

Quantiles and Convexity

Michel Valadier

*Université de Montpellier, IMAG, Case courrier 051,
place Eugène Bataillon, 34095 Montpellier cedex 5, France*

Dedicated to the memory of Jean Jacques Moreau.

Received: November 11, 2014

Accepted: December 14, 2014

A real random variable admits median(s) and quantiles. These values minimize convex functions on $\mathbb{R}$. This is known. We discuss these results and their relationship with some notions about functions of bounded variation developed by J. J. Moreau in his mathematical works in view of the mechanical phenomena which are shocks and friction. Specially *filled-in graphs* and Stieltjes measures of products of BV functions.

Keywords: Median, quantile, convexity, subdifferential, monotone operator, minimizer, BV function, Jean Jacques Moreau

1991 Mathematics Subject Classification: 62E99, 52A99

1. Remembrance

I began research in Mathematics under the supervision of Robert Pallu de La Barrière (and with the companionship of Charles Castaing). In the middle of the sixties Pallu de La Barrière foreboded the future importance of convexity. Around 1967 I received from Castaing a copy of *Fonctionnelles convexes* (Collège de France) by Jean Jacques Moreau [24]. An impressive basic fundamental text! Some time later (in October 1970) we all three (Moreau, Castaing and me) were together in Montpellier for more than a decade, at the big time of the *Séminaire d'Analyse Convexe*.

J. J. Moreau was not only mathematician (and a very good one): he always claimed he was a Mechanician (in the academic sense) and he developed Mathematics in view of Mechanics. But in Mathematics he had a taste for generality and avoided useless hypotheses: for examples in [24] the framework was (not necessarily locally convex) infinite dimensional topological vector spaces, and in [26] he proved basic results about BV functions with values in a metric space whereas he needed only results in a Hilbert space (BV abbreviates *bounded variation*)!

This paper provides “Convex Analysis” proofs of an elementary result of Statistics. We discuss the hypotheses and involved structures and invoke some favorite

notions of J. J. Moreau, specially concerning BV functions. We follow a logical way which throws light on the naturalness of the result. I began this paper (for a first version see [36]) being unaware of Koltchinskii [18]. This 1997 work is more devoted to the multivariate case, as well as the 1996 paper by Chaudhuri [5]. See some other comments in Section 4.

2. Definitions and notations

Let $\mathbf{P}$ be a probability law on $\mathbb{R}$. If necessary X will denote a random variable obeying the law $\mathbf{P}$. The law $\mathbf{P}$ is of order p ($p \in \{1, 2\}$) if X is of order p , which writes $X \in L^p$. If no integrability condition on X is assumed, specialists write $X \in L^0$ (order 0). When $\mathbf{P}$ is of order 1 the mean $\mathbb{E}(X)$ does exist. When $\mathbf{P}$ is of order 0 there exists a non-empty compact interval of *medians* (the definition of a median is included in that of a quantile below). Let F be the right-continuous *distribution function*:

$$F(x) := \mathbf{P}(]-\infty, x]) = \mathbf{P}(X \leq x).$$

For $\tau \in]0, 1[$ (usually in Statistics 0.95 or 0.99 etc.) a real number $\bar{q}$ is a τ -quantile if

$$\mathbf{P}(X \leq \bar{q}) \geq \tau \quad \text{and} \quad \mathbf{P}(X \geq \bar{q}) \geq 1 - \tau$$

or maybe more plainly

$$\mathbf{P}(X < \bar{q}) \leq \tau \leq \mathbf{P}(X \leq \bar{q})$$

(note that $x \mapsto \mathbf{P}(X < x)$ is the left-continuous version F^- of F). When $\tau = 1/2$ one recovers median(s). A geometrical definition of τ -quantiles is the following: their set is the projection on $\mathbb{R}$ of the intersection of $\mathbb{R} \times \{\tau\}$ with the *filled-in graph*¹ G of F . That is, the set of τ -quantiles is

$$\{\bar{q} \in \mathbb{R}; (\bar{q}, \tau) \in G\}.$$

This geometrical definition operates when $\tau = 0$ or 1. For example the set of 1-quantiles is the closed, possibly empty, interval $\{q \in \mathbb{R}; F(q) = 1\}$ and the set of 0-quantiles is the closed, possibly empty, interval $\{q \in \mathbb{R}; F^-(q) = 0\}$.

