mathrm{i}}}(\mathbf{x})\phi(\mathbf{x}-\xi)\,\mathrm{d}\mathbf{x}\right].
\end{aligned}
$$

To calculate this integral, observing that

$$
\frac{\partial \phi}{\partial x_{\mathrm{i}}}(\mathbf{x}-\boldsymbol{\xi}) = -\phi(\mathbf{x}-\boldsymbol{\xi})(x_{\mathrm{i}}-\xi_i), \tag{1}
$$

we get that,

$$
\begin{aligned}
\int\limits_{\mathbb{R}^{\mathrm{p}}}(\mathrm{x_i}-\xi_{\mathrm{i}})\frac{1}{\mathrm{p}(\mathbf{x})}\frac{\partial \mathrm{p}}{\partial x_{\mathrm{i}}}(\mathbf{x})\,\phi(\mathbf{x}-\xi)\,\mathrm{d}\mathbf{x} &= \int\limits_{\mathbb{R}^{\mathrm{p}}}(\mathrm{x_i}-\xi_{\mathrm{i}})\phi(\mathbf{x}-\xi)\frac{1}{\mathrm{p}(\mathbf{x})}\frac{\partial \mathrm{p}}{\partial x_{\mathrm{i}}}(\mathbf{x})\,\mathrm{d}\mathbf{x} \\
&= -\int\limits_{\mathbb{R}^{\mathrm{p}}}\left(\frac{\partial}{\partial x_{\mathrm{i}}}\phi(\mathbf{x}-\xi)\frac{1}{\mathrm{p}(\mathbf{x})}\frac{\partial \mathrm{p}}{\partial x_{\mathrm{i}}}(\mathbf{x})\right)\mathrm{d}\mathbf{x}.
\end{aligned}
$$

By integration by parts,

$$
\begin{aligned}
\int\limits_{\mathrm{p}}(\mathrm{x_i}-\xi_{\mathrm{i}})\frac{1}{\mathrm{p}(\mathbf{x})}\frac{\partial \mathrm{p}}{\partial x_{\mathrm{i}}}(\mathbf{x})\,\phi(\mathbf{x}-\xi)\mathrm{d}\mathbf{x} &= -\left[\phi(\mathbf{x}-\xi)\left(\frac{1}{\mathrm{p}(\mathbf{x})}\frac{\partial \mathrm{p}}{\partial x_{\mathrm{i}}}(\mathbf{x})\right)\Big|_{-\infty}^{\infty}\right. \\
&\quad \left. -\int\limits_{\mathbb{R}^{\mathrm{p}}}\left(\frac{1}{\mathrm{p}(\mathbf{x})}\frac{\partial^2 \mathrm{p}}{\partial {x_{\mathrm{i}}}^2}(\mathbf{x}) - \frac{\partial \mathrm{p}}{\partial x_{\mathrm{i}}}(\mathbf{x})\frac{\frac{\partial \mathrm{p}}{\partial x_{\mathrm{i}}}(\mathbf{x})}{\mathrm{p}(\mathbf{x})^2}\right)\phi(\mathbf{x}-\boldsymbol{\xi})\,\mathrm{d}\mathbf{x}\right] \\
&= \int\limits_{\mathbb{R}^{\mathrm{p}}}\left[\frac{1}{\mathrm{p}(\mathbf{x})}\frac{\partial^2 \mathrm{p}}{\partial {x_{\mathrm{i}}}^2}(\mathbf{x}) - \left(\frac{1}{\mathrm{p}(\mathbf{x})}\frac{\partial \mathrm{p}}{\partial x_{\mathrm{i}}}(\mathbf{x})\right)^2\right]\phi(\mathbf{x}-\boldsymbol{\xi})\,\mathrm{d}\mathbf{x}.
\end{aligned}
$$

