J. Ital. Statist. Soc. (1993)
1, pp. 17-34

CONCENTRATION FUNCTION AND COEFFICIENTS OF DIVERGENCE FOR SIGNED MEASURES

Sandra Fortini[1]

Fabrizio Ruggeri

Consiglio Nazionale delle Ricerche
Istituto per le Applicazioni della Matematica e dell'Informatica

Summary

Comparisons among probability measures are rather frequent in many statistical problems and they are sometimes performed through the coefficients of divergence or the concentration functions with respect to a reference measure. Extending the notion of Lorenz-Gini curve, the concentration function studies the discrepancy between two probability measures Π and Π_0.

In this paper, both the concentration function and the coefficients have been defined and studied for a signed measure Π, as an extension of the concentration curve for real valued statistical variables. Signed measures are relevant in statistical analysis, even if unusual, because real problems require them, especially in descriptive statistics, like the simple one presented here.

Keywords: Concentration function, coefficients of divergence, Gini's concentration ratio, Pietra index, signed measure.

1. Introduction

Comparisons among two measures on the same measurable space have been the object of many researches, leading to different approaches; a well-known approach is given by the Lorenz-Gini concentration curve (Marshall and Olkin, 1979, p. 5) which compares the actual distribution of wealth among n individuals with the uniform one. Cifarelli and Regazzini (1987) defined the concentration function of a probability measure Π with respect to another Π_0, extending the classical notion of Lorenz-Gini curve. By the concentra-

1. *Address for correspondence*: Sandra Fortini, Fabrizio Ruggeri, Consiglio Nazionale delle Ricerche, Istituto per le Applicazioni della Matematica e dell'Informatica, Via A.M. Ampère, 56, I-20131 Milano, Italy.

tion function, the discrepancy between two measures defined on the same probability space is studied, comparing the different concentrations of probabilities determined by the measures. In particular, Cifarelli and Regazzini (1987) proved, under very general conditions, that it is possible to determine, under Π, the range of the probabilities of all the sets with equal probability content under Π_0. Probability measures have been also compared through the general class of coefficients of divergence defined by Ali and Silvey (1966), including, e.g., the Kullback-Leibler and the variational distances.

Comparisons among two or more measures could be made by some concentration indices, like in Ragazzini (1992), where they induce rankings among the measures. Besides, Ragazzini (1992) proved that the rankings due to the concentration function and the coefficients of divergence coincide, under very general conditions.

At the same time, actual problems require that the comparisons are to be made also with signed measures, e.g. when some of the n individuals in the above Lorenz-Gini scheme have debts. Many authors have already studied such a problem (e.g. Wold (1935), Castellano (1938) and Michetti and Dall'Aglio (1957)), extending the notion of Lorenz-Gini curve to statistical variables also taking negative values.

In this paper, the notions of concentration function and coefficients of divergence are extended to compare a signed measure with a probability one, because of both mathematical and statistical interest. Their main properties are then proved and the links in terms of induced rankings are again proved. Furthermore, both of them are split into the two parts corresponding to the Jordan decomposition of the signed measure. Finally, the results are applied to some examples and further developments are discussed.

2. Comparison among Probability Measures

Consider two probability measures Π and Π_0 on the same measurable space $(\Theta, \mathcal{F})$. According to the Radon-Nikodym theorem, there is a unique partition $\{N, N^C\} \subset \mathcal{F}$ of Θ and a non-negative function h on N^C such that

$$\Pi(E) = \int_{E \cap N^C} h(\vartheta)\, \Pi_0(d\vartheta) + \Pi_s(E \cap N),$$

$\forall E \in \mathcal{F}$, $\Pi_0(N) = 0$, $\Pi_s(N) = \Pi_s(\Theta)$, where

$$\Pi_a(\cdot) = \int_{\cdot \cap N^C} h(\vartheta)\, \Pi_0(d\vartheta)$$

and Π_s denote the absolutely continuous and the singular part of Π with regard to Π_0, respectively. Set $h(\vartheta) = \infty$ over N and define $H(y) = \Pi_0(\{\vartheta \in \Theta : h(\vartheta) \leq y\})$, $c(x) = inf\{y \in \Re : H(y) \geq x\}$. Finally, let $L(X) = \{\vartheta \in \Theta : h(\vartheta) \leq c(x)\}$ and $L^-(x) = \{\vartheta \in \Theta : h(\vartheta) < c(x)\}$.

Definition 1. The function $\varphi : [0,1] \to [0,1]$ defined by

$$\varphi(x) = \begin{cases} 0 & x = 0 \\ \Pi(L^-(x)) + c(x)\{x - H(c(x)^-)\} & x \in (0,1) \\ \Pi_a(\Theta) & x = 1 \end{cases} \tag{1}$$

is said to be the concentration function of Π with respect to (w.r.t.) Π_0.
Observe that

$$\varphi(x) = \begin{cases} \Pi(L(x)) & x = H(c(x)) = \Pi_0(L(x)) \\ \Pi(L^-(x)) & x = H(c(x)^-) = \Pi_0(L^-(x)) \end{cases}$$

while $\varphi(x)$ is defined by linear interpolation on $\{x : H(c(x)^-) < x < H(c(x))\}$, if it is not empty. Furthermore, as proved in Cifarelli and Regazzini (1987), $\varphi(x)$ is a nondecreasing, continuous and convex function such that $\varphi(x) \equiv 0 \Leftrightarrow \Pi \perp \Pi_0, \varphi(x) = x\ \forall x \in [0,1] \Leftrightarrow \Pi = \Pi_0$ and

$$\varphi(x) = \int_0^{c(x)} \{x - H(t)\}dt = \int_0^x c(t)dt. \tag{2}$$

An interesting interpretation of the concentration function is provided by the following Theorem, due to Cifarelli and Regazzini (1987).