A real number candidate to be the mean or a median will be denoted by m . A candidate to be a quantile will be denoted by q .

¹ This graph is obtained from the graph of F by adding vertical segments when there are gaps, thus obtaining an arcwise connected curve:

$$G = \{(x, y); x \in \mathbb{R} \text{ and } y \in [F^-(x), F(x)]\}.$$

This comes from Monteiro Marques and Moreau's works on the sweeping process: [21, p. 147], [22, p. 15], some ideas coming back to [25]. In Section 4 the filled-in graph will be seen as a maximal monotone operator.

3. Medians and quantiles as minimizers

Well known² is the result: *If $\mathbf{P}$ is of order 1, $\bar{m}$ is a median if and only if it minimizes on $\mathbb{R}$ the function*

$$m \mapsto \int_{\mathbb{R}} |x - m| d\mathbf{P}(x).$$

Surprising are: the assumption about order 1, and the fact that medians depend only on the structure of ordered space of $\mathbb{R}$ and not on its metric (nor on its group structure, nor on Haar measure). An “answer”, at least relatively to the “order 1” assumption, is the following.

Well known too³ is the result: *If $\mathbf{P}$ is any law, $x \mapsto |x - m| - |x|$ is $\mathbf{P}$ -integrable (obviously $||x - m| - |x|| \leq |m|$) and $\bar{m}$ is a median if and only if it minimizes on $\mathbb{R}$ the function*

$$m \mapsto \int_{\mathbb{R}} (|x - m| - |x|) d\mathbf{P}(x).$$

The notion of median extends to $\mathbb{R}^d$ and to Banach spaces: see [15, 18, 20] and there exist conditional medians [35]; Kemperman [15] and Milasevic & Ducharme [20] gave in the multivariate case a sufficient condition implying uniqueness of the median (see already in 1948 Haldane [12]). These questions will not be considered here.

The interest of medians comes from robustness, i.e., stability with respect to outliers values. Maybe the notions of means in metric spaces going back to Fréchet (see [11] and many other papers by the same author) should be revisited. See the papers by Armatte [1, 2], the first one containing more than six pages of Fréchet’s references. Numerous authors studied random variables with values in a metric space: for instance [4, 27].

Now we turn to quantiles. T. S. Ferguson [8, Exercise 1.8.3 p. 51, solution in [9]] says that⁴: *If $\tau \in]0, 1[$ and $\mathbf{P}$ is of order 1, $\bar{q}$ is a τ -quantile if and only if it minimizes on $\mathbb{R}$ the function*

$$q \mapsto \int \rho_{\tau}(x - q) d\mathbf{P}(x)$$

where

$$\rho_{\tau}(x) = \frac{1}{2} |x| + \left(\tau - \frac{1}{2} \right) x = \begin{cases} (\tau - 1) x & \text{if } x \leq 0, \\ \tau x & \text{if } x \geq 0. \end{cases} \quad (1)$$

My purpose here is to provide a Convex Analysis proof of the extension to any probability law following a logical way which makes clear the naturalness of the result. See also Koltchinskii [18, Example 2.2 p. 440].

² In French textbooks: [6, Exercice E.11 p. 71] (solution page 28 of [7]), and Théorème 10.1 page 93 in [10]. See also [8, ex. 1.8.2 p. 51] and its solution in [9].

³ Exercises in French Universities and surely in many countries...

⁴ This result is appreciated by specialists: [16, first chapter] and already [17, p. 38].

4. A Convex Analysis point of view

The filled-in graph G defined in Section 2 is a subset of $\mathbb{R}^2$ which is a *maximal monotone operator* [30, Chapter 12] (and therefore a maximal cyclically monotone one because of dimension 1, cf. [30, 12.6 pp. 547–548]). It is known ([30, Th. 12.25 p. 547]) that a maximal cyclically monotone graph is the graph of the subdifferential (definition recalled below) of a l.s.c. convex function (unique up to an additive constant). This is due to Rockafellar; for comments see the *alinea* pages 575–576 of the first edition of [30] (a small shift occurs in further editions) where Moreau is quoted and credited for his famous 1965 paper [23]. Here we will get such a function in a direct way.

