Sankhyā: The Indian Journal of Statistics
1997, Volume 59, Series A, Pt. 3, pp. 345-364

DIFFERENTIAL PROPERTIES OF THE CONCENTRATION FUNCTION

By SANDRA FORTINI
Università "L. Bocconi", Milano
and
FABRIZIO RUGGERI
CNR-IAMI, Milano

SUMMARY. The Lorenz curve and its extensions, as the concentration function, are used in some statistical problems, mainly related to economics, to describe the discrepancies between a probability measure P and a reference measure P_0. Some applications require the consideration of variations in the concentration of P with respect to P_0 due to small changes of P. This leads to the study of the differential properties of the concentration function at a point, regarded as a functional defined on the space of the finite signed measures. This paper proves that the Gâteaux differential of the concentration function exists in any direction and it provides its expression. Moreover, it shows that the concentration function is not Fréchet differentiable. Some applications are described.

1. Introduction

Classical analysis of inequality describes the tendency of a transferable quantity to be concentrated on few individuals. This analysis is traditionally carried out using the Lorenz curve, which describes the discrepancy between a probability measure P, which refers to the distribution of the transferable quantity, and the uniform distribution.

Marshall and Olkin (1979, p.5) give the following definition of Lorenz concentration curve (also known as the Lorenz-Gini curve):
"Consider a population of n individuals, and let x_i be the wealth of individual $i, i = 1, \ldots, n$. Order the individuals from poorest to richest to obtain $x_{(1)}, \ldots, x_{(n)}$. Now plot the points $(k/n, S_k/S_n), k = 0, \ldots, n$, where $S_0 = 0$ and $S_k = \sum_{i=1}^{k} x_{(i)}$ is the total wealth of the poorest k individuals in the population.

Paper received. August 1996; revised July 1997.
AMS (1991) *subject classifications.* Primary 62E10; secondary 60E05.
Key words and phrases. Concentration function; Ali-Silvey indices; Gâteaux differential.

Join these points by line segments to obtain a curve connecting the origin with the point $(1,1)$. Notice that if total wealth is uniformly distributed in the population, then the Lorenz curve is a straight line. Otherwise, the curve is convex and lies under the straight line."

Therefore, the classical definition of concentration refers to the discrepancy between a probability P, which gives mass $x_{(i)}/S_n$ to $\theta_i, i = 1,\dots,n$, and the uniform distribution P_0 on $\Theta = \{\theta_1,\dots,\theta_n\}$. Cifarelli and Regazzini (1987) defined the concentration function of P with respect to P_0, where P and P_0 are two probability measures on the same measurable space $(\Omega,\mathcal{A})$; Fortini and Ruggeri (1993) extended their definition, considering a signed measure M instead of the probability measure P (see Definition 1), so that the concentration function can be considered as a functional on the vector space of finite, signed measures. From the definition of concentration function $\varphi(M,x)$, it can be shown, for any $x \in [0,1]$, that

$$\varphi(M,x) = \min\{M(A) : A \in \mathcal{A} \text{ and } P_0(A) \geq x\},$$

under very general conditions (see Theorem 1). Therefore the concentration function provides lower bounds on set probabilities.

The Lorenz curve is mainly applied in the field of economics, where it is used to study social inequalities, as determined by the wealth concentration. Actually, studies on the concentration apply to many other situations, all the more so since Cifarelli and Regazzini (1987) introduced the concentration function.

Many concrete problems suggest studying the dynamical behaviour of the concentration function for a measure subject to changes. Consider, for example, the distribution of wealth in a population and suppose that such distribution is subject to a change due, say, to a new tax policy. The corresponding variation in the concentration function accounts for the impact of the tax policy on the social inequality. If the trend of the social inequality is to be analysed, then infinitesimal increments (i.e. differentials) of the concentration function must be considered: the larger the increment (in absolute value), the faster the modification of the social inequality.

The above example suggests developing a differential calculus for the concentration function. Actually this has a wide range of practical applications; particularly interesting are those involving optimization problems. For example, suppose that an economic policy is sought to achieve a reduction of social inequality. Such a policy will act in transferring wealth from the richest to the poorest classes and questions arise about which class to start from and how to move wealth in order to reduce the inequality in the shortest time. More effective policies are those leading to larger (in absolute value) increments in the concentration function. Hence one should start from the policy for which the derivative of the concentration function is the largest. This policy consists in moving wealth from the richest class to the poorest one, leaving the others unchanged (see Example 2).

The paper is concerned with developing a differential calculus for both the concentration function and some concentration indices, viz. the Ali-Silvey indices, regarded as functions of the involved measures. As far as we know, this approach, although natural, has never been proposed before. Gâteaux derivatives are obtained for the concentration function evaluated at any fixed point x and, as well, when considering φ as a map taking values in the space of the continuous functions on the interval $[0, 1]$. An example will show that the concentration function is not Fréchet differentiable with respect to the variational norm in the space of the finite measures.

Despite the very technical proofs, presented in the Appendix, the expression of the Gâteaux derivative is quite simple and has an immediate interpretation, which makes it easy to apply in practical problems (see Examples 2 and 3).

Section 2 recalls some basic results for the concentration function and the Ali-Silvey indices, and their extensions to signed measures. Section 3 provides the expressions for the Gâteaux derivatives of both the concentration function and the Ali-Silvey indices, while some applications are discussed in Section 4.

2. Preliminaries

This Section recalls the definitions of concentration function and Ali-Silvey indices of a signed measure, along with their main properties. We refer to Fortini and Ruggeri (1993) for further details.

Let $(\Omega, \mathcal{A})$ be a measurable space. Denote by $\mathcal{M}$ the space of all the finite, signed measures on $(\Omega, \mathcal{A})$ endowed with the variational norm. The subspace of $\mathcal{M}$ consisting of all the probability measures will be denoted by $\mathcal{P}$, while $\mathcal{D}$ will be the subspace of $\mathcal{M}$ of all the signed measures λ such that $\lambda(\Omega) = 0$. Moreover, in the following, the extended real line $\overline{\mathbf{R}}$ will be considered, with the usual algebra and ordering.

Let P_0 and M be, respectively, a probability measure and a bounded, signed measure on $(\Omega, \mathcal{A})$ with $M(\Omega) = 1$. From the Jordan decomposition (cf. Ash (1972), pp. 60 and 61) there exist two positive measures M^+ and M^- such that $M = M^+ - M^-$ and a set D such that $M^+(A) = M(A \cap D)$ and $M^-(A) = -M(A \cap D^C)$ for any $A \in \mathcal{A}$. Let $|M| = M^+ + M^-$.

According to Lebesgue decomposition and Radon-Nikodym theorems (cf. Ash (1972), pp. 63 and 66) there exist a partition $\{N, N^C\} \subset \mathcal{A}$ of Ω, a real valued function h_M defined on N^C and a signed measure M_s such that

$$P_0(N) = 0, \;\; M_s^+(N) = M_s^+(\Omega), \;\; M_s^-(N) = M_s^-(\Omega)$$

and, $\forall E \in \mathcal{A}$,

$$M(E) = \int_{E \cap N^C} h_M(\omega) P_0(d\omega) + M_s(E \cap N).$$

M_s and $M_a(\,\cdot\,) = \int_{\cdot\,\cap N^C} h_M(\omega)P_0(d\omega)$ are called the singular part and the absolutely continuous part of M w.r.t. P_0. Set $h_M(\omega) = +\infty$ all over the subset $N^+ = D \cap N$ and $h_M(\omega) = -\infty$ all over the subset $N^- = D^C \cap N$. Define, for any $y \in \mathbf{R}$ and any $x \in [0,1]$, $H_M(y) = P_0(\{\omega \in \Omega : h_M(\omega) \leq y\})$, $c_M(x) = \inf\{y \in \mathbf{R} : H_M(y) \geq x\}$. Finally, let $L_M(x) = \{\omega \in \Omega : h_M(\omega) \leq c_M(x)\}$ and $L_M^-(x) = \{\omega \in \Omega : h_M(\omega) < c_M(x)\}$.

DEFINITION 1. Let M be a finite signed measure and let P_0 be a probability measure. The function $\varphi(M,\,\cdot\,)$ defined on $[0,1]$ by

$$\varphi(M,x) = \begin{cases} -M_s^-(\Omega) & x=0 \\ M\left(L_M^-(x)\right) + c_M(x)\left\{x - H_M(c_M(x)^-)\right\} - M_s^-(\Omega) & x \in (0,1) \\ M_a(\Omega) - M_s^-(\Omega) & x = 1 \end{cases}$$

is said to be the concentration function of M with respect to P_0.