Theorem 1. If $A \in \mathcal{F}$, $\Pi_0(A) = x$, then $\varphi(x) \leq \Pi_a(A)$. Moreover if $x \in [0,1]$ is adherent to the range of H, then B_x exists such that $\Pi_0(B_x) = x$ and

$$\varphi(x) = \Pi_a(B_x) = \min\{\Pi(A) : A \in \mathcal{F} \text{ and } \Pi_0(A) \geq x\}. \tag{3}$$

If Π_0 is nonatomic, then (3) holds for any $x \in [0,1]$.

Therefore, the concentration function of Π w.r.t. Π_0 synthesizes the discrepancy between Π and Π_0 on the sets with the same measure under Π_0; in fact, for nonatomic Π_0 or for x adherent to the range of H,

$$x - \varphi(x) = \sup_{\Pi_0(A)=x} \{\Pi_0(A) - \Pi(A)\} = \sup_{\Pi_0(B)=1-x} \{\Pi(B) - \Pi_0(B)\}.$$

Such a result has been throughly explored in Fortini and Ruggeri (1990), where the concentration functions were applied in a robust Bayesian analysis to compare a class of ε-contaminated priors with a reference probability measure.

Consider now the class $\mathcal{P}$ of all the probability measures on $(\Theta,\mathcal{F})$; it makes sense to compare them in terms of their concentrations w.r.t. a fixed measure Π_0 in $\mathcal{P}$. In this case the concentration function of Π w.r.t. Π_0 will be denoted by $\varphi(\Pi,x)$, to stress the dependence on Π. From Theorem 1, it follows that the smaller $\varphi(\Pi,x)$ is, the greater the concentration of Π w.r.t. Π_0.

The comparison of probability measures is also possible when introducing a partial ordering in the space $\mathcal{P}$.

Definition 2. If $\varphi(\Pi_2,x) < \varphi(\Pi_1,x)$ $\forall x \in [0,1]$, we will say that Π_2 is not less concentrated than Π_1 w.r.t. Π_0. Afterwards we will denote it by $\Pi_1 \leqslant \Pi_2$.

As stated in the next Theorem, due to Regazzini (1992), the previous partial ordering is equivalent to the one induced by the class of indices, considered in Ali and Silvey (1966) and Csiszár (1967),

$$\varrho(\Pi,g) = \int_{[0,\infty)} g(t)dH_\Pi(t) + \Pi_s(\Theta)\lim_{t\to\infty} \{g(t)/t\}$$

where $g : [0,\infty) \to \Re$ is continuous and convex, while H_Π and Π_s are defined as before, for any $\Pi \in \mathcal{P}$, with respect to a fixed $\Pi_0 \in \mathcal{P}$.

Theorem 2. For any pair of probability measures Π_1, $\Pi_2 \in \mathcal{P}$, $\Pi_1 \leqslant \Pi_2$ holds w.r.t. Π_0 if and only if $\varrho(\Pi_1,g) \leqslant \varrho(\Pi_2,g)$ for all continuous, convex g for which $\varrho(\Pi_1,g)$ and $\varrho(\Pi_2,g)$ are finite.

Observe that the well-known Gini's concentration ratio (Gini, 1914) $C(\Pi) = 2 \int_0^1 \{x - \varphi(x)\}\, dx$ and the index $G(\Pi) = \sup_{x\in[0,1]} \{x - \varphi(x)\}$ proposed by Pietra (1915), which is equal to twice the variational distance $\sup_{A\in F} |\Pi(A) - \Pi_0(A)|$, are obtained as particular cases of $\varrho(\Pi,g)$, taking, respectively,

$$g(t) = 1/2 \int_{\Re} |t - u|dH_\Pi(u) + 1/2\Pi_S(\Theta) \text{ and } g(t) = |t - 1|.$$

3. Signed Measures

Cifarelli and Regazzini (1987) gave a general definition of concentration function connected to the classical Lorenz-Gini curve defined for statistical variable taking only nonnegative values. Many authors (e.g. Wold (1935), Castellano (1938), Michetti and Dall'Aglio (1957)) have considered variables assuming also negative values, e.g. when interested in the concentration of the gain of some industrial categories. Besides, signed measures are worthwhile when considering slight changes, even infinitesimal, in probability measures, like in local sensitivity analysis in Bayesian robustness (see Ruggeri and Wasserman, 1993, and Fortini and Ruggeri, 1992, about infinitesimal properties of the concentration function). In this paper, all the concepts introduced in the previous Section about the concentration function and the coefficients of divergence between probability measures are extended to bounded signed measures. Further results about them are proved and applied to some examples.

Let Π_0 be a probability measure and Π a bounded signed measure on the same measurable space Θ (observe that the extension to any positive measure Π_0 is straightforward). Then, from the Jordan decomposition (see Kolmogorov and Fomin, 1980, p. 347) there exist two positive measures Π^+ and Π^- such that $\Pi = \Pi^+ - \Pi^-$.

The notion of concentration function of Π with respect to Π_0 can be introduced as follows. Let h be the Radon-Nikodym derivative of Π with respect to Π_0 (see Ash, 1972, p. 68). Set $h(\vartheta) = +\infty$ all over the subset N^+ where $\Pi \equiv \Pi^+$ and Π is singular with respect to Π_0 and set $h(\vartheta) = -\infty$ all over the subset N^- where $\Pi \equiv \Pi^-$ and Π is singular with respect to Π_0. Let $H(y)$, $c(x)$, $L(x)$, $L^-(x)$ be as in Section 2, while the definition of the concentration function slightly differs from (1).