Let $\mathbf{F}$ be the primitive (antiderivative) of F null at 0:

$$\mathbf{F}(x) = \begin{cases} \int_0^x F(u) du & \text{if } x \geq 0, \\ -\int_x^0 F(u) du & \text{if } x < 0. \end{cases}$$

Since F is nondecreasing $\mathbf{F}$ is convex. The *subdifferential* of $\mathbf{F}$ at x is

$$\partial \mathbf{F}(x) = \{\ell \in \mathbb{R}; \forall h \in \mathbb{R}, \ell h \leq \mathbf{F}(x+h) - \mathbf{F}(x)\}.$$

Lemma 4.1. *The graph of the multifunction $x \mapsto \partial \mathbf{F}(x)$ is nothing else but the filled-in graph G of the graph of F :*

$$G = \{(x, \ell) \in \mathbb{R}^2; \ell \in \partial \mathbf{F}(x)\}.$$

Proof. Let $\mathbf{F}'(x; w)$ denote the directional derivative of $\mathbf{F}$ at x in the direction w . There holds

$$\partial \mathbf{F}(x) = [-\mathbf{F}'(x; -1), \mathbf{F}'(x; 1)].$$

Then, since F is right continuous,

$$\begin{aligned} \mathbf{F}'(x; 1) &= \lim_{h \searrow 0} \frac{\mathbf{F}(x+h) - \mathbf{F}(x)}{h} \\ &= \lim_{h \searrow 0} \frac{1}{h} \int_x^{x+h} F(u) du \\ &= F(x) \end{aligned}$$

and, since $\mathbf{F}$ is as well the primitive of F^- ,

$$\begin{aligned} \mathbf{F}'(x; -1) &= \lim_{h \searrow 0} \frac{\mathbf{F}(x-h) - \mathbf{F}(x)}{h} \\ &= -\lim_{h \searrow 0} \frac{1}{h} \int_{x-h}^x F^-(u) du \\ &= -F^-(x). \end{aligned}$$

Therefore

$$\partial \mathbf{F}(x) = [F^-(x), F(x)].$$

□

In such a situation a natural question is: what does the minimisation of $\mathbf{F}$ give? Obviously the infimum may be $-\infty$: this is why we cannot consider $\int_{-\infty}^x F(u) du$. But with a slope τ belonging to $]0, 1[$ the function

$$q \mapsto \mathbf{F}(q) - \tau q$$

does achieve minimum(s) and $\bar{q}$ is a minimum if and only if $0 \in -\tau + \partial \mathbf{F}(\bar{q})$. This is equivalent to $(\bar{q}, \tau) \in G$, that is, $\bar{q}$ is a τ -quantile.

Scholium. Let $\mathbf{F}$ denote a primitive of F . For $\tau \in]0, 1[$,

$$\bar{q} \text{ is a } \tau\text{-quantile} \iff \bar{q} \text{ minimizes } q \mapsto \mathbf{F}(q) - \tau q. \quad (2)$$

Next Lemma is an integration by parts result. We will indicate its connexion with Moreau [26, Section 11] and give other references in the Appendix.

Lemma 4.2. Let $a < b$ in $\mathbb{R}$. Then (note the half-open interval $]a, b[$)

$$\int_{]a, b[} x d\mathbf{P}(x) = b F(b) - a F(a) - \int_a^b F(x) dx. \quad (3)$$