The concentration function $\varphi(M,\,\cdot)$ is a continuous, convex function. The following fundamental property holds.

THEOREM 1. *For any $A \in \mathcal{A}$ such that $P_0(A) = x$, then*

$$\varphi(M,x) \leq M(A).$$

Moreover if $x \in [0,1]$ is adherent to the range of H, then $B_x \in \mathcal{A}$ exists such that $P_0(B_x) = x$ and

$$\varphi(M,x) = M(B_x) = min\left\{M(A) : A \in \mathcal{A}\ and\ P_0(A) \geq x\right\}. \qquad \ldots(1)$$

If P_0 is nonatomic, then (1) holds for any $x \in [0,1]$.

Consider now the space $\mathcal{M}_\alpha$ of signed measures M on $(\Omega, \mathcal{A})$ such that $M(\Omega) = \alpha$; $\alpha \in \Re$. From (1) it follows that the measure M is much more concentrated w.r.t. P_0 as its concentration function is far from the straight line connecting $(0,0)$ with $(1,\alpha)$. Therefore, the concentration function induces a partial ordering in $\mathcal{M}_\alpha$: given any M_1 and M_2 in $\mathcal{M}_\alpha$, then M_2 is said to be not less concentrated than M_1 if and only if $\varphi(M_1,x) \geq \varphi(M_2,x)$, for any $x \in [0,1]$.

An important class of concentration indices, also known as coefficients of divergence, can be defined.

Let $g : \mathbf{R} \to \mathbf{R}$ be a continuous and convex function such that there exist finite $g^+ = \lim_{t\to\infty}\{g(t)/t\}$ and $g^- = \lim_{t\to-\infty}\{g(t)/t\}$.

DEFINITION 2. The Ali-Silvey index of the signed measure M, relative to g, is defined by

$$\rho(M,g) = \int_{\mathbf{R}} g(t)dH_M(t) + M_s^+(\Omega)g^+ - M_s^-(\Omega)g^-. \qquad \ldots(2)$$

Ali-Silvey indices induce an ordering among probability measures which is consistent with the one induced by the concentration function itself.

3. Differential Calculus for the Concentration Function

The Fréchet and Gâteaux derivatives (see Lusternik and Sobolev (1989), pp. 215 and 221) are the most widespread derivatives of functionals and mappings.

Definition 3. Let X and Y be normed linear spaces. The Fréchet derivative of a mapping $f : X \to Y$ at a point x_0 is the linear continuous operator $\Lambda : X \to Y$ (if it exists) satisfying the condition

$$f(x_0 + h) = f(x_0) + \Lambda h + \varepsilon(h),$$

where

$$\lim_{||h|| \to 0} \frac{||\varepsilon(h)||}{||h||} = 0.$$

Definition 4. Let X and Y be linear topological spaces. The Gâteaux differential at a point x_0 and in the direction of $h \in X$ of a mapping $f : X \to Y$ is the limit (if it exists)

$$\lim_{\lambda \to 0^+} \frac{f(x_0 + \lambda h) - f(x_0)}{\lambda}.$$

If X and Y are normed linear spaces and the operator f is Fréchet differentiable, then it is also Gâteaux differentiable. The converse is not true (cf. Lusternik and Sobolev (1989), p. 222).

Given P_0 and P in $\mathcal{P}$, consider a signed finite measure $\Delta \in \mathcal{D}$. For any $x \in [0,1]$, the Gâteaux differential of $\varphi(\cdot, x)$ at the point P and in the direction of Δ is the limit

$$\varphi'_\Delta(P,x) = \lim_{\lambda \to 0^+} \left\{ \varphi(P + \lambda\Delta, x) - \varphi(P,x) \right\} / \lambda,$$

if it exists.

It will be shown that for any $x \in [0,1]$ and any $\Delta \in \mathcal{D}$, the Gâteaux derivative of $\varphi(\cdot, x)$ at the point P and in the direction of Δ exists. An expression for the Gâteaux derivative will be supplied and conditions will be found ensuring its linearity in $\Delta \in \mathcal{D}$.

Let P_0, $P \in \mathcal{P}$ and $\Delta \in \mathcal{D}$ be defined as earlier and let h_P, c_P, $L_P(x)$ and $L_P^-(x)$ be defined as in Section 2. For any fixed $x \in [0,1]$, consider the extended real function

$$h_x(\omega) = \begin{cases} -\infty & \text{if } h_P(\omega) < c_P(x) \\ h_\Delta(\omega) & \text{if } h_P(\omega) = c_P(x) \\ +\infty & \text{if } h_P(\omega) > c_P(x). \end{cases}$$

For any $x, y \in [0,1]$ and $z \in \overline{\mathbf{R}}$, let

$$H_x(z) = P_0(\{\omega \in \Omega : h_x(\omega) \le z\}),$$

$$c_x(y) = \inf\left\{z \in \overline{\mathbf{R}} : H_x(z) \ge y\right\},$$

$$L_x(y) = \{\omega \in \Omega : h_x(\omega) \le c_x(y)\}, \quad L_x^-(y) = \{\omega \in \Omega : h_x(\omega) < c_x(y)\}.$$

and

$$c_P(x)^+ = \begin{cases} \lim_{y \to \overline{x}^+} c_P(y) & \text{if } \overline{x} < 1 \\ c_P(x) & \text{if } \overline{x} = 1, \end{cases}$$

where $\overline{x} = \sup\{z \in [0,1] : c_P(z) \le c_P(x)\}$.

REMARK 1. Notice that

$$c_x(z) = \begin{cases} -\infty & \text{if } z \le \inf\limits_{y \in \mathbf{R}} H_x(y) = P_0(L_P^-(x)) \\ \inf\{y \in \mathbf{R} : H_x(y) \ge z\} & \text{if } \inf\limits_{y \in \mathbf{R}} H_x(y) < z \le \sup\limits_{y \in \mathbf{R}} H_x(y) \\ +\infty & \text{if } z > \sup\limits_{y \in \mathbf{R}} H_x(y) = P_0(L_P(x)) \end{cases}$$

and

$$h_x(\omega) = \begin{cases} -\infty & \text{if } \omega \in L_P^-(x) \\ h_\Delta(\omega) & \text{if } \omega \in L_P(x) \setminus L_P^-(x) \\ +\infty & \text{if } \omega \in (L_P(x))^C. \end{cases}$$

From now on, the notation will be simplified by avoiding, when possible, any mention of the dependence of functions on ω.

Theorem 2 states that, for most of the x's, the derivative of the concentration function equals the value of the measure Δ on the subset $L_P^-(x)$, i.e. where P is less concentrated w.r.t. P_0.

THEOREM 2. *Let P_0 and P be probability measures and let Δ be a bounded signed measure such that $\Delta(\Omega) = 0$. Then*

$$\varphi'_\Delta(P, x) = \begin{cases} \Delta(L_P^-(x)) & \text{if } x = H_P(c_P(x)^-) \\ \Delta(L_x^-(x)) + c_x(x)\{x - H_x(c_x(x)^-)\} & \text{if } x \ne H_P(c_P(x)^-). \end{cases} \qquad \ldots(3)$$

PROOF. See Appendix.

REMARK 2. It can be easily proved that $\varphi'_\Delta(P,x)$ is a continuous function of $x \in [0,1]$. Hence, if $\lambda_n \downarrow 0$, then $\varphi_{P+\lambda_n\Delta}(x) - \varphi_P(x) - \lambda_n \varphi'_\Delta(P,x)$ converges uniformly to 0 as $n \to \infty$. Consider the space $\mathcal{C}[0,1]$ of the continuous functions on $[0,1]$, with the topology of uniform convergence, and the map $\varphi : \mathcal{M} \to \mathcal{C}[0,1]$. From uniform convergence, it follows that $\varphi'_\Delta(P,\cdot)$ is the Gâteaux derivative of φ in the direction Δ evaluated at P.