From this identity follows that,

$$\int\limits_{\mathbb{R}^{\mathrm{p}}} \left(\frac{1}{\mathrm{p}}\frac{\partial \mathrm{p}}{\partial x_{\mathrm{i}}}(\mathbf{x})\right)^2 \phi(\mathbf{x}-\boldsymbol{\xi})\,\mathrm{d}\mathbf{x} = \int\limits_{\mathbb{R}^{\mathrm{p}}} \left[\frac{1}{\mathrm{p}(\mathbf{x})}\frac{\partial^2 \mathrm{p}}{\partial {x_{\mathrm{i}}}^2}(\mathbf{x})\,\phi(\mathbf{x}-\boldsymbol{\xi}) - \frac{1}{\mathrm{p}(\mathbf{x})}\frac{\partial \mathrm{p}}{\partial x_{\mathrm{i}}}(\mathbf{x})\,\phi(\mathbf{x}-\boldsymbol{\xi})(\mathrm{x}_{\mathrm{i}}-\xi_{\mathrm{i}})\right]\mathrm{d}\mathbf{x}$$

$$= \int\limits_{\mathbb{R}^{\mathrm{p}}} \frac{1}{\mathrm{p}(\mathbf{x})}\left[\frac{\partial^2 \mathrm{p}}{\partial {x_{\mathrm{i}}}^2}(\mathbf{x}) - (\mathrm{x}_{\mathrm{i}}-\xi_{\mathrm{i}})\frac{\partial \mathrm{p}}{\partial x_{\mathrm{i}}}(\mathbf{x})\right]\phi(\mathbf{x}-\boldsymbol{\xi})\,\mathrm{d}\mathbf{x}$$

Replacing this equality in the expression $\mathrm{R}_{\mathbf{T}}(\boldsymbol{\xi}) - \mathrm{R}_{\delta}(\boldsymbol{\xi})$ we have

$$\begin{aligned}
\mathrm{R}_{\mathbf{T}}(\boldsymbol{\xi}) - \mathrm{R}_{\delta}(\boldsymbol{\xi}) &= \int\limits_{\mathbb{R}^{\mathrm{p}}} \sum_{\mathrm{i}=1}^{\mathrm{p}} \left[\left(\frac{1}{\mathrm{p}(\mathbf{x})}\frac{\partial \mathrm{p}}{\partial x_{\mathrm{i}}}(\mathbf{x})\right)^2 + 2(\mathrm{x}_{\mathrm{i}}-\boldsymbol{\xi}_{\mathrm{i}})\frac{1}{\mathrm{p}(\mathbf{x})}\frac{\partial \mathrm{p}}{\partial x_{\mathrm{i}}}(\mathbf{x})\right]\phi(\mathbf{x}-\boldsymbol{\xi})\,\mathrm{d}\mathbf{x} \\
&= \int\limits_{\mathbb{R}^{\mathrm{p}}} \sum_{\mathrm{i}=1}^{\mathrm{p}} \left[\frac{1}{\mathrm{p}(\mathbf{x})}\left(\frac{\partial^2 \mathrm{p}}{\partial x_i^2}(\mathbf{x}) - (\mathrm{x}_{\mathrm{i}}-\boldsymbol{\xi}_{\mathrm{i}})\frac{\partial \mathrm{p}}{\partial x_{\mathrm{i}}}(\mathbf{x})\right)\phi(\mathbf{x}-\boldsymbol{\xi})\right. \\
&\quad \left. +2(\mathrm{x}_{\mathrm{i}}-\boldsymbol{\xi}_{\mathrm{i}})\frac{1}{\mathrm{p}(\mathbf{x})}\frac{\partial \mathrm{p}}{\partial x_{\mathrm{i}}}(\mathbf{x})\,\phi(\mathbf{x}-\boldsymbol{\xi})\right]\mathrm{d}\mathbf{x} \\
&= \int\limits_{\mathbb{R}^{\mathrm{p}}} \sum_{\mathrm{i}=1}^{\mathrm{p}} \frac{1}{\mathrm{p}(\mathbf{x})}\left[\left(\frac{\partial^2 \mathrm{p}}{\partial x_i^2}(\mathbf{x}) - (\mathrm{x}_{\mathrm{i}}-\boldsymbol{\xi}_{\mathrm{i}})\frac{\partial \mathrm{p}}{\partial x_{\mathrm{i}}}(\mathbf{x})\right)\phi(\mathbf{x}-\boldsymbol{\xi})\,\mathrm{d}\mathbf{x}.\right.
\end{aligned} \tag{2}$$