Definition 3. The function $\varphi : [0,1] \to [-\Pi^-(\Theta), \Pi_a(\Theta) - \Pi_s^-(\Theta)]$ defined by

$$\varphi(x) = \begin{cases} -\Pi_s^-(\Theta) & x = 0 \\ \Pi(L^-(x)) + c(x)\{x - H(c(x)^-)\} & x \in (0,1) \\ \Pi_a(\Theta) - \Pi_s^-(\Theta)) & x = 1 \end{cases} \qquad (4)$$

is said to be the concentration function of Π w.r.t. Π_0.

It can be easily that (2) becomes

$$\varphi(x) = xc(x) - \int_{-\infty}^{c(x)} H(t)dt - \Pi_s^-(\Theta) = \int_0^x c(t)dt - \Pi_s^-(\Theta). \qquad (5)$$

This is obvious if $x = 0,1$ while, if $x \in (0,1)$,

$$\varphi(x) + \Pi_s^-(\Theta) = \int_{-\infty}^{c(x)-} t dH(t) + c(x)\{x - H(c(x)^-)\} = xc(x) - \int_{-\infty}^{c(x)} H(t)dt.$$

The concentration function of a signed measure Π can be expressed in terms of the concentration functions of the positive measures Π^+ and Π^-, as shown in Theorem 3. To prove it, a lemma is needed; it should be noted that it has its own interest, because it gives the relation between $\varphi(\Pi,x)$ and $\varphi(-\Pi,x)$.

Lemma 1. $\varphi(-\Pi,x) = \varphi(\Pi,1 - x) - \Pi(\Theta)$ *for any* $x \in [0,1]$.

Proof. Because of their definitions, it follows that

$$H_{-\Pi}(y) = 1 - H_\Pi(-y) + \Pi_0(\{\vartheta \in \Theta : h_\Pi(\vartheta) = -y\}) \text{ and}$$
$$c_{-\Pi}(x) = -c_\Pi(1 - x) - v(1 - x),$$

where $v(1 - x) = \sup\{z \in \Re : H_\Pi(z^-) \leq 1 - x\} - c_\Pi(1 - x)$ and the meaning of the subscripts is evident. Applying (5), it follows that

$$\varphi(-\Pi,x) = -\varphi(\Pi,1) + \varphi(\Pi,1 - x) - \Pi_s^+(\Theta) - E_x,$$

where $E_x = \int_0^x v(1 - t)dt$.

Taking $x = 1$, it follows that $E_1 = 0$ so that the thesis is proved because $E_x = 0$, since $v(1 - x) \geq 0$ for any $x \in [0,1]$.

Theorem 3. Given the Jordan decomposition $\Pi = \Pi^+ - \Pi^-$, *it follows that*

$$\varphi(\Pi,x) = \varphi(\Pi^+,x) + \varphi(\Pi^-,1 - x) - \Pi^-(\Theta).$$

Proof. Let $h(\omega)$ be decomposed into its positive and negative parts $h = h^+ - h^-$ and take $\underline{h}^- = -h^-$ and $\underline{\Pi}^- = -\Pi^-$. Denote $H(t)$ and $c(x)$ corresponding to Π^+ and Π^- with the subscripts + and −, respectively. It follows that

$$H(t) = \begin{cases} H^+(t) & t \geq 0 \\ H^-(t) & t < 0 \end{cases}$$

and

$$c(x) = \begin{cases} c^-(x) & x \leqslant H^-(0^-) \\ 0 & H^-(0^-) < x \leqslant H^+(0) \\ c^+(x) & x > H^+(0) \end{cases}$$

Because of (5), it results that

$$\varphi(\Pi,x) = \begin{cases} \varphi(\tilde{\Pi}^-,x) & x \leqslant H^-(0^-) \\ \varphi(\Pi^+,x) + \tilde{\Pi}^-(\Theta) & x > H^-(0^-) \end{cases}$$

so that $\varphi(\Pi,x) = \varphi(\Pi^+,x) + \varphi(\tilde{\Pi}^-,x)$ for any $x \in [0,1]$, while the thesis is a consequence of Lemma 1.

Because of Theorem 3, it is possible to apply the results in Cifarelli and Regazzini (1987) to $\varphi(\Pi^+,x)$ and $\varphi(\Pi^-,1 - x)$, so that φ is a continuou, convex function, but not necessarily nondecreasing as before; moreover Theorem 1 still holds, provided that $\Pi_0(A) = x$ is substituted in (3).

Consider now the space $\mathcal{M}_\alpha$ of the signed measures Π on Θ such that $\Pi(\Theta) = \alpha$, $\alpha \in \Re$. The concentration function induces a partial ordering in such a space: for any Π_1 and Π_2 in $\mathcal{M}_\alpha$, $\Pi_1 \leqslant \Pi_2$ if and only if $\varphi_1 \geqslant \varphi_2$, where φ_1 and φ_2 are the concentration function of Π_1 and Π_2 w.r.t. Π_0, respectively.

Define, for any $\Pi \in \mathcal{M}_\alpha$, the extension of the Ali-Silvey index:

$$\varrho(\Pi,g) = \int_{\Re} g(t)dH_\Pi(t) + \Pi_s^+(\Theta) \lim_{t\to+\infty} \{g(t)/t\} - \Pi_s^-(\Theta) \lim_{t\to-\infty} \{g(t)/t\} \quad (6)$$

where $g : \Re \to \Re$ is continuous and convex. Such an index can be interpreted as a coefficient of divergence $d(\Pi,\Pi_0)$ of Π from Π_0, satisfying four basic properties which are the natural extension of the Ali-Silvey's (1966) ones. Such properties are:

P.1. The coefficient $d(\Pi,\Pi_0)$ should be defined for all pairs of measures Π and Π_0 on $(\Theta,\mathcal{F})$.