REMARK 3. If $H_P(c_P(x)^-) = H_P(c_P(x))$, then (3) becomes

$$\varphi'_\Delta(P,x) = \Delta\left(L_P(x)\right).$$

so that $\varphi'_\Delta(P,x)$ is a linear, bounded functional of $\Delta \in \mathcal{D}$. Despite the linearity of $\varphi'_\Delta(P,x)$, $\varphi(P,x)$ is not Fréchet differentiable on $\mathcal{D}$ with respect to the variational norm, as shown by the following counterexample. Consider $\Omega = [0,1]$ and let $\mathcal{B}$ be the Borel σ–field of $[0,1]$ and P_0 the Lebesgue measure on $[0,1]$. Consider the measures P and Δ_ε on $(\Omega,\mathcal{B})$ with Radon-Nikodym derivative w.r.t. P_0 given by $h_P(\omega) = 1/2 + \omega$, $\omega \in [0,1]$ and

$$h_{\Delta_\varepsilon}(\omega) = \begin{cases} 1 & \omega \in [0,\varepsilon] \\ -1 & \omega \in [1-\varepsilon,1] \\ 0 & \text{elsewhere,} \end{cases}$$

respectively. Then, for any $x \in (0,1)$ and for ε sufficiently small, then $L_P(x) = [0,x]$ and $L_{P+\Delta_\varepsilon}(x) = [\varepsilon,x] \cup [1-\varepsilon,1]$. Therefore, it follows that

$$\varphi(P+\Delta_\varepsilon,x) - \varphi(P,x) - \Delta(L_P(x)) = -\varepsilon.$$

Since $\|\Delta_\varepsilon\| = \varepsilon$, $\varphi(\cdot,x)$ is not Fréchet differentiable at P.

EXAMPLE 1. Consider $\Omega = \mathbf{R}$ and let P_0 and P be, respectively, an exponential distribution $\mathcal{E}(1)$ and a Gamma one $\mathcal{G}(2,1)$. It can be easily shown that $h_P(\omega) = \omega$ for any $\omega \in \Omega$ and that $L_P(x) = [0,-\log(1-x)]$ for any $x \in [0,1]$. Consider the signed measure $\Delta = \varepsilon(P_0 - P)$, with $0 < \varepsilon < 1$, so that $P + \Delta = (1-\varepsilon)P + \varepsilon P_0$ becomes a probability measure called ε–contamination of P (see Huber, 1981). Here, it follows that $\varphi'_\Delta(P,x) = -\varepsilon(1-x)\log(1-x)$, i.e. a nonnegative function, for every $x \in [0,1]$. Therefore there is an increase in the concentration function, showing, as expected, that the modified probability measure $P + \Delta$ gets closer to P_0.

Consider the Gâteaux derivative of $\varphi(P,x)$, evaluated at P_0; since $c_x(x) = c_\Delta(x)$ for any $x \in [0,1]$, the following result holds.

COROLLARY 1. *If P_0 is a probability measure and Δ is a finite signed measure such that $\Delta(\Omega) = 0$, then $\varphi'_\Delta(P_0, x) = \varphi(\Delta, x)$.*

REMARK 4. By the definition of concentration function, it is easy to show that a stronger property holds. In fact

$$\varphi(P_0 + \Delta, x) - \varphi(P_0, x) = \varphi'_\Delta(P_0, x) = \varphi(\Delta, x).$$

Like the concentration function, the Ali-Silvey index $\rho(M, g)$ (Definition 2) is Gâteaux differentiable in any direction. We now compute its derivative.

THEOREM 3. *Let P_0, P be probability measures and let Δ be a finite signed measure such that $\Delta(\Omega) = 0$. Then*

$$\rho'_\Delta(P, g) = \int_\Omega g'_+(h_P(\omega))h_{\Delta^+}(\omega)P_0(d\omega) - \int_\Omega g'_-(h_P(\omega))h_{\Delta^-}(\omega)P_0(d\omega) +$$

$$+(\Delta_s(N) + \Delta_s^+(N^C))g^+ - \Delta_s^-(N^C)g^-,$$

where g'_+, g'_- denote the right and left derivatives of g, respectively, and $P_s(N) = P_s(\Omega)$, $P_0(N) = 0$.

PROOF. See Appendix.

REMARK 5. If g is differentiable, then

$$\rho'_\Delta(P, g) = \int_{\mathbf{R}} g'(t)dH_\Delta(d\omega) + (\Delta_s(N) + \Delta_s^+(N^C))g^+ - \Delta_s^-(N^C)g^-.$$

COROLLARY 2. *If P_0 is a probability measure and Δ is a finite signed measure such that $\Delta(\Omega) = 0$, then*

$$\rho'_\Delta(P_0, g) = \Delta_a^+(\Omega)g'_+(1) - \Delta_a^-(\Omega)g'_-(1) + \Delta_s^+(\Omega)g^+ - \Delta_s^-(\Omega)g^-.$$

5. **Applications**

The analysis of the differential properties of the concentration function allows the comparison of two measures, through their concentration, in dynamical situations when the measures themselves go through changes. In particular, functional differentials of the concentration function can be useful in solving optimization problems involving the concentration function, as illustrated by the following example.

EXAMPLE 2. Let Ω be a population and let P_0 be the distribution of the total wealth in Ω; consider the social ordering induced by P_0 among individuals in Ω. Denote by $\varphi(P_0, x)$ the concentration function of P_0 w.r.t. the uniform distribution on Ω, say U, corresponding to the ideal situation of complete equality. For simplicity, U nonatomic and P_0 absolutely continuous w.r.t. U are assumed.

Suppose that income is transferred in order to reduce inequality (e.g through tax increases and deductions). Moreover suppose that the change is brought about in several steps in such a way that, at each step, the social ordering is preserved and the fraction of the total income transferred does not exceed δ $(0 < \delta < 1)$.

It is wished to find the strategy that, at each step, achieves the maximum reduction of inequality.

If Δ_i denotes the transfer of income at the i-th step $(i = 1, \ldots, n)$, then the distribution of income after the i-th step will be denoted by

$$P_i(\underline{\Delta}_i) = P_0 + \sum_{j=1}^{i} \Delta_j,$$

where $\underline{\Delta}_i = (\Delta_1, \ldots, \Delta_i)$. Let Ψ denote the desired concentration of the income and let S_n be the set of the transfer measures $\underline{\Delta}_n$ which satisfy the following conditions:

$$(i) \quad \Delta_i \ll P_0 \qquad i = 1, \ldots, n;$$

$$(ii) \quad \varphi(P_n(\underline{\Delta}_n)) = \Psi;$$

$$(iii) \quad \|\Delta_i\| < \delta \qquad i = 1, \ldots, n;$$

$$(iv) \quad (h_{P_i(\underline{\Delta}_i)}(\omega_1) - h_{P_i(\underline{\Delta}_i)}(\omega_2))(h_{P_0}(\omega_1) - h_{P_0}(\omega_2)) \geq 0, \forall \omega_1, \omega_2 \in \Omega \qquad i = 1, \ldots, n,$$

where, for any measure M, h_M is the Radon-Nikodym derivative of M w.r.t. U. The problem consists in solving, w.r.t. $(\underline{\Delta}_n, n)$,

$$\varphi(P_i(\underline{\Delta}_i), x) = \max_{(\underline{\Delta}'_n, n) \in S} \varphi(P_i(\underline{\Delta}'_i), x) \quad \forall x \in [0, 1],$$

where $S = \{S_n \times \{n\} : n \in \mathbf{N}\}$. An approximate solution can be obtained, for small δ's, by the method of the steepest ascent, that is by solving, w.r.t. $\underline{\Delta}_i$ and n,

$$\varphi'_{\Delta_i}(P_{i-1}(\underline{\Delta}_{i-1}), x) = \max_{((\underline{\Delta}_{i-1}, \Delta'_i, \Delta'_{i+1}, \ldots, \Delta'_n), n) \in S} \varphi'_{\Delta'_i}(P_{i-1}(\underline{\Delta}_{i-1}), x), \qquad \ldots (4)$$

for all $x \in [0, 1]$, $i = 1, \ldots, n$.