Note that $\mathrm{p}(\mathbf{x}) = \int\limits_{\mathrm{C}} \phi(\mathbf{x}-\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta}$, $\frac{\partial \mathrm{p}}{\partial x_i}(\mathbf{x}) = -\int\limits_{C} \phi(\mathbf{x}-\boldsymbol{\theta})(x_i-\theta_i)d\boldsymbol{\theta}$,

$$\begin{aligned}
\frac{\partial^2 \mathrm{p}}{\partial x_i^2}(\mathbf{x}) &= -\int\limits_{C} \left[-\phi(\mathbf{x}-\boldsymbol{\theta})(x_i-\theta_i)^2 + \phi(\mathbf{x}-\boldsymbol{\theta})\right]d\boldsymbol{\theta} \\
&= \int\limits_{C} \phi(\mathbf{x}-\boldsymbol{\theta})\left[(x_i-\theta_i)^2 - 1\right]d\boldsymbol{\theta}
\end{aligned}$$

hence

$$\frac{\partial^2 \mathrm{p}}{\partial x_i^2}(\mathbf{x}) + (x_i-\xi_i)\frac{\partial p}{\partial x_i}(\mathbf{x}) = \int\limits_{C} \phi(\mathbf{x}-\boldsymbol{\theta})\left[(x_i-\theta_i)^2 - 1\right]d\boldsymbol{\theta} - (x_i-\xi_i)\int\limits_{C} \phi(\mathbf{x}-\boldsymbol{\theta})(x_i-\theta_i)d\boldsymbol{\theta}$$

$$= \int\limits_{C} \left\{\left[(x_i-\theta_i)^2 - 1\right]\phi(\mathbf{x}-\boldsymbol{\theta}) - (x_i-\xi_i)(x_i-\theta_i)\phi(\mathbf{x}-\boldsymbol{\theta})\right\}d\boldsymbol{\theta}. \tag{3}$$

Using the following equality

$$\begin{aligned} \frac{\partial \phi}{\partial \theta_i}(\mathbf{x}-\boldsymbol{\theta}) &= -\phi(\mathbf{x}-\boldsymbol{\theta})(x_i-\theta_i) \\ \frac{\partial^2 \phi}{\partial \theta_i^2}(\mathbf{x}-\boldsymbol{\theta}) &= \frac{\partial \phi}{\partial \theta_i}(\mathbf{x}-\boldsymbol{\theta})(x_i-\theta_i) - \phi(\mathbf{x}-\boldsymbol{\theta}) \\ &= \phi(\mathbf{x}-\boldsymbol{\theta})(x_i-\theta_i)^2 - \phi(\mathbf{x}-\boldsymbol{\theta}) = \left[(x_i-\theta_i)^2-1\right]\phi(\mathbf{x}-\boldsymbol{\theta}), \end{aligned} \tag{4}$$

replacing in equation (3) and sum in i we have

$$\sum_i \left[\frac{\partial^2 p}{\partial x_i^2}(\mathbf{x}) + (x_i-\xi_i)\frac{\partial p}{\partial x_i}(\mathbf{x})\right] = \int_C \sum_i \left[\frac{\partial^2 \phi}{\partial \theta_i^2}(\mathbf{x}-\boldsymbol{\theta}) + (x_i-\xi_i)\frac{\partial \phi}{\partial \theta_i}(\mathbf{x}-\boldsymbol{\theta})\right] d\boldsymbol{\theta}.$$