P.2. Let $\vartheta = t(\vartheta)$ be a measurable transformation from $(\Theta,\mathcal{F})$ onto the measurable space $(\Omega,\mathcal{G})$. Then, it should follow that $d(\Pi,\Pi_0) \geqslant d(\Pi t^{-1},\Pi_0 t^{-1})$, where Πt^{-1} and $\Pi_0 t^{-1}$ are the induced measures on Ω corresponding, respectively, to Π and Π_0.

P.3. The coefficient $d(\Pi,\Pi_0)$ should be minimised when $\Pi \equiv \Pi_0$ and maximised, among all the measures Π sharing the same $\Pi^+(\Theta)$ and $\Pi^-(\Theta)$, when $\Pi \perp \Pi_0$.

P.4. Let $\{\Pi_\omega;\ \omega \in (a,b) \subseteq \Re\}$ be a family of equivalent (mutually absolutely continuous) measures on $\Re$ such that the family of densities $\pi_\omega(\vartheta)$ with respect to a fixed measure μ has monotone likelihood ratio in ϑ (see Lehmann, 1986, p. 78). Given $a < \omega_1 < \omega_2 < \omega_3 < b$, it should follow that $d(\Pi_{\omega 1},\Pi_{\omega 2}) \leqslant d(\Pi_{\omega 1},\Pi_{\omega 3})$.

The detailed proofs of the index (6) satisfying the properties are omitted, since they are very similar to the Ali and Silvey's ones, and just few changes are mentioned.

In proving the property P.1, it should be observed that $g(x) \geqslant (x - a)/(b - a) \cdot (g(b) - g(a)) + g(a)$ holds for any $x \notin [a,b]$ and that $\lim_{t \to -\infty} \{g(t)/t\} < \infty$.

The bound $g(x + y) \leqslant g(y) + x \lim_{x \to -\infty} \{g(x)/x\}$, for $x < 0$, is used in the proofs of the properties P.2 and, for $y = 0$, P.3. Finally, P.4 holds, provided that the monotone likelihood ratio is defined also for signed measure.

Like the concentration function, the index (6) can be split in two parts too, corresponding to the Jordan decomposition $\Pi = \Pi^+ - \Pi^-$. In the next Theorem, the index (6) should be better denoted $\varrho(\Pi,g(t))$.

Theorem 4. Given the Jordan decomposition $\Pi = \Pi^+ - \Pi^-$, it follows that

$$\varrho(\Pi,g(t)) = \varrho(\Pi^+,g(t)) + \varrho(\Pi^-,g(-t)) = \varrho(\Pi^+,g(t)) - \varrho(-\Pi^-,g(t)).$$

Proof. Let H^+ and H^- be defined for Π^+ and Π^-, respectively, so that

$$H(t) = \begin{cases} H^+(t) & t \geqslant 0 \\ 1 - H^-(-t) + \Pi_0(\{\vartheta \in \Theta : h_\Pi(\vartheta) = -t\}) & t < 0 \end{cases}$$

Because of Lemma 1 and the definition of the index (6), the result can be proved.

As another property of the index (6), it should be remarked that the corresonding partial ordering coincides with that induced by the concentration function, as proved by Regazzini (1992) about probability measures.

Lemma 2. Let Π_1 and Π_2 be signed measures such that $\Pi_1(\Theta) = \Pi_2(\Theta)$; let h_1, h_2 be the Radon-Nikodym derivatives of Π_1 and Π_2 with respect to Π_0; let H_1 and H_2 be their distribution functions and c_1, c_2 their quantile functions. If $\Pi_1 \leqslant \Pi_2$, then, for any $x \in \Re$, $\Pi^-_{2s}(\Theta) - \Pi^-_{1s}(\Theta) + \int_{-\infty}^{x} \{H_2(t) - H_1(t)\}\, dt \geqslant 0$.

Proof. Let $x \in \Re$ and $y = H_1(x)$. Since

$$\varphi_i(y) = -\Pi^-_{is}(\Theta) + yc_i(y) - \int_{-\infty}^{c_i(y)} H_i(t)dt,\ i = 1,2$$

then

$$(\Pi^-_{2s} - \Pi^-_{1s})\,(\Theta) + \int_{-\infty}^{c_2(y)} H_2(t)dt - \int_{-\infty}^{c_1(y)} H_1(t)dt \geqslant y(c_2(y) - c_1(y)).$$

Therefore

$$(\Pi^-_{2s} - \Pi^-_{1s})\,(\Theta) + \int_{-\infty}^{c_1(y)} \{H_2(t) - H_1(t)\}dt \geqslant y(c_2(y) - c_1(y)) - \int_{c_1(y)}^{c_2(y)} H_2(t)dt.$$

It follows

$$(\Pi^-_{2s} - \Pi^-_{1s})\,(\Theta) + \int_{-\infty}^{x} \{H_2(t) - H_1(t)\}dt \geqslant$$
$$\geqslant \int_{0}^{c_1(y)} \{H_2(t) - H_1(t)\}dt + \int_{c_1(y)}^{x} \{H_2(t) - H_1(t)\}dt \geqslant$$
$$\geqslant y(c_2(y) - c_1(y)) - \int_{c_1(y)}^{c_2(y)} H_2(t)dt + \int_{c_1(y)}^{x} \{H_2(t) - H_1(t)\}dt \geqslant 0.$$

Theorem 5. Under the same hypotheses as Lemma 2, $\Pi_1 \leqslant \Pi_2$ if and only if $\varrho(\Pi_1,g) \leqslant \varrho(\Pi_2,g)$ for all continuous, convex g for which $\varrho(\Pi_1,g)$ and $\varrho(\Pi_2,g)$ are finite.