Proof (from Schilling [31, Exercise 13.13 p. 133] and its solution on the Net [32, pp. 11–12]). For the product of the Lebesgue measure (we denote it by dx) and of $\mathbf{P}$

$$(dx \otimes d\mathbf{P})(]a, b]^2) = (b - a) [F(b) - F(a)]. \quad (4)$$

And moreover

$$\begin{aligned} & (dx \otimes d\mathbf{P})(]a, b]^2) = \\ &= \iint \mathbf{1}_{]a, b]}(x) \mathbf{1}_{]x, b]}(y) dx d\mathbf{P}(y) + \iint \mathbf{1}_{]a, b]}(x) \mathbf{1}_{]a, x]}(y) dx d\mathbf{P}(y) \\ &= \iint \mathbf{1}_{]a, b]}(x) \mathbf{1}_{]x, b]}(y) dx d\mathbf{P}(y) \stackrel{\text{Tonelli}}{+} \iint \mathbf{1}_{]a, b]}(y) \mathbf{1}_{]y, b]}(x) dx d\mathbf{P}(y) \\ &= \int_{]a, b]} [F(b) - F(x)] dx + \int_{]a, b]} (b - y) d\mathbf{P}(y) \\ &= (b - a) F(b) - \int_a^b F(x) dx + b [F(b) - F(a)] - \int_{]a, b]} y d\mathbf{P}(y). \end{aligned} \quad (5)$$

Comparing (4) and (5) one gets (3). $\square$

Now let⁵ (ρ_τ is defined in (1) above)

$$\Phi(q) := \int_{\mathbb{R}} (\rho_\tau(x - q) - \rho_\tau(x)) d\mathbf{P}(x) \quad (6)$$

(note that the function $x \mapsto \rho_\tau(x - q) - \rho_\tau(x)$ is bounded on $\mathbb{R}$).

⁵ Koltchinskii [18] considers the same functional. See specially two lines after his formula (1.1) page 436 where he sets $f_{P,t}(s) := f_P(s) - st$: there s is my variable q and t my threshold τ . Comparisons are not easy. He also uses (on pp. 439–440) a multivalued integration result [14, Th. 8.3.4]. I got similar results earlier in [33, 34] but Ioffe and Tihomirov published also some papers on these questions before their book.

Theorem 4.3. *Let $\tau \in]0, 1[$ and $\mathbf{P}$ be a law on $\mathbb{R}$. There holds*

$$\Phi(q) = \begin{cases} -\tau q + \int_{]0, q]} F(x) dx & \text{if } q \geq 0, \\ -\tau q - \int_{]q, 0]} F(x) dx & \text{if } q < 0, \end{cases}$$

that is, $\Phi(q) = \mathbf{F}(q) - \tau q$. The function Φ is convex and inf-compact. The value $\bar{q}$ is a τ -quantile if and only if it minimizes the function Φ .

Comments. Part 3) below is a bit technical. Surely careless calculus would be quicker: the formal calculus

$$\begin{aligned} \Phi'(q) &= \int_{\mathbb{R}} \frac{d}{dq} (\rho_{\tau}(x - q) - \rho_{\tau}(x)) d\mathbf{P}(x) \\ &= \int_{-\infty}^q (1 - \tau) d\mathbf{P} + \int_q^{+\infty} (-\tau) d\mathbf{P} \\ &= (1 - \tau) F(q) - \tau (1 - F(q)) \\ &= F(q) - \tau \end{aligned}$$

leads to

$$\Phi'(q) = 0 \iff F(q) = \tau !$$

Proof. 1) Firstly $x \mapsto \rho_{\tau}(x - q) - \rho_{\tau}(x)$ is less than $\tau |q|$ if $\tau \geq 1/2$ or than $(1 - \tau) |q|$ if $\tau \leq 1/2$. Or more roughly $\forall x, |\rho_{\tau}(x - q) - \rho_{\tau}(x)| \leq |q|$. So the integral in (6) is well defined.

2) The function ρ_{τ} defined by (1) being convex, the function

$$q \mapsto \rho_{\tau}(x - q) - \rho_{\tau}(x)$$

is also convex, hence the convexity of Φ .

3) We now prove $\Phi(q) \rightarrow +\infty$ as $|q| \rightarrow +\infty$.