If $g(x_1,...,x_p) = (g_1(x_1,...,x_p),...,g_p(x_1,...,x_p))$ is a vector field and $\psi = \psi(x_1,...,x_p)$ a scalar function, the divergence of the fiel g is defined as $\nabla \cdot g = \frac{\partial g_1}{\partial x_1} + ... + \frac{\partial g_p}{\partial x_p}$ and the gradient of ψ is the vector field $\mathsf{grad}\psi(x_1,...,x_p) = \left(\frac{\partial \psi}{\partial x_1},...,\frac{\partial \psi}{\partial x_1}\right)$. The classical divergence theorem (Gauss Theorem) ensures us that the volume integral of the divergent of the field in a ball is equal to surface integral of inner product of the vector field with unit normal field η

$$\int_C \nabla \cdot g dv = \int_{\partial C} (g \cdot \eta) ds.$$

To use this theorem consider the vector field

$$\begin{aligned} &grad\phi(\mathbf{x}-\theta) - (\mathbf{x}-\xi)\phi(\mathbf{x}-\theta) = \\ &\quad \left(\frac{\partial \phi}{\partial \theta_1}(\mathbf{x}-\theta) - (x_1-\xi_1)\phi(\mathbf{x}-\theta),...,\frac{\partial \phi}{\partial \theta_p}(\mathbf{x}-\theta) - (x_p-\xi_p)\phi(\mathbf{x}-\theta)\right), \end{aligned}$$

$\mathsf{grad}\phi(\mathbf{x}-\boldsymbol{\theta})$ is taken in relation to θ with x fixed in which $\mathtt{grad}\phi(\mathbf{x}-\boldsymbol{\theta})$ is the gradient field of the function $\phi(\mathbf{x}-\boldsymbol{\theta})$. Follow from (4) that $\mathtt{grad}\phi(\mathbf{x}-\boldsymbol{\theta}) = (\mathbf{x}-\boldsymbol{\theta})\phi(\mathbf{x}-\boldsymbol{\theta})$. Adding and subtracting ξ the gradient field can be expressed as

$$\mathsf{grad}\phi(\mathbf{x}-\boldsymbol{\theta}) - (\mathbf{x}-\boldsymbol{\xi})\phi(\mathbf{x}-\boldsymbol{\theta}) = (\boldsymbol{\xi}-\boldsymbol{\theta})\phi(\mathbf{x}-\boldsymbol{\theta}).$$

Then $\nabla \cdot (\boldsymbol{\xi}-\boldsymbol{\theta}))\phi(\mathbf{x}-\boldsymbol{\theta}) = \nabla \cdot [\mathsf{grad}\phi(\mathbf{x}-\boldsymbol{\theta}) - (\mathbf{x}-\boldsymbol{\xi})\phi(\mathbf{x}-\boldsymbol{\theta})] =$
$\nabla \cdot \mathsf{grad}\phi(\mathbf{x}-\boldsymbol{\theta}) - \nabla \cdot [(\mathbf{x}-\boldsymbol{\xi})\phi(\mathbf{x}-\boldsymbol{\theta})] = \sum_i \left[\frac{\partial^2 \phi}{\partial \theta_i^2}(\mathbf{x}-\boldsymbol{\theta}) + (x_i-\xi_i)\frac{\partial \phi}{\partial \theta_i}(\mathbf{x}-\boldsymbol{\theta})\right]$.
Applying the divergence theorem we get the equality

$$\int_C \nabla \cdot (\boldsymbol{\xi}-\boldsymbol{\theta})\phi(\mathbf{x}-\boldsymbol{\theta}) dv = \int_{S^{p-1}} (\boldsymbol{\xi}-\boldsymbol{\theta})\phi(\mathbf{x}-\boldsymbol{\theta}) \cdot \left(\frac{\theta}{\|\theta\|}\right) ds.$$

Observe that $\eta(\mathbf{x}) = \frac{\theta}{\|\theta\|}$ is the unit normal field to sphere S^{p-1}. Replacing this result in (2) we have

$$\mathrm{R}_{\mathbf{T}}(\boldsymbol{\xi}) - \mathrm{R}_{\delta}(\boldsymbol{\xi}) = \sum_{i=1}^{p} \left[\int_{\mathbb{R}^p} \frac{1}{p(\mathbf{x})} \left[\int_{S^{p-1}} (\boldsymbol{\xi} - \boldsymbol{\theta}) \phi(\mathbf{x} - \boldsymbol{\theta}) \cdot \left(\frac{\theta}{\|\theta\|} \right) ds \right] \phi(\mathbf{x} - \boldsymbol{\xi}) d\mathbf{x} \right].$$