Proof. The argument partially follows the one in Regazzini (1992).

Necessity. Since g is continuous and convex, there exists a non-decreasing function γ on $\Re$ such that $g(y) = g(a) + \int_a^y \gamma(t)dt$, for $a,\ y \in \Re$. Since

$$\int_{\Re}\int_{a}^{y} \gamma(t)dtd(H_2 - H_1)\,(y) = \int_{\Re} \gamma(y)\{H_1(y) - H_2(y)\}dy$$
$$= \left[\gamma(y)\int_{-\infty}^{y} \{H_1(t) - H_2(t)\}dt\right]_{-\infty}^{+\infty} - \int_{\Re}\int_{-\infty}^{y} \{H_1(t) - H_2(t)\}dtd\gamma(y)$$

then

$$\varrho(\Pi_2,g) - \varrho(\Pi_1,g) =$$

$$\int_{\Re} g(t)\, d(H_2 - H_1)(t) + (\Pi^+_{2s} - \Pi^+_{1s})(\Theta) \lim_{y\to+\infty} g(y)/y -$$

$$(\Pi^-_{2s} - \Pi^-_{1s})(\Theta) \lim_{y \to -\infty} g(y)/y$$

$$= (\Pi_{2a} - \Pi_{1a})(\Theta) \lim_{y \to +\infty} \gamma(y) + (\Pi_{2s} - \Pi_{1s})(\Theta) \lim_{y \to +\infty} g(y)/y$$

$$+ \int_{\Re} \left\{ (H^-_{2s} - \Pi^-_{1s})(\Theta) + \int_{-\infty}^{y} (H_2(t) - H_1(t))dt \right\} d\gamma(y) =$$

$$= \int_{\Re} \left\{ (\Pi^-_{2s} - \Pi^-_{1s})(\Theta) + \int_{-\infty}^{y} \{H_2(t) - H_1(t)\}dt \right\} d\gamma \geqslant 0$$

from Lemma 2, $\Pi_2(\Theta) = \Pi_1(\Theta), \lim_{y \to \pm\infty} \{g(y)/y - \gamma(y)\} = 0$ and

$$\int_{-\infty}^{0} \mathrm{ydH_i(y)} = \int_{\{\mathrm{h_i}(\vartheta)<0\}} h_i(\vartheta)\Pi_0(d\vartheta) = \int_{\{h_i(\vartheta)<0\}} \Pi_i(d\vartheta)$$

$$\int_{0}^{\infty} ydH_i(y) = \int_{\{h_i(\vartheta)>0\}} h_i(\vartheta)\Pi_0(d\vartheta) = \int_{\{h_i(\vartheta)>0\}} \Pi_i(d\vartheta).$$

Sufficiency. For $y \in (0,1)$ belonging to the range of H_1, let $g(t) = |t - c_2(y)|$. Then

$$\varrho(\Pi_1,g) = \int_{\Re} |t - c_2(y)|dH_1(t) + \Pi^+_{1s}(\Theta) + \Pi^-_{1s}(\Theta) \geqslant$$

$$\geqslant \int_{-\infty}^{c_1(y)} (c_2(y) - t)dH_1(t) + \int_{c_1(y)}^{+\infty} (t - c_2(y))dH_1(t) + \Pi_{1s}(\Theta) + 2\Pi^-_{1s}(\Theta) \geqslant$$

$$\geqslant 2yc_2(y) - c_2(y) + \Pi_1(\Theta) - 2\int_{-\infty}^{c_2(y)} tdH_1(t) + 2\Pi^-_{1s}(\Theta).$$

On the other hand

$$\varrho(\Pi_1,g) \leqslant \varrho(\Pi_2,g) \leqslant \int_{\Re} |t - c_2(y)|dH_2(t) + \Pi^+_{2s}(\Theta) + \Pi^-_{2s}(\Theta) \leqslant$$

$$\leqslant 2yc_2(y) - c_2(y) + \Pi_2(\Theta) - 2\int_{-\infty}^{c_2(y)} tdH_2(t) + 2\Pi^-_{2s}(\Theta).$$

It follows that

$$-\Pi^-_{2s}(\Theta)\int_{-\infty}^{c_2(y)} tdH_2(t) \leqslant -\Pi^-_{1s}(\Theta)\int_{-\infty}^{c_2(y)} tdH_1(t)$$

and, hence

$$\varphi_2(x) \leqslant \varphi_1(x) \tag{7}$$

holds for any x belonging to the range of H_1. Since φ_1 and φ_2 are continuous and convex functions on $[0,1]$, and φ_1 is linear on $(H_1(t^-),H_1(t))$, (7) holds necessarily for all $x \in [0,1]$.

Setting $g(t) = |t - 1|$, the index (6) becomes the Pietra's index $G(\Pi)$ as before and it coincides with twice the variational distance between Π and Π_0, while

$$g(t) = 1/2 \left\{ \int_{\mathfrak{R}} |t - u| dH_\Pi(u) + \Pi_s^+(\Theta) + \Pi_s^-(\Theta) \right\} + 1 - \Pi(\Theta)$$

gives the Gini's concentration ratio $C(\Pi)$.

Proposition 1. Given $g(t) = |t - 1|$, it follows that

$$\varrho(\Pi,g) = 2 \sup_{x \in [0,1]} \{x - \varphi(x)\} + \Pi(\Theta) - 1 = 2\left\{1/2 \int_{\Theta} |\pi - \pi_0| d\mu\right\},$$

where π and π_0 are the densities, respectively, of Π and Π_0 with respect to a dominating measure μ and $1/2 \int_\Theta |\pi - \pi_0| d\mu = \sup_{A \in \mathcal{F}} |\Pi(A) - \Pi_0(A)|$ is the variational distance between Π and Π_0 (cf. Strasser, 1985, Definition 2.1 and Lemma 2.4).