The solution to problem (4) can be found iteratively. For simplicity, it is supposed that Ψ is strictly convex and that $U(h_{P_0} = x) = 0$ for every $x \in \mathbf{R}$. An analogous solution can be found otherwise. Notice that, under the above conditions, there exists a unique $P_n(\underline{\Delta}_n)$ satisfying conditions (ii) and (iv). Denote it by P_n. For a given $P_{i-1} = P_{i-1}(\underline{\Delta}_{i-1})$, define

$$h_{\Delta_i}^{(1)}(\omega; x) = \min(h_{\Delta^{(i)}}(\omega), c_{P_{i-1}}(x) - h_{P_{i-1}}(\omega)),$$

$$h_{\Delta_i}^{(2)}(\omega; x) = \max(h_{\Delta^{(i)}}(\omega), c_{P_{i-1}}(x)^+ - h_{P_{i-1}}(\omega)),$$

where $\Delta^{(i)} = P_n - P_{i-1}$, $i = 1, \ldots, n$. Moreover, let

$$x_i^{(1)} = \sup\{x \in [0,1] : \int_{L^-_{P_{i-1}}(x)} h_{\Delta_i}^{(1)}(\omega; x) U(d\omega) \leq \delta\},$$

$$x_i^{(2)} = \inf\{x \in [0,1] : \int_{(L_{P_{i-1}}(x))^C} h_{\Delta_i}^{(2)}(\omega; x) U(d\omega) \geq -\delta\}.$$

Then the solution is given by

$$h_{\Delta_i}(\omega) = \begin{cases} h_{\Delta_i}^{(1)}(\omega, x_i^{(1)}) & \text{if } \omega \in L^-_{P_{i-1}}(x_i^{(1)}) \\ h_{\Delta_i}^{(2)}(\omega, x_i^{(2)}) & \text{if } \omega \in (L_{P_{i-1}}(x_i^{(2)}))^C \\ 0 & \text{elsewhere.} \end{cases} \qquad \square$$

Example 3. Consider the distribution of incomes among individuals in a population. According to its definition, let $L_P(x)$ be the set of the poorest $100x\%$ part of the population. For simplicity, it will be supposed that

$$\varphi'_\Delta(P, x) = \Delta\left(L_P(x)\right)$$

holds, at least as an approximation (reasonable in a large population). two possible applications are briefly considered.

Suppose that a subclass $\mathcal{D}_0$ of $\mathcal{D}$ is given, corresponding to different policies of income redistribution (e.g. different tax systems). Focusing on $L_P(x)$ (the group of the poorest individuals), for a given $x \in (0,1)$, the Gàteaux differential gives the tendency to changes in income distribution among this group. It is worth looking for the most effective policies, that is those leading to the largest values (in absolute value) of the differentials.

Conversely, consider a given signed measure Δ and let x vary in $(0,1)$. Here the largest values (in absolute value) of the differentials give the groups $L_P(x)$ of the poorest individuals on which the policy Δ has been more effective.

EXAMPLE 4. Functional derivatives of the concentration function can be used in robust Bayesian analysis, to assess the effects of infinitesimal changes in the prior probability measure on some quantity of interest. (Different techniques have been considered in, e.g., Diaconis and Freedman, (1986) and Ruggeri and Wasserman, (1993). Some results can be found in Fortini and Ruggeri (1995c). When inferencing about the whole posterior distribution, the concentration function can be used as a measure of the discrepancy between posterior distributions, corresponding to different prior assessments (see Fortini and Ruggeri (1994, 1995a, 1995b)).

Suppose, now, that a specific prior Π_0 is suggested by some particular features of the inferential problem but, nevertheless, an analysis of sensitivity to small variations in Π_0 is deemed necessary. For any posterior distribution Π^*, corresponding to a prior Π, the concentration function of Π^* w.r.t. Π_0^*, say $\varphi(\Pi^*, x)$, describes the discrepancy between Π^* and Π_0^*. Hence, if Δ is a signed measure such that $\Delta(\Omega) = 0$ and $||\Delta|| \leq 1$, then a measure of robustness for small variations of Π in the direction of Δ is given by

$$\lim_{\lambda \to 0} \frac{\varphi((\Pi_0 + \lambda\Delta)^*, x) - \varphi(\Pi_0^*, x)}{\lambda}$$

which coincides with the differential of the functional $\Phi(\Pi) = \varphi(\Pi^*(\Pi), x)$ in Π_0 and in the direction of Δ, say $\Phi'_\Delta(\Pi_0)$.

We end this section by noticing that it would be interesting to connect the results of the paper with mean equalizing transfers, which get a probability measure to coincide with the uniform one; see Regazzini (1992).

Appendix

This Appendix includes the proofs of Theorem 2 and 3 along with some Lemmas needed for the proof of Theorem 2. Let P_0 and P be probability measures and let Δ be a finite, signed measure such that $\Delta(\Omega) = 0$. Let $x \in (0,1)$ be a fixed real number. P_0 will be assumed to be a nonatomic probability measure, throughout.

LEMMA 1. *For any $\delta \in \mathbf{R}^+$, there exists $\lambda_0 = \lambda_0(\delta)$ such that, for any λ with $0 \leq \lambda \leq \lambda_0$,*

$$c_P(x) - \delta < c_{P+\lambda\Delta}(x) < c_P(x)^+ + \delta. \qquad \ldots(5)$$

PROOF. Since it is equivalent to the first inequality in (5), it will be first proved that, for $\lambda > 0$ sufficiently small,

$$P_0(h_{P+\lambda\Delta} \leq c_P(x) - \delta) < x. \qquad \ldots(6)$$

From the left-continuity of c_P in x, for any $\eta \in (0, \delta)$ there exists $\gamma \in \mathbf{R}^+$ such that

$$c_P(y) \leq c_P(x) - \delta + \eta \text{ implies } y \leq x - \gamma. \qquad \ldots (7)$$

Taking $y = H_P(c_P(x) - \delta + \eta)$, it follows that

$$c_P(y) = \inf\{z \in \overline{\mathbf{R}} : H_P(z) \geq H_P(c_P(x) - \delta + \eta)\} \leq c_P(x) - \delta + \eta$$

and, from (7), that $y \leq x - \gamma$.

Let $\lambda_0^{(1)}$ be such that $P_0(h_\Delta \leq -\eta/\lambda_0^{(1)}) < \gamma/2$. Therefore, for any λ, $0 \leq \lambda \leq \lambda_0^{(1)}$, it holds

$$\begin{aligned} &P_0(h_{P+\lambda\Delta} \leq c_P(x) - \delta) \\ &\quad \leq P_0(h_P \leq c_P(x) - \delta - \lambda h_\Delta, -\lambda h_\Delta < \eta) + P_0\{h_\Delta \leq -\eta/\lambda\} \\ &\quad < y + \gamma/2 < x, \end{aligned}$$

which proves (6).

For the arbitrariness of $\delta > 0$, the second inequality in (5) is equivalent to

$$P_0(h_{P+\lambda\Delta} \leq c_P(x)^+ + \delta) \geq x.$$

Fix $\eta \in (0, \delta)$; there exists $\gamma \in \mathbf{R}^+$ such that,

$$y \leq x + \gamma \text{ implies } c_P(y) \leq c_P(x)^+ + \delta - \eta.$$

Let $\lambda_0^{(2)}$ be such that $P_0(\lambda_0^{(2)} h_\Delta \geq \eta) < \gamma/2$. Then

$$\begin{aligned} &P_0(h_{P+\lambda\Delta} \leq c_P(x)^+ + \delta) \geq 1 - P_0(h_P > c_P(x)^+ + \delta - \lambda h_\Delta) \\ &\quad \geq 1 - P_0(h_P > c_P(x)^+ + \delta - \lambda h_\Delta, -\lambda h_\Delta > -\eta) - P_0(\lambda h_\Delta \geq \eta) \\ &\quad \geq P_0(h_P \leq c_P(x)^+ + \delta - \eta) - \gamma/2 \geq H_P(c_P(x + \gamma)) - \gamma/2 \geq x + \gamma/2. \end{aligned}$$

The proof is completed by taking $\lambda_0 = \min\{\lambda_0^{(1)}, \lambda_0^{(2)}\}$. □

LEMMA 2. For any sequence $\lambda_n \downarrow 0$,

a) $\bigcap_{N=1}^\infty \bigcup_{n=N}^\infty (L_P^-(x) \cap (L_{P+\lambda_n\Delta}^-(x))^C) = \emptyset$;

b) $\lim_{n\to\infty} P_0((L_{P+\lambda_n\Delta}^-(x))^C \cap L_P^-(x)) = 0$;

c) if $c_P(x) = c_P(x)^+$, $\bigcap_{N=1}^\infty \bigcup_{n=N}^\infty (L_{P+\lambda_n\Delta}(x) \cap (L_P(x))^C) = \emptyset$;

d)$\lim_{n\to\infty} P_0(L_{P+\lambda_n\Delta}^-(x) \cap (L_P(x))^C) = 0$.