Now it is possible to have an interesting geometrical interpretation as described on Figure 2. The angle between the vector $\boldsymbol{\xi} - \boldsymbol{\theta}$ and the unit normal vector $\eta(\theta)$, for any $\boldsymbol{\xi}$ in the interior of the ball and any vector $\boldsymbol{\theta}$ in the border S^{p-1}, is bigger than 90^0, so the inner product of these vectors is always negative that is,

$$(\boldsymbol{\xi} - \boldsymbol{\theta}) \cdot \left(\frac{\theta}{\|\theta\|} \right) < 0 \quad \text{so} \quad \mathrm{R}_{\mathbf{T}}(\boldsymbol{\xi}) - \mathrm{R}_{\delta}(\boldsymbol{\xi}) < 0,$$

and hence $\mathbf{T}$ estimator dominates mean estimator δ. See Figure 2.

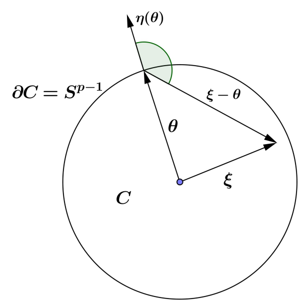

Figure 2 - The angle between $\eta(\theta)$ and $\xi - \theta$.

2.1 Some new results

Comparison between Bayes estimator $\mathbf{T}$ and the usual estimator $\delta(X) = X$ is not really fair, because δ estimator does not consider that the mean vector parameter is bounded. A more suitable estimator to compare is the maximum likelihood estimator in relation to parametric restriction.

$$\delta_m^{ML}(\mathbf{x}) = \begin{cases} m\dfrac{\mathbf{x}}{\|\mathbf{x}\|}, & \|\mathbf{x}\| \geq m \\ \mathbf{x}, & \|\mathbf{x}\| < m. \end{cases}$$

A Hartigan-kind integral formula for the difference between the risks of these estimators would be excessively complex. However, an interesting observation is that the comparison can be obtained through the analysis of unidimensional case developed by Gatsonis et al. (1987). They also considered an uniform priori on the interval $[-m, m]$, getting the Bayes estimator

$$\delta_m(x) = \frac{\int_{-m}^{m} \theta\phi(x-\theta)\, d\theta}{\int_{-m}^{m} \phi(x-\theta)\, d\theta}.$$

δ_m dominates the usual δ estimator, and dominates the maximum likelihood unit estimator δ_m^{ML} for θ in an interval close to $\left[-\frac{3}{4}m,\ \frac{3}{4}m\right]$.

To use the unidimensional result consider the case in which the parametric restriction defines a hypercube centered at origin, i.e., $\boldsymbol{\theta} = (\theta_1, ..., \theta_p)$ with $|\theta_i| \leq m$ $i = 1, ..., p$. Observe that for practical applications this restriction can be more natural than $\|\boldsymbol{\theta}\| \leq m$. In this case, Bayes estimator $\mathbf{T}$ has components in the form

$$\begin{aligned} T_i(x) &= \frac{\int\limits_C \theta_i \phi(x-\theta)\, d\theta}{\int\limits_C \phi(x-\theta)\, d\theta} \\ &= \frac{\int\limits_C \theta_i \prod\limits_{j=1}^{p} \phi(x_j-\theta_j) d\theta}{\int\limits_C \prod\limits_{j=1}^{p} \phi(x_j-\theta_j) d\theta} \\ &= \frac{\int_{-m}^{m} \theta_i \phi(x_i-\theta_i) d\theta_i \int_{-m}^{m} ... \int_{-m}^{m} \prod\limits_{\substack{j=1 \\ j\neq i}}^{p} \phi(x_j-\theta_j) d\theta}{\int_{-m}^{m} \phi(x_i-\theta_i) d\theta_i \int_{-m}^{m} ... \int_{-m}^{m} \prod\limits_{\substack{j=1 \\ j\neq i}}^{p} \phi(x_j-\theta_j) d\theta} \\ &= \frac{\int_{-m}^{m} \theta_i \phi(x_i-\theta_i) d\theta_i}{\int_{-m}^{m} \phi(x_i-\theta_i) d\theta_i} \\ &= \delta_m(x_i). \end{aligned}$$