Proof. Substituting $g(t) = |t - 1|$ in (6), it follows that

$$\varrho(\Pi,g) = 2\int_{-\infty}^{1} (1-t)dH_\Pi(t) + \int_{\mathfrak{R}} (t-1)dH_\Pi(t) + \Pi_s^+(\Theta) + \Pi_s^-(\Theta)$$

$$= 2\left\{H_\Pi(1) - \int_{-\infty}^{1} t dH_\Pi(t)\right\} + \Pi_a(\Theta) - 1 + \Pi_s(\Theta) + 2\Pi_s^-(\Theta) \tag{8}$$

$$= 2\{\hat{x} - \varphi(\hat{x})\} + \Pi(\Theta) - 1,$$

where $\hat{x} = H_\Pi(1)$ is such that $c(\hat{x}) \leqslant 1 \leqslant c(\hat{x}^+)$. Since $\varphi(x)$ is a.e. differentiable, it follows from (5) that $\hat{x}$ maximise $x - \varphi(x)$ all over $[0,1]$. The second part of the Proposition is proved considering the three possible situations:

1) $c(y) > 1$, $\forall y \in \mathfrak{R}$. In this case, $\pi(\vartheta) > \pi_0(\vartheta)$ a.e. and $\hat{x} = 0$, so that

$$\varrho(\Pi,g) = \Pi_a(\Theta) - 1 + \Pi_s^+(\Theta) + \Pi_s^-(\Theta) = 2\left\{1/2 \int_\Theta |\pi - \pi_0| d\mu\right\}.$$

2) $c(y) < 1$, $\forall y \in \Re$. Now, $\pi(\vartheta) < \pi_0(\vartheta)$ a.e. and $\ell = 1$. Since here $\int_{-\infty}^{1} t dH_\Pi(t) = \Pi_a(\Theta)$ and $H_\Pi(1) = 1$, it follows from (8) that

$$\varrho(\Pi,g) = 1 - \Pi_a(\Theta) + \Pi_s^+(\Theta) + \Pi_s^-(\Theta) = 2\left\{1/2 \int_\Theta |\pi - \pi_0| d\mu\right\}.$$

3) $\lim_{y\to 0} c(y) \leq 1 \leq \lim_{y\to 1} c(y)$. Suppose that $c(\ell) = 1$; otherwise $L^-(1)$ should substitute $L(1)$ in the next proof. It follows that

$$\varrho(\Pi,g) = \int_{L(1)} \{\pi_0(\vartheta) - \pi(\vartheta)\} d\mu + \int_{[L(1)]^c} \{\pi(\vartheta) - \pi_0(\vartheta)\} d\mu =$$
$$2\left\{1/2 \int_\Theta |\pi - \pi_0| d\mu\right\}.$$

Proposition 2. Given

$$g(t) = 1/2 \left\{ \int_\Re |t - u| dH_\Pi(u) + \Pi_s^+(\Theta) + \Pi_s^-(\Theta)\right\} + 1 - \Pi(\Theta),$$

it follows that

$$\varrho(\Pi,g) = 2\int_0^1 \{x - \varphi(x)\} dx = 1/2\Delta_H + 1 - \Pi(\Theta) + \Pi_s^+(\Theta) + \Pi_s^-(\Theta), \quad (9)$$

where $\Delta_H = 2\int_\Re H_\Pi(t)\{1 - H_\Pi(t)\} dt$ *is the mean difference of* H_Π.

Proof. Observe that $g(t)$ equals

$$1/2 \left\{2\int_{-\infty}^{t} (t-u) dH_\Pi(u) + \int_\Re (u-t) dH_\Pi(u) + \Pi_s^+(\Theta) + \Pi_s^-(\Theta)\right\} + 1 - \Pi(\Theta) \quad (10)$$

which becomes, after integrating the first integral by parts and computing the second one,

$$\int_{-\infty}^{t} H_\Pi(u) du + 1 - t/2 - 1/2\, \Pi(\Theta) + \Pi_s^-(\Theta),$$

so that $g(t) \sim -t/2$ and $g(t)/t \to -1/2$ as $t \to -\infty$ while $g(t) \sim t/2$ from (10) and $g(t)/t \to 1/2$ as $t \to \infty$.

Substituting the above limits in $\varrho(\Pi,g)$, then (9) is obtained after some computations, including another integration by parts. Because of (5), it follows that $2\int_0^1 \{x - \varphi(x)\}dx$ equals

$$1 + 2\Pi_s^-(\Theta) - 2\int_0^1 xc(x)dx + 2\int_0^1\int_{-\infty}^{c(x)} H_\Pi(t)dtdx.$$

Taking $x = H_\Pi(t)$, then $c(x) = t + \{t^* - t\}$, where $t^* = inf\{z : H_\Pi(z) = H_\Pi(t)\}$, and $\int_t^{t^*} dH_\Pi(u) = 0$. The proof is complete, observing that

$$2\int_0^1 xc(x)dx = 1/2\ \Delta_H + \Pi_a(\Theta) \text{ and } 2\int_0^1\int_{-\infty}^{c(x)} H_\Pi(t)dtdx = \Delta_H.$$

It should be noted that not all the indices, defined through the Ali-Silvey index for the probability measures, can be extended to the signed measures. As an example, consider the Kullback-Leibler's index which is obtained from $\varrho(\Pi,g)$ setting $g(x) = x \log x$, which cannot be extended to negative x preserving both continuity and convexity.