PROOF. *a)* Fix $\omega \in L_P^-(x)$. Then $h_P(\omega) < c_P(x) - \varepsilon$ for some $\varepsilon > 0$. Let $N_0 \in \mathbf{N}$ be such that, for any $n \geq N_0$, $\lambda_n |h_\Delta(\omega)| < \varepsilon/4$. Let $N_1 \in \mathbf{N}$ be such that, for any $n \geq N_1$, $c_{P+\lambda_n h_\Delta}(x) > c_P(x) - \varepsilon/4$, according to Lemma 1. Then, for any $n \geq N = \max(N_0, N_1)$,

$$h_{P+\lambda_n\Delta}(\omega) \leq h_P(\omega) + \lambda_n|h_\Delta(\omega)| \leq c_P(x) - 3\varepsilon/4 \leq c_{P+\lambda_n h_\Delta}(x) - \varepsilon/2.$$

Hence, if $\omega \in L_P^-(x)$, then $\omega \in L^-_{P+\lambda_n\Delta}(x)$ for any $n \geq N$. Therefore $L_P^-(x) \subset \bigcup_{N=1}^\infty \bigcap_{n=N}^\infty L^-_{P+\lambda_n\Delta}(x)$, which implies *a)*.

b) follows directly from *a)* and from the continuity of probability measures.

c) Consider $\omega \in \bigcap_{N=1}^\infty \bigcup_{n=N}^\infty L_{P+\lambda_n\Delta}(x)$. Fix a real number $\varepsilon > 0$. Let $N_0 \in \mathbf{N}$ be such that, for any $n \geq N_0$, $c_{P+\lambda_n\Delta}(x) < c_P(x)^+ + \varepsilon/2$, according to Lemma 1. Let $N_1 \in \mathbf{N}$ be such that, for any $n \geq N_1$, $\lambda_n|h_\Delta(\omega)| < \varepsilon/2$. Since $\omega \in \bigcap_{N=1}^\infty \bigcup_{n=N}^\infty L_{P+\lambda_n\Delta}(x)$, there exists $n \geq N = \max(N_0, N_1)$ such that

$$h_P(\omega) = h_{P+\lambda_n\Delta}(\omega) - \lambda_n h_\Delta(\omega) \leq c_{P+\lambda_n\Delta}(x) + \lambda_n|h_\Delta(\omega)| < c_P(x)^+ + \varepsilon = c_P(x) + \varepsilon.$$

Since this holds for every $\varepsilon > 0$, $h_P(\omega) \leq c_P(x)$ and, therefore, $\omega \in L_P(x)$. It follows that $\bigcap_{N=1}^\infty \bigcup_{n=N}^\infty L_{P+\lambda_n\Delta}(x) \subset L_P(x)$.

d) follows from *c)* for the x's such that $c_P(x) = c_P(x)^+$. We will prove it for $c_P(x)^+ - c_P(x) > 0$.

Suppose that there exists a sequence $\lambda_n \downarrow 0$ such that $P_0((L_P(x))^C \cap L^-_{P+\lambda_n\Delta}(x)) > 2\delta$ for some $\delta > 0$ and for every $n \in \mathbf{N}$. Choose $k > 0$ such that $P_0(|h_\Delta| > k) < \delta/2$ and let $H = (|h_\Delta| < k)$. Then $P_0((L_P(x))^C \cap L^-_{P+\lambda_n\Delta}(x)) \cap H) > \delta$ for every $n \in \mathbf{N}$. Since $P_0(c_P(x) < h_P < c_P(x)^+) = 0$, then $P_0((h_P > c_P(x)^+) \cap L^-_{P+\lambda_n\Delta}(x) \cap H) > \delta$ for every $n \in \mathbf{N}$. Take N_0 such that, for every $n \geq N_0$, $\lambda_n k < \varepsilon$ with $0 < \varepsilon < (c_P(x)^+ - c_P(x))/2$ and take $\omega \in (h_P > c_P(x)^+) \cap L^-_{P+\lambda_n\Delta}(x) \cap H$. Then, for every $n \geq N_0$, $c_{P+\lambda_n\Delta}(x) > h_{P+\lambda_n\Delta}(\omega) \geq h_P(\omega) - \lambda_n|h_\Delta(\omega)| > c_P(x)^+ - \varepsilon$.

Let $K = (|h_\Delta| < k) \cap L_P(x)$. Then $P_0(K) \geq P_0(L_P(x)) - \delta/2 \geq x - \delta/2$. Fix $N_1 \in \mathbf{N}$ such that, for any $n > N_1$, $\lambda_n < \varepsilon/k$. Then, if $n \geq \max(N_0, N_1)$ and $\omega \in K$,

$$h_{P+\lambda_n\Delta}(\omega) \leq h_P(\omega) + \lambda_n|h_\Delta(\omega)| \leq c_P(x) + \varepsilon < c_P(x)^+ - \varepsilon < c_{P+\lambda_n\Delta}(x).$$

It follows that $K \subset L^-_{P+\lambda_n\Delta}(x)$, for $n > \max(N_0, N_1)$. Moreover

$$x \geq P_0(L^-_{P+\lambda_n\Delta}(x)) \geq P_0(K) + P_0(L^-_{P+\lambda_n\Delta}(x) \cap (L_P(x))^C) \geq x + \delta/2$$

which is a contradiction. □

LEMMA 3. a) *There exist* $A_\lambda \subset \{\omega : h_{P+\lambda\Delta}(\omega) = c_{P+\lambda\Delta}(x)\}$, $A_x \subset \{\omega : h_x(\omega) = c_x(x)\}$ *and* $A \subset \{\omega : h_P(\omega) = c_P(x)\}$ *such that* $P_0(A_\lambda) =$

$x - H_{P+\lambda\Delta}(c_{P+\lambda\Delta}(x)^-)$, $P_0(A_x) = x - H_x(c_x(x)^-)$, $P_0(A) = x - H_P(c_P(x)^-)$ and $L_P^-(x) \cup A = L_x^-(x) \cup A_x$, a.s. $- P_0$.

b) For every sequence $\lambda_n \downarrow 0$, $\lim_{n\to\infty} P_0(A_{\lambda_n} \cap (L_P(x))^C) = 0$.

c) If, for every $\lambda > 0$, $B_\lambda = \{\omega : h_P(\omega) = c_P(x), h_\Delta(\omega) < (c_{P+\lambda\Delta}(x) - c_P(x))/\lambda\}$, $B_x = \{\omega : h_P(\omega) = c_P(x), h_\Delta(\omega) < c_x(x)\}$, then, for every sequence $\lambda_n \downarrow 0$,

$$\lim_{n\to\infty} |P_0(B_{\lambda_n} \cup A_{\lambda_n}) - P_0(B_x \cup A_x)| = 0.$$

PROOF. *a)* Since P_0 is nonatomic, there exist A_λ and A such that $A_\lambda \subset (h_{P+\lambda\Delta} = c_{P+\lambda\Delta}(x))$, $A_x \subset (h_x = c_x(x))$, $P_0(A_\lambda) = x - H_{P+\lambda\Delta}(c_{P+\lambda\Delta}(x)^-)$ and $P_0(A_x) = x - H_x(c_x(x)^-)$. If $A = (L_x^-(x) \setminus L_P^-(x)) \cup A_x$, then $L_P^-(x) \cup A = L_x^-(x) \cup A_x$ and $P_0(A) = x - H_P(c_P(x)^-)$.

b) If $c_P(x)^+ = c_P(x)$, *b)* follows from Lemma 2 *c)*. We will prove *b)* when $c_P(x)^+ - c_P(x) > 2\delta$ for some $\delta > 0$. Suppose that there exist $\varepsilon > 0$ and a sequence λ_n such that, for every n, $P_0(A_{\lambda_n} \cap (L_P(x))^C) > \varepsilon$. Let $K = (|h_\Delta| < k) \cap L_P(x)$, where k is such that $P_0(K) > x - \varepsilon/2$. As in the proof of Lemma 2 d), it can be shown that there exists N_0 such that, for any $n \geq N_0$, $c_P(x)^+ < c_{P+\lambda_n\Delta}(x) + \delta$; moreover let N_1 be such that, for any $n \geq N_1$, $\lambda_n < \delta/(2k)$. For every $n \geq N = \max(N_0, N_1)$, if $\omega \in K$,

$$\begin{aligned} h_{P+\lambda_n\Delta}(\omega) &\leq c_P(x) + \lambda_n k \leq c_P(x)^+ - 2\delta + \lambda_n k \\ &\leq c_{P+\lambda_n\Delta}(x) - \delta + \lambda_n k \leq c_{P+\lambda_n\Delta}(x) - \delta/2; \end{aligned}$$

it follows that $K \subset L_{P+\lambda_n\Delta}^-(x)$, for any $n \geq N$ but then

$$x = P_0(L_{P+\lambda_n\Delta}^-(x) \cup A_{\lambda_n}) \geq P_0(K) + P_0(A_{\lambda_n} \cap (L_P(x))^C) \geq x - \varepsilon/2$$

which is a contradiction.