Then, it follows that Hartigan estimator for the hypercube is expressed in terms of the unidimensional estimator obtained by Gatsonis et al. (1987) as $T(\mathbf{x}) = (\delta_{\mathbf{m}}(\mathbf{x_1}), ..., \delta_{\mathbf{m}}(\mathbf{x_p}))$. This estimator will be denoted by $\delta_m(\mathbf{x})$. In this case the difference is

$$\begin{aligned} \mathrm{R}_{\mathbf{T}}(\boldsymbol{\xi}) - \mathrm{R}_{\delta_m^{ML}}(\boldsymbol{\xi}) &= E\left[\|T(x)-\boldsymbol{\xi}\|^2\right] - E\left[\left\|\delta_m^{ML}(x)-\boldsymbol{\xi}\right\|^2\right] \\ &= \sum_{i=1}^{p} \left\{E\left[|T_i(x)-\xi_i|^2\right] - E\left[\left|\delta_m^{ML}(x_i)-\xi_i\right|^2\right]\right\} \\ &= \sum_{i=1}^{p} \left\{E\left[|\delta_m(x_i)-\xi_i|^2\right] - E\left[\left|\delta_m^{ML}(x_i)-\xi_i\right|^2\right]\right\}. \end{aligned}$$

By the result of Galtonis et al. (1987) for $-\frac{3}{4}m \leq x_i \leq \frac{3}{4}m$, we have $\mathrm{R}_{\mathbf{T}}(\boldsymbol{\xi}) - \mathrm{R}_{\delta_m^{ML}}(\boldsymbol{\xi}) \leq 0$ and therefore $\mathbf{T}$ dominates δ_m^{ML}. The case where the restriction is a sphere will be studied through computational simulation.

2.2 A study through computational simulation

Aiming to relate Hartigan's result (2004) with the estimators studied in Casella and Strawderman (1981) and Gatsonis et al. (1987) a computational simulation in dimension 2 was performed. In these papers, the estimators are

$$\begin{aligned} &\delta_m^{ML} \quad , \quad \text{Maximum likelihood restricted estimator} \\ &\delta_m(x) = x + \frac{g'(x)}{g(x)} \quad , \quad g(x) = \Phi(m-x) + \Phi(m+x) - 1 \\ &\delta_m^0(x) = m\ tanh(mx) \end{aligned}$$

where Φ is the standard normal cumulative function and the mean is limited to the interval $(-m, m)$.
The results were:

- δ_m^0 is minimax for $0 \leq m \leq m_0 \simeq 1,056742$ and dominates δ_m^{ML} for $m < 1$.
- δ_m dominates usual mean estimator, e.g, it has risk smaller than 1.
- δ_m dominates δ_m^{ML} in the approximated interval $\left(-\frac{3}{4}m, \frac{3}{4}m\right)$.

.

For a multivariate normal mean case with the vector mean limited to a sphere of radius m in $\mathbb{R}^p$, these results can be generalized:

- Hartigan $\mathbf{T}$ estimator generalizes estimator δ_m.
- Estimator δ_m^0 will be generalized to the form

$$\delta_m^0(\mathbf{x}) = (m\,tanh(mx_1), ..., m\,tanh(mx_p)).$$

- δ_m can be also generalized to $\delta_m(\mathbf{x}) = (\delta_m(x_1), ..., \delta_m(x_p))$.

Observe that the estimators $\delta_m^0(\mathbf{x})$ and $\delta_m(\mathbf{x})$ are in fact generating estimates on a p-dimensional m sided hypercube, $\{(x_1, ..., x_p) \in \mathbb{R}^p,\ -m \leq x_i \leq m,\ i = 1, ..., p\}$. However by simulation is observed that these estimmatives are more concentrated in the m radius sphere inside the hipercube. Therefore it makes sense to compare $\delta_m^0(\mathbf{x})$ and $\delta_m(\mathbf{x})$ with the estimators δ_m^{ML} and $\mathbf{T}$ which generate only estimates in the sphere. Since there isn't an explicit formula for the risks of these estimators the comparation were obtained through simulation.