4. Examples

Two examples are now presented; the former has essentially a mathematical interest because it gives the analytical expression and the plot of the concentration function when the measures are absolutely continuous with respect to the Lebesgue measure; the latter case is more interesting, from a statistical viewpoint, because it compares the spreads of people's natural movement and of the population among the Italian regions, confirming the different behaviour across the country.

Example 1. Let Π_0 have density $\pi_0(\vartheta) = 6\vartheta(1 - \vartheta)$, for $\vartheta \in [0,1]$ while the density of Π is

$$\pi(\vartheta) = \begin{cases} 3/2 & \vartheta \in [1/12, 11/12) \\ -3/2 & \vartheta \in [0,1/12) \cup [11/12, 1] \end{cases}$$

The concentration function of Π with respect to Π_0 is shown in Fig. 1, and it

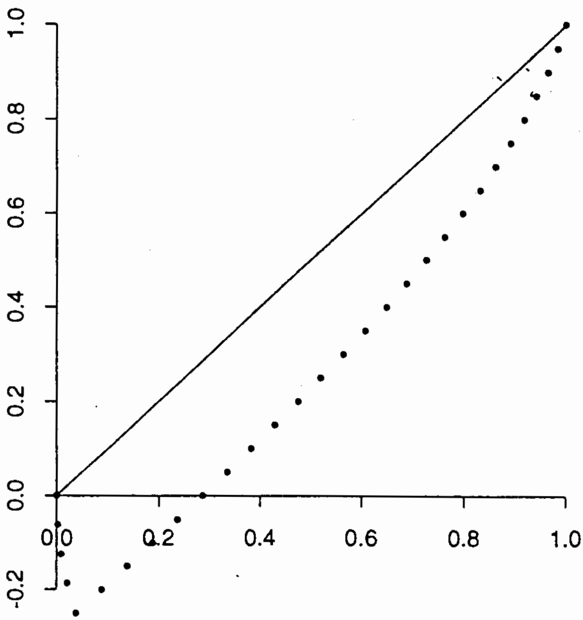

Fig. 1 – Concentration function of the signed measure Π w.r.t. $\Pi_0 \sim Beta(2,2)$ $(\cdots)$ compared with the minimum concentration curve (—).

is given in a parametric form by $x = 6\alpha^2 - 4\alpha^3$ and $\varphi(x) = -3\alpha$ when $0 \leqslant \alpha \leqslant 1/12$ while $x = 449/432 - 6\alpha^2 + 4\alpha^3$ and $\varphi(x) = 5/4 - 3\alpha$ when $1/12 < \alpha < 1/2$.

Example 2. The data in Table 1 (extract from Table 2.17 in the Annuario Statistico Italiano by Istituto Centrale di Statistica, 1989) refer to the natural movement of the resident population in the twenty Italian regions during the year 1988; such a movement is given by the difference between the births and the deaths.

It could be observed that the increase in population in each region is not proportional to the number of resident people; furthermore, the deaths are more than the births in many regions, all of them from Northern and Central Italy. To analyse such a phenomenon, two statistical variables are defined on the space of the Italian regions: the former assigns them the population and the latter its increase, both of them expressed by their percentage with respect to the national figures. Afterwards, the variables are compared using the concentration function, when the first of them is taken as the reference

Table 1
Births and deaths in the Italian Regions in 1988

REGIONS	births	deaths	difference	population
Piemonte	33,424	48,754	−15,330	4,365,911
Valle d'Aosta	1,003	1,224	−221	114,760
Lombardia	76,085	83,170	−7,085	8,898,951
Trentino-Alto Adige	9,336	8,108	1,228	884,039
Veneto	38,708	40,557	−1,849	4,380,587
Friuli-Venezia Giulia	8,646	14,963	−6,317	1,206,362
Liguria	11,246	22,733	−11,487	1,738,263
Emilia-Romagna	26,305	43,068	−16,763	3,921,281
Toscana	26,641	39,490	−12,849	3,565,280
Umbria	6,791	8,505	−1,714	819,562
Marche	12,136	14,005	−1,869	1,429,223
Lazio	50,362	43,217	7,145	5,156,053
Abruzzi	12,678	11,887	791	1,262,692
Molise	3,618	3,309	309	335,211
Campania	84,424	44,788	39,636	5,773,067
Puglia	53,337	29,794	23,543	4,059,309
Basilicata	7,671	5,547	2,124	622,658
Calabria	29,749	17,448	12,301	2,151,357
Sicilia	68,895	44,393	24,502	5,164,266
Sardegna	16,801	12,585	4,216	1,655,859
Italia	577,856	537,545	40,311	57,504,691

measure; in fact, such a concentration function can be defined and analysed according to Section 3. The concentration curve is shown in Fig. 2 and $\varphi(x)$ gives here the increase rate of the $100x\%$ of the population, resident in the regions with the smallest rate. As expected, since the concentration function in Fig. 2 is very different from the straight line, the increase rate is not proportional to the number of residents.

The data could have been analysed also using the indices (6) and observing that they become

$$\varrho(\Pi,g) = \sum_{i=1}^{20} g(h(\omega_i))\ \Pi_0(\{\omega_i\}),$$

where $\omega_i, i = 1, \ldots, 20$ denote the Italian regions. As an example, the Pietra's index confirms the discrepancy between the two measures because it becomes $G(\Pi) = 4.8085$.