c) Since $P_0(L_{P+\lambda\Delta}^-(x) \cup A_\lambda) = P_0(L_x^-(x) \cup A_x)$ and $L_x^-(x) = L_P^-(x) \cup B_x$,

$$\begin{aligned} 0 &= P_0(L_{P+\lambda\Delta}^-(x)) + P_0(A_\lambda) - P_0(L_P^-(x)) - P_0(B_x) - P_0(A_x) \\ &= P_0(L_{P+\lambda\Delta}^-(x) \cap (L_P^-(x))^C) + P_0(A_\lambda) - P_0(L_P^-(x) \cap (L_{P+\lambda\Delta}^-(x))^C) \\ &\quad - P_0(B_x) - P_0(A_x) \\ &= P_0(B_\lambda) + P_0(L_{P+\lambda\Delta}^-(x) \cap (L_P(x))^C) + P_0(A_\lambda) - P_0(L_P^-(x) \cap (L_{P+\lambda\Delta}^-(x))^C) \\ &\quad - P_0(B_x) - P_0(A_x). \end{aligned}$$

Therefore, from Lemma 2, for any $\lambda_n \downarrow 0$,

$$\begin{aligned} |P_0(B_{\lambda_n} \cup A_{\lambda_n}) - P_0(B_x \cup A_x)| \leq\ & P_0(L_{P+\lambda_n\Delta}^-(x) \cap (L_P(x))^C) \\ & + P_0(L_P^-(x) \cap (L_{P+\lambda_n\Delta}^-(x))^C) \to 0 \end{aligned}$$

as $n \to \infty$ □

Lemma 4. *Let $\lambda_n \downarrow 0$ as $n \to \infty$. Then, for every $\delta > 0$ there exists N_0 such that, for every $n \geq N_0$,*

$$c_x(x) - \delta < \frac{c_{P+\lambda_n \Delta}(x) - c_P(x)}{\lambda_n} < c_x(x)^+ + \delta.$$

Proof. Suppose that there exist $\delta > 0$ and a sequence $\lambda_n \downarrow 0$ such that, for every n,

$$\frac{c_{P+\lambda_n \Delta}(x) - c_P(x)}{\lambda_n} > c_x(x)^+ + \delta.$$

Since $A_x \subseteq (h_P = c_P(x))$, then

$$P_0(A_{\lambda_n} \cup B_{\lambda_n}) \geq P_0(B_{\lambda_n}) = P_0(h_P = c_P(x), h_\Delta < \frac{c_{P+\lambda_n \Delta}(x) - c_P(x)}{\lambda_n})$$

and

$$P_0(A_x \cup B_x) \leq P_0(h_P = c_P(x), h_\Delta \leq c_x(x)).$$

Hence

$$\begin{aligned} &P_0(A_{\lambda_n} \cup B_{\lambda_n}) - P_0(A_x \cup B_x) \\ &\quad\geq \; P_0(h_P = c_P(x), c_x(x) < h_\Delta < \frac{c_{P+\lambda_n \Delta}(x) - c_P(x)}{\lambda_n}) \\ &\quad\geq \; P_0(h_P = c_P(x), c_x(x) < h_\Delta < c_x(x)^+ + \delta) \geq P_0(c_x(x) < h_x < c_x(x)^+ + \delta\}) \not\to 0 \end{aligned}$$

as $n \to \infty$, which contradicts Lemma 3.

Suppose now that there exist $\delta > 0$ and a sequence $\lambda_n \downarrow 0$ such that, for every n,

$$\frac{c_{P+\lambda_n \Delta}(x) - c_P(x)}{\lambda_n} < c_x(x) - \delta.$$

Since $B_{\lambda_n} \subset (h_P = c_P(x)) \cap B_x$ and $A_{\lambda_n} \cap (h_P = c_P(x)) \subseteq (h_P = c_P(x), h_\Delta = [c_{P+\lambda_n \Delta}(x) - c_P(x)]/\lambda_n)$, then

$$
\begin{aligned}
&P_0(A_x \cup B_x) - P_0(A_{\lambda_n} \cup B_{\lambda_n})\\
&\quad = P_0(A_x) + P_0(B_x) - P_0(A_{\lambda_n}) - P_0(B_{\lambda_n})\\
&\quad = P_0(A_x) + P_0(h_P = c_P(x), [c_{P+\lambda_n\Delta}(x) - c_P(x)]/\lambda_n \le h_\Delta < c_x(x)) - P_0(A_{\lambda_n})\\
&\quad \ge P_0(A_x) + P_0(h_P = c_P(x), [c_{P+\lambda_n\Delta}(x) - c_P(x)]/\lambda_n < h_\Delta < c_x(x))\\
&\quad + P_0(h_P = c_P(x), h_\Delta = [c_{P+\lambda_n\Delta}(x) - c_P(x)]/\lambda_n)\\
&\quad - P_0(A_{\lambda_n} \cap L_P^-(x)) - P_0(A_{\lambda_n} \cap (L_P(x))^C) - P_0(A_{\lambda_n} \cap (h_P = c_P(x)))\\
&\quad \ge P_0(A_x) + P_0(h_P = c_P(x), c_x(x) - \delta < h_\Delta < c_x(x)\}\\
&\quad - P_0(A_{\lambda_n} \cap (L_P(x))^C) - P_0((L_{P+\lambda_n\Delta}^-(x))^C \cap L_P^-(x)).
\end{aligned}
$$

Hence,

$$
\begin{aligned}
P_0(A_x \cup B_x) - P_0(A_{\lambda_n} \cup B_{\lambda_n})&\\
\ge\ & P_0(A_x) + P_0(h_P = c_P(x), c_x(x) - \delta < h_\Delta < c_x(x))\\
& - P_0((L_{P+\lambda_n\Delta}^-(x))^C \cap L_P^-(x)) - P_0(A_{\lambda_n} \cap (L_P(x))^C).
\end{aligned}
$$

From Lemma 2 b) and Lemma 3 b), for every $\varepsilon > 0$, there exists $N(\varepsilon)$ such that, for any $n \ge N(\varepsilon)$,

$$P_0((L_{P+\lambda_n\Delta}^-(x))^C \cap L_P(x)^-) < \varepsilon/4 \text{ and } (P_0(A_{\lambda_n} \cap (L_P(x))^C) < \varepsilon/4.$$

If $x \ne H_x(c_x(x)^-)$, then there exists $\varepsilon > 0$ such that $P_0(A_x) > \varepsilon$ and, therefore, for any $n \ge N(\varepsilon)$,

$$P_0(A_x \cup B_x) - P_0(A_{\lambda_n} \cup B_{\lambda_n}) \ge P_0(A_x) - \varepsilon/2 \ge \varepsilon/2,$$

which contradicts Lemma 3.
On the other hand, if $x = H_x(c_x(x)^-)$, there exists $\varepsilon > 0$ such that

$$P_0(c_x(x) - \delta < h_x < c_x(x)) > \varepsilon.$$

Then, for every $n \ge N(\varepsilon)$,

$$P_0(A_x \cup B_x) - P_0(A_{\lambda_n} \cup B_{\lambda_n}) \ge P_0(c_x(x) - \delta < h_x < c_x(x)) - \varepsilon/2 > \varepsilon/2,$$

which contradicts Lemma 3. □

LEMMA 5. *Consider a sequence* $\lambda_n \downarrow 0$ *as* $n \to \infty$ *and suppose that* $x \ne H_P(c_P(x)^-)$. *The following results hold, for every* $\varepsilon > 0$.

a) *There exists* N_1 *such that, for every* $n \geq N_1$, $\Delta(A_{\lambda_n}) \geq c_x(x)P_0\ (A_{\lambda_n}) -\varepsilon/2$.

b) *There exists* N_2 *such that, for every* $n \geq N_2$, $\Delta(L^-_{P+\lambda_n\Delta}(x) \cap\ (L_P(x))^C) \geq -\varepsilon$.

c) *There exists* N_3 *such that, for every* $n \geq N_3$, $\Delta(B_{\lambda_n} \cup A_{\lambda_n}) - \Delta(B_x \cup A_x) > -\varepsilon$.