Simulation was carried out at is software R environment. A bidimensional reticulate with 2500 points $\theta = (\theta_1, \theta_2)$ was built in the sphere. For each value of θ a sample of the two dimensional normal $N_2(\theta, I)$ is generated. Then, it is obtained a discret version of the risk graphics (Figures 3-7).

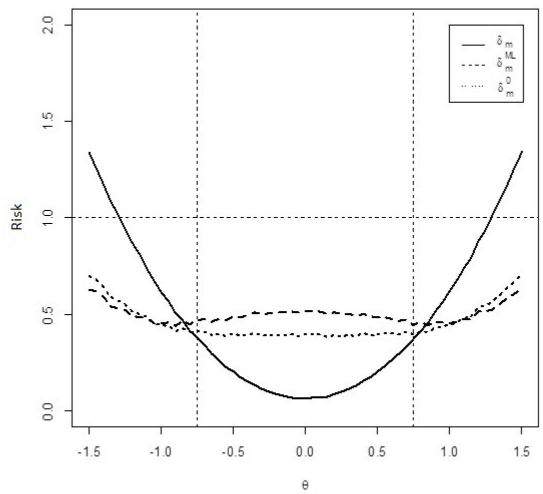

Figure 3 - Estimator risks for dimension 1 and m=1.

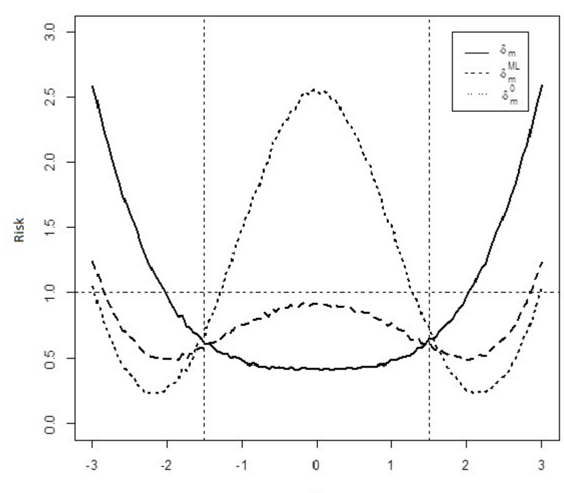

Figure 4 - Estimator risks for dimension 1 and m=2.

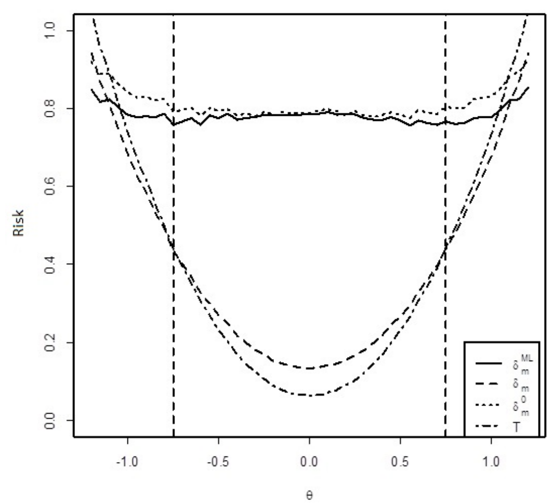

Figure 5 - Estimator risks for dimension 2 and m= 1.

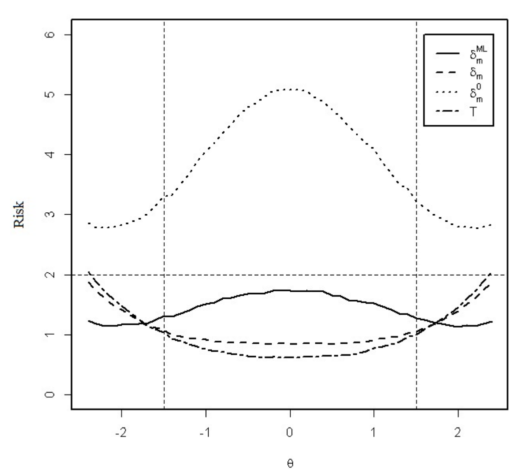

Figure 6 - Estimator risks for dimension 2 and m= 2.