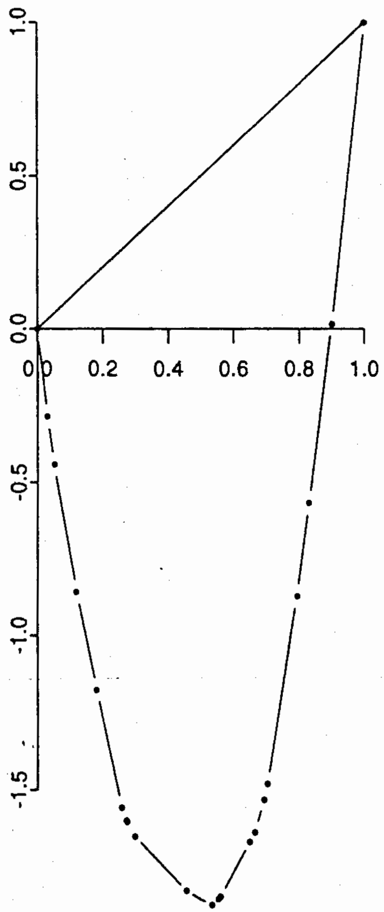

Fig. 2 – Concentration function of the natural movement w.r.t. the population in the Italian regions (— · —) compared with the minimum concentration curve (—).

5. Discussion

In this paper the concentration function and the coefficients of divergence have been defined and studied to compare a signed measure with a positive one, as suggested by actual problems (e.g. populations in which individuals might have debts). Actually, the signed measures could be considered in some problems in which only positive measures are to be compared. As an example, it is worth looking for their connections with the mean equalizing transfers (MET), studied in Regazzini (1992), that «provide an *average evaluation* of transfers which get Π to coincide with Π_0». The MET have an evident importance in many fields, like economics, where it might be relevant the study of how to move income in order to reduce social inequalities.

Comparisons among measures are also interesting when a measure is changed along a «direction», given, for example, by a signed measure with null total mass. In such a case, infinitesimal changes could be still interesting so that the Gâteaux differentials should be studied, like in Fortini and Ruggeri (1992).

The concentration function and the coefficients of divergence, have been already considered in the robust Bayesian analysis, when the posterior probability measures, from a given class of priors, are compared with a reference posterior one (e.g. Fortini and Ruggeri, 1990, 1993a, and Dey and Birmiwal, 1990). The study of their infinitesimal properties, expressed through their Gâteaux differentials, could be useful in performing a local sensitivity analysis, when posterior effects of small departures (even infinitesimal) from a given prior probability measure are studied (like in Ruggeri and Wasserman, 1991, 1993).

Besides, it could be interesting, from a mathematical point of view, to define neighbourhoods of a probability measure, which include signed measures too, extending the results in Fortini and Ruggeri (1993b). Finally, it is worth observing that the definition of concentration function between signed measures is not a straightforward extension of the definition given in this paper.

REFERENCES

ALI, S. M. and SILVEY, S. D. (1966), A general class of coefficients of divergence of one distribution from another, *J. R. Statist. Soc. B*, 28, 131-142.

ASH, R. B. (1972), *Real Analysis and Probability*, Academic Press, San Diego.

CASTELLANO, V. (1938), Sugli indici relativi di variabilità e sulla concentrazione dei caratteri con segno, *Metron*, 15, 31-49.

CIFARELLI, D. M. and REGAZZINI, E. (1987), On a general definition of concentration function, *Sankhyā B*, 49, 307-319.

CSISZÁR, I. (1967), Information-type measures of difference of probability distributions and indirect observations, *Studia Sci. Math. Hung.*, 2, 299-318.

DEY, D. K. and BIRMIWAL, L. R. (1990), Robust Bayesian analysis using entropy and divergence measures, *Technical Report*, Department of Statistics, University of Connecticut, Storrs.

FORTINI, S. and RUGGERI, F. (1990), Concentration function in a robust Bayesian framework, *Quaderno IAMI* 90.6. Milano: CNR-IAMI.

FORTINI, S. and RUGGERI, F. (1992), Infinitesimal properties of the concentration function, *Unpublished Manuscript*.

FORTINI, S. and RUGGERI, F. (1993a), Concentration functions and Bayesian robustness. To appear in *J. Statist. Plann. Inference*.

FORTINI, S. and RUGGERI, F. (1993b), On defining neighbourhoods of measures through the concentration function. To appear in *Sankhyā A*.

GINI, C. (1914), Sulla misura della concentrazione della variabilità dei caratteri, *Atti del Reale Istituto Veneto di S.L.A., A.A. 1913-1914*, 73, parte II, 1203-1248.

ISTITUTO CENTRALE DI STATISTICA (1989), *Annuario Statistico Italiano*, ISTAT, Roma.

KOLMOGOROV, A. N. and FOMIN, S. V. (1980), *Elementi di teoria delle funzioni e di analisi funzionale*, Mir, Moscow.

LEHMANN, E. L. (1986), *Testing Statistical Hypotheses*, Wiley, New York.

MARSHALL, A. W. and OLKIN, I. (1979), *Inequalities: theory of majorization and its applications*, Academic Press, New York.

MICHETTI, B. and DALL'AGLIO, G. (1957), La differenza semplice media, *Statistica*, 2, 159-255.

PIETRA, G. (1915), Delle relazioni tra gli indici di variabilità, *Atti del Reale Istituto Veneto di S.L.A. A.A. 1914-1915*, 74, parte II, 775-792.

REGAZZINI, E. (1992), Concentration comparisons between probability measures. To appear in *Sankhyā B*, 54.

RUGGERI, F. and WASSERMAN, L. A. (1991), Density based classes of priors: infinitesimal properties and approximations, *Technical Report # 528*, Department of Statistics, Canergie Mellon University, Pittsburgh.

RUGGERI, F. and WASSERMAN, L. A. (1993), Infinitesimal sensitivity of posterior distribution. To appear in *Canad. J. Statist.*

STRASSER, H. (1985), *Mathematical Theory of Statistics*, Walter de Gruyter, Berlin.

WOLD, H. (1935), A study on the mean difference, concentration curves and concentration ratio, *Metron*, 12, 39-58.