PROOF. Since $x \neq H_P(c_P(x)^-)$, $c_x(x) \neq -\infty$. According to Lemma 4 there exists N_0 such that, for every $n \geq N_0$,

$$\frac{c_{P+\lambda_n\Delta}(x) - c_P(x)}{\lambda_n} > c_x(x) - \varepsilon/4.$$

a) If $c_P(x) = c_P(x)^+$, then by Lemma 2 c), there exists $N_1 > N_0$ such that, for $n \geq N_1$,

$$\max\{|\Delta|(A_{\lambda_n} \cap (L_P(x))^C), c_x(x)P_0(A_{\lambda_n} \cap (L_P(x))^C)\} < \varepsilon/8.$$

Hence,

$$\begin{aligned} \Delta(A_{\lambda_n}) &= \Delta(A_{\lambda_n} \cap L_P(x)) + \Delta(A_{\lambda_n} \cap (L_P(x))^C) \\ &\geq [c_x(x) - \varepsilon/4]P_0(A_{\lambda_n} \cap L_P(x)) - \varepsilon/8 \\ &\geq c_x(x)P_0(A_{\lambda_n}) - 3/8\varepsilon - c_x(x)[P_0(A_{\lambda_n}) - P_0(A_{\lambda_n} \cap L_P(x))] \\ &\geq c_x(x)P_0(A_{\lambda_n}) - \varepsilon/2. \end{aligned}$$

On the other hand, if $c_P(x) < c_P(x)^+ - \eta$, then, by Lemma 4, there exists $M_1 > N_0$ such that, for $n \geq M_1$,

$$c_{P+\lambda_n\Delta}(x) < c_P(x) + \lambda_n c_x(x)^+ + \eta/4 < c_P(x)^+ - \eta/2.$$

Let k be such that, if $H = \{h_\Delta < -k\}$, then

$$\max\{|\Delta|(H), c_x(x)P_0(H)\} < \varepsilon/8.$$

Choose $N_1 > M_1$ such that, for every $n \geq N_1$, $\lambda_n k < \eta/4$. Then, if $\omega \in A_{\lambda_n} \cap H^C$,

$$h_P(\omega) = c_{P+\lambda_n\Delta}(x) - \lambda_n h_\Delta < c_{P+\lambda_n\Delta}(x) + \lambda_n k < c_{P+\lambda_n\Delta}(x) + \eta/4 < c_P(x)^+ - \eta/4.$$

Therefore, since $P_0(c_P(x) < h_P < c_P(x)^+ - \eta/4) = 0$, it follows that

$$\begin{aligned}\Delta(A_{\lambda_n}) &= \Delta(A_{\lambda_n} \cap H^C) + \Delta(A_{\lambda_n} \cap H)\\ &\geq (c_x(x) - \varepsilon/4)P_0(A_{\lambda_n} \cap H^C) - \varepsilon/8\\ &\geq c_x(x)P_0(A_{\lambda_n}) - 3/8\varepsilon - c_x(x)[P_0(A_{\lambda_n}) - P_0(A_{\lambda_n} \cap H^C)\\ &\geq c_x(x)P_0(A_{\lambda_n}) - \varepsilon/2.\end{aligned}$$

b) Let k be such that, if $H = \{h_\Delta < -k\}$, then $|\Delta|(H) < \varepsilon/2$. According to Lemma 2 d), let N_2 be such that, for every $n \geq N_2$, $kP_0(L^-_{P+\lambda_n\Delta}(x) \cap (L_P(x))^C) < \varepsilon/2$. Then

$$\begin{aligned}&\Delta(L^-_{P+\lambda_n\Delta}(x) \cap (L_P(x))^C)\\ &= \Delta(L^-_{P+\lambda_n\Delta}(x) \cap (L_P(x))^C \cap H^C) + \Delta(L^-_{P+\lambda_n\Delta}(x) \cap (L_P(x))^C \cap H)\\ &\geq -kP_0(L^-_{P+\lambda_n\Delta}(x) \cap (L_P(x))^C) - \varepsilon/2\\ &\geq -\varepsilon.\end{aligned}$$

c) From Lemma 3 and Lemma 2 c), there exists $N_3 \geq N_0$ such that, for every $n \geq N_3$,

$$c_x(x)(P_0(B_{\lambda_n} \cup A_{\lambda_n}) - P_0(B_x \cup A_x)) > -\varepsilon/2.$$

and

$$\max\{|\Delta|(A_{\lambda_n} \cap (L_P(x))^C), c_x(x)P_0(A_{\lambda_n} \cap (L_P(x))^C)\} < \varepsilon/8.$$

Then, it follows that

$$\begin{aligned}&\Delta(B_{\lambda_n} \cup A_{\lambda_n}) - \Delta(B_x \cup A_x)\\ \geq\ & \Delta(h_P = c_P(x), h_\Delta < [c_{P+\lambda_n\Delta}(x) - c_P(x)]/\lambda_n, h_\Delta \geq c_x(x))\\ & -\Delta(h_P = c_P(x), h_\Delta < c_x(x), h_\Delta \geq [c_{P+\lambda_n\Delta}(x) - c_P(x)]/\lambda_n)\\ & +\Delta(A_{\lambda_n} \cap (L_P(x))^C) + \Delta(A_{\lambda_n} \cap L_P(x)) - \Delta(A_x)\\ \geq\ & c_x(x)P_0(B_{\lambda_n}) - c_x(x)P_0(B_x) + c_x(x)P_0(A_{\lambda_n}) - c_x(x)P_0(A_x) - \varepsilon/2\\ \geq\ & -\varepsilon.\end{aligned}$$

□

Proof of Theorem 2. Suppose that P_0 is nonatomic. Consider a sequence $\lambda_n \downarrow 0$ as $n \to \infty$. Then, from Lemma 3 a) in the Appendix, there exist

A_{λ_n}, A and A_x such that

$$\varphi_{P+\lambda_n\Delta}(x) - \varphi_P(x) - \lambda_n\varphi'_\Delta(x)$$

$$= (P+\lambda_n\Delta)(L^-_{P+\lambda_n\Delta}(x)\cup A_{\lambda_n}) - P(L^-_P(x)\cup A) - \lambda_n\Delta(L^-_x(x)\cup A_x)$$

$$= (P+\lambda_n\Delta)(L^-_{P+\lambda_n\Delta}(x)\cup A_{\lambda_n}) - (P+\lambda_n\Delta)(L^-_P(x)\cup A) \le 0,$$

since $(P+\lambda_n\Delta)(C) \ge (P+\lambda_n\Delta)(L^-_{P+\lambda_n\Delta}(x)\cup A_{\lambda_n})$, for any C such that $P_0(C)\ge x$. On the other hand,

$$\varphi_{P+\lambda_n\Delta}(x) - \varphi_P(x) - \lambda_n\varphi'_\Delta(x)$$

$$= P(L^-_{P+\lambda_n\Delta}(x)\cup A_{\lambda_n}) - P(L^-_P(x)\cup A) + \lambda_n\Delta(L^-_{P+\lambda_n\Delta}(x)\cup A_{\lambda_n})$$

$$-\lambda_n\Delta(L^-_x(x)\cup A_x)$$

$$\ge \lambda_n\Delta(L^-_{P+\lambda_n\Delta}(x)\cup A_{\lambda_n}) - \lambda_n\Delta(L^-_x(x)\cup A_x),$$

since $P(C)\ge P(L^-_P(x)\cup A)$, for any C such that $P_0(C)\ge x$. Moreover, if B_{λ_n} and B_x are defined as in Lemma 3 c),

$$\Delta(L^-_{P+\lambda_n\Delta}(x)\cup A_{\lambda_n}) - \Delta(L^-_x(x)\cup A_x)$$

$$\ge \Delta(L^-_{P+\lambda_n\Delta}(x)) - \Delta(L^-_P(x)) - \Delta(B_x) + \Delta(A_{\lambda_n}) - \Delta(A_x)$$

$$\ge \Delta(B_{\lambda_n}) + \Delta(L^-_{P+\lambda_n\Delta}(x)\cap(L_P(x))^C) - \Delta(L^-_P(x)\cap(L^-_{P+\lambda_n\Delta}(x))^C)$$

$$-\Delta(B_x) + \Delta(A_{\lambda_n}) - \Delta(A_x)$$

$$\ge \Delta(B_{\lambda_n}\cup A_{\lambda_n}) - \Delta(B_x\cup A_x) + \Delta(L^-_{P+\lambda_n\Delta}(x)\cap(L_P(x))^C) -$$

$$-\Delta(L^-_P(x)\cap(L^-_{P+\lambda_n\Delta}(x))^C) \to 0,$$

as $n\to\infty$, from Lemma 2 and Lemma 5 in the Appendix.