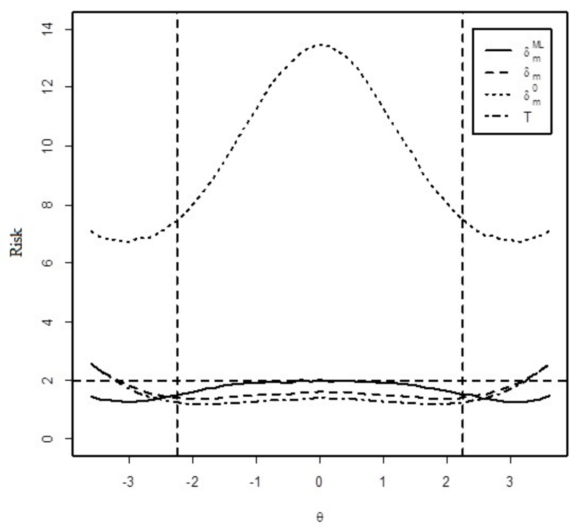

Figure 7 - Estimator risks for dimension 2 and m= 3.

Through graphic analysis we obtain:

- Unidimensional minimax estimator δ_m^0 when generalized to dimension 2 completely loses its relative advantages to other estimators.
- Hartigan estimator $\mathbf{T}(\mathbf{x})$, dominates $\delta_m(\mathbf{x})$ on the sphere of radius $m = 3$. For radius 2 and 1 and θ close to the border of the sphere an inversion happens with $\delta_m(\mathbf{x})$ dominating $\mathbf{T}(\mathbf{x})$.
- Hartigan estimator $\mathbf{T}(\mathbf{x})$ dominates the maximum likelihood estimator $\delta_m^{ML}(\mathbf{x})$ except when θ is close to the sphere's border. This fact was intuitively expected since, $\delta_m^{ML}(\mathbf{x})$ tends to produce estimates at the border.

Conclusion

The mathematical theory of Hartigan's estimator is accessible and for the sphere and it has a geometrical meaning. The properties of unidimensional estimators for the bounded mean changes when generalized to the two dimensional space.

GAJO, C. A.; PEREIRA, L. S.; CHAVES, L. M.; SOUZA, D. J. Estimação limitada do vetor de médias na normal multivariada: A geometria do estimador de Hartigan. *Rev.Bras. Biom.,* Lavras, v.34, n.2, p.304-316, 2016.

▪ RESUMO: O problema de se estimar a média de uma normal multivariada $N_p(\theta, I)$ quando se supõe que o vetor de médias é limitado desperta o interesse prático e teórico. Sob tal hipótese é possível obter estimadores que dominam o estimador média amostral em relação a perda quadrática. Generalizando resultados obtidos anteriormente, para a normal univariada, J.A. Hartigan obteve, para a normal multivariada com componentes independentes, um estimador de Bayes definido sobre um conjunto fechado limitado e convexo, com interior não vazio, que também domina o estimador média amostral. Neste trabalho, este resultado é apresentado com detalhes para o caso em que o conjunto restrito é uma esfera centrada na origem. A interpretação geométrica, útil para a compreensão desse fenômeno, é apresentada. Outros estimadores baseados em Gatsonis et. al. (1987) são propostos e os riscos de todos esses estimadores são comparados por simulação computacional, para os casos de dimensão p=1 e p-=2.

▪ PALAVRAS-CHAVE: Normal multivariada; conjuntos convexos; priori uniforme; estimador de Bayes.

References

CASELLA, G.; STRAWDERMAN,W. Estimating a bounded normal mean. *The Annals Statistics*, v.9,n.4, p.870-878, 1981.

GATSONIS, C.; MACGIBBON,B.; STRAWDERMAN,W. On the estimation of a restricted normal mean. *Statistics and Probability Letters*, v.6,n.1, p.21-30, 1987.

HARTIGAN, J.A. Uniform prioris on convex sets improve risk. *Statistics and Probability Letters*, v.67, p.285-288, 2004.

Received in 14.08.2015.

Approved after revised in 17.01.2016.