Suppose now that P_0 is any probability measure on $(\Omega,\mathcal{A})$. Consider the probability space $([0,1],\mathcal{B}([0,1]),\hat{P}_0)$, where $\mathcal{B}([0,1])$ and $\hat{P}_0$ denote the σ-algebra of the Borel sets and the Lebesgue measure on $[0,1]$, respectively. Let $c_P(x)$ be the quantile function of the Radon-Nikodym derivative of P w.r.t. P_0. Consider the probability measure $\hat{P}$ on $([0,1],\mathcal{B}([0,1]))$ whose density, w.r.t. $\hat{P}_0$, is $\hat{h}(x) = c_P(x)$. Consider a random variable $\hat{h}_\Delta$ on $([0,1],\mathcal{B}([0,1]))$ such that the joint distribution of $(\hat{h},\hat{h}_\Delta)$ under $\hat{P}_0$ coincides with the joint distribution of (h_P,h_Δ) under P_0. Denote by $\hat{\Delta}$ the signed measure on $([0,1],\mathcal{B}([0,1]))$ which has density $\hat{h}_\Delta$ w.r.t. $\hat{P}_0$. Then $\hat{\Delta}([0,1]) = 0$; moreover, for any $\lambda > 0$, the distribution of $\hat{h}+\lambda\hat{h}_\Delta$ under $\hat{P}_0$ coincides with the distribution of $h_P+\lambda h_\Delta$ under P_0. It follows that the concentration function of $\hat{P}+\lambda\hat{\Delta}$ w.r.t. $\hat{P}_0$, $\varphi_{\hat{P}+\lambda\hat{\Delta}}$, coincides with the concentration function $\varphi_{P+\lambda\Delta}$ of $P+\lambda\Delta$ w.r.t. P_0, for every $\lambda>0$. Since $\hat{P}_0$ is nonatomic,

$$\begin{aligned}\varphi'_\Delta(P,x)&=\varphi'_{\hat{\Delta}}(\hat{P},x)\\&=\begin{cases}\hat{\Delta}(\hat{L}^-(x)) & \text{if } x=H_{\hat{P}}(c_{\hat{P}}(x)^-)\\ \hat{\Delta}(\hat{L}_x^-(x))+c_{\hat{P},x}(x)\{x-H_{\hat{P},x}(c_{\hat{P},x}(x)^-)\} & \text{if } x\neq H_{\hat{P}}(c_{\hat{P}}(x)^-),\end{cases}\end{aligned}$$

where $\hat{L}^-(x)=\{y\in[0,1]:\hat{h}(y)<c_{\hat{P}}(x)\}$ and $c_{\hat{P}}(x)$ is the quantile function of $\hat{h}$. Moreover, since the distribution of $(\hat{h},\hat{h_\Delta})$ under $\hat{P}_0$ coincides with the joint distribution of (h_P,h_Δ) under P_0, $x=H_{\hat{P}}(c_{\hat{P}}(x)^-)$ holds if and only if $x=H_P(c_P(x)^-)$, $\hat{\Delta}(\hat{L}^-(x))=\Delta(L_P^-(x))$, $\hat{\Delta}(\hat{L}_x^-(x))=\Delta(L_x^-(x))$, $H_{\hat{P},x}=H_{P,x}$ and $c_{\hat{P},x}=c_{P,x}$. The Theorem follows immediately. □

Proof of Theorem 3. By the definition of Gâteaux derivative,

$$\rho'_\Delta(P,g)=\lim_{\lambda\to0^+}\frac{\rho(P+\lambda\Delta,g)-\rho(P,g)}{\lambda}=l_1+l_2,$$

where

$$\begin{aligned}l_1&=\lim_{\lambda\to0^+}\frac{1}{\lambda}\left\{\int_{\mathbf{R}}g(t)dH_{P+\lambda\Delta}(t)-\int_{\mathbf{R}}g(t)dH_P(t)\right\},\\ l_2&=\lim_{\lambda\to0^+}\frac{1}{\lambda}\left\{\left[(P+\lambda\Delta)_s^+(\Omega)-P_s(\Omega)\right]g^+-(P+\lambda\Delta)_s^-(\Omega)g^-\right\},\end{aligned}$$

if such limits exist. Observe that

$$l_1=\lim_{\lambda\to0^+}\int_{(h_\Delta<0)}I_\lambda(\omega)h_\Delta(\omega)P_0(d\omega)+\lim_{\lambda\to0^+}\int_{(h_\Delta>0)}I_\lambda(\omega)h_\Delta(\omega)P_0(d\omega),$$

where

$$I_\lambda(\omega)=\frac{g(h_P(\omega)+\lambda h_\Delta(\omega))-g(h_P(\omega))}{\lambda h_\Delta(\omega)}.$$

Since g is a convex function,

$$|I_\lambda(\omega)|\le\max\left(\left|\lim_{t\to\infty}g(t)/t\right|,\left|\lim_{t\to-\infty}g(t)/t\right|\right)<\infty.$$

Hence, from the dominated convergence theorem, it follows that

$$l_1=\int_\Omega g'_+(h_P(\omega))h_{\Delta^+}(\omega)P_0(d\omega)-\int_\Omega g'_-(h_P(\omega))h_{\Delta^-}(\omega)P_0(d\omega).$$

Furthermore, if $P_s(N) > 0$, there exists λ_0 such that $(P+\lambda\Delta)(N) > 0$ for any $\lambda < \lambda_0$. Hence, it follows that

$$\begin{aligned} l_2 &= \lim_{\lambda\to 0^+} \frac{1}{\lambda}((P+\lambda\Delta)_s(N) - P_s(N))g^+ \\ &\quad + \lim_{\lambda\to 0^+} \frac{1}{\lambda}((P+\lambda\Delta)_s^+(N^C)g^+ - (P+\lambda\Delta)_s^-(N)g^- - (P+\lambda\Delta)_s^-(N^C)g^-) \\ &= \Delta_s(N)g^+ + \Delta_s^+(N^C)g^+ - \Delta_s^-(N^C)g^-. \quad \square \end{aligned}$$

References

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

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

DIACONIS, P. and FREEDMAN, D. (1986). On the consistency of Bayes estimates. *Ann. Statist.*, 14, 1-67.

FORTINI, S. and RUGGERI, F. (1993). Concentration function and coefficients of divergence for signed measures. *J. Ita. Stat. Soc.*, 2, 17-34.

−−−− (1994). Concentration function and Bayesian robustness. *J. Stat. Plann. Inference*, 40, 205-220.

−−−− (1995a). On defining neighbourhoods of measures through the concentration function. *Sankhyā* A, 56, 444-457.

−−−− (1995b). Concentration function and sensitivity to the prior. To appear in *J. Ita. Stat. Soc.*

−−−− (1995c). Infinitesimal Properties of the Concentration Function and Bayesian Robustness. *Quaderno IAMI 95.8*, Milano: CNR-IAMI.

LUSTERNIK, L. and SOBOLEV, V. (1989). *Précis d'Analyse Fontionnelle.* Moscow: Mir.

HUBER, P.J. (1981). *Robust Statistics.* New York: Wiley.

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

REGAZZINI, E. (1992). Concentration comparisons between probability measures. *Sankhyā B*, 54, 129-149.

RUGGERI, F. and WASSERMAN, L.A. (1993). Infinitesimal sensitivity of posterior distributions. *Canad. J. Statist.* 21, 195-203.

ISTITUTO DI METODI QUANTITATIVI
UNIVERSITÀ L. BOCCONI
VIA SARFATTI 25
I-20136 MILANO, ITALY

ISTITUTO PER LE APPLICAZIONI
DELLA MATEMATICA E DELL'INFORMATICA
CONSIGLIO NAZIONALE DELLE RICERCHE
VIA AMPÉRE 56
I-20131 MILANO, ITALY
fabrizio@iami.mi.cnr.it