For $q \geq 0$,

$$\rho_{\tau}(x - q) - \rho_{\tau}(x) = \begin{cases} (1 - \tau) q & \text{if } x \in]-\infty, 0], \\ (1 - \tau) q - x & \text{if } x \in]0, q], \\ -\tau q & \text{if } x \in]q, +\infty[, \end{cases}$$

and, for $q < 0$,

$$\rho_{\tau}(x - q) - \rho_{\tau}(x) = \begin{cases} (1 - \tau) q & \text{if } x \in]-\infty, q], \\ x - \tau q & \text{if } x \in]q, 0], \\ -\tau q & \text{if } x \in]0, +\infty[. \end{cases}$$

Hence, if $q \geq 0$ (we will apply formula (3) between lines 2 and 3),

$$\begin{aligned}
 \Phi(q) &= (1 - \tau) q \mathbf{P}(]-\infty, 0]) + \int_{]0, q]} ((1 - \tau) q - x) d\mathbf{P}(x) - \tau q \mathbf{P}(]q, +\infty[) \\
 &= (1 - \tau) q \mathbf{P}(]-\infty, q]) - \int_{]0, q]} x d\mathbf{P}(x) - \tau q \mathbf{P}(]q, +\infty[) \\
 &= (1 - \tau) q \mathbf{P}(]-\infty, q]) - [x F(x)]_0^q + \int_{]0, q]} F(x) dx - \tau q \mathbf{P}(]q, +\infty[) \\
 &= (1 - \tau) q F(q) - q F(q) + \int_{]0, q]} F(x) dx - \tau q (1 - F(q)) \\
 &= -\tau q + \int_{]0, q]} F(x) dx. \tag{7}
 \end{aligned}$$

When $q \rightarrow +\infty$ the mean value of F over $]0, q]$ ultimately exceeds any value in $]\tau, 1[$, so $\Phi(q) \rightarrow +\infty$. (More plainly for $q > \tau$, $\Phi(q) = \int_{]0, \tau]} (F(x) - \tau) dx + \int_{] \tau, q]} (F(x) - \tau) dx$ and $F(x) - \tau \rightarrow 1 - \tau > 0$ as $x \rightarrow \infty$.)

On the other hand, if $q < 0$, (we again apply (3) between lines 2 and 3),

$$\begin{aligned}
 \Phi(q) &= (1 - \tau) q \mathbf{P}(]-\infty, q]) + \int_{]q, 0]} (x - \tau q) d\mathbf{P}(x) - \tau q \mathbf{P}(]0, +\infty[) \\
 &= (1 - \tau) q \mathbf{P}(]-\infty, q]) + \int_{]q, 0]} x d\mathbf{P}(x) - \tau q \mathbf{P}(]q, +\infty[) \\
 &= (1 - \tau) q \mathbf{P}(]-\infty, q]) + [x F(x)]_q^0 - \int_{]q, 0]} F(x) dx - \tau q \mathbf{P}(]q, +\infty[) \\
 &= (1 - \tau) q F(q) - q F(q) - \int_{]q, 0]} F(x) dx - \tau q (1 - F(q)) \\
 &= -\tau q - \int_{]q, 0]} F(x) dx. \tag{8}
 \end{aligned}$$

When $q \rightarrow -\infty$ the mean value of F over $]q, 0]$ ultimately passes under any value in $]0, \tau[$, so $\Phi(q) \rightarrow +\infty$.

Thus the finite valued convex function Φ is inf-compact, so it achieves its infimum over a non-empty compact interval. By (7) and (8)

$$\Phi(q) = \mathbf{F}(q) - \tau q.$$

From (2) the minimizers are the τ -quantiles. □

Corollary 4.4. *Let $\tau \in]0, 1[$ and $\mathbf{P}$ be a first order law on $\mathbb{R}$ then*

$$\bar{q} \text{ is a } \tau\text{-quantile} \iff \bar{q} \text{ minimizes } q \mapsto \int_{\mathbb{R}} \rho_{\tau}(x - q) d\mathbf{P}(x).$$

Proof. The term

$$\int_{\mathbb{R}} \rho_{\tau}(x) d\mathbf{P}(x)$$

is finite and does not depend on q . Hence one can subtract it from the right-hand side of (6). $\square$

5. A direct proof

Quantiles are basic knowledge for Statisticians, and their existence is not a problem. Then their characterization as minimizers is an exercise. Classically the convex function Φ on $\mathbb{R}$ achieves a minimum at $\bar{q}$ if and only if the left and right derivatives of Φ at $\bar{q}$ are respectively ≤ 0 and ≥ 0 .

Theorem 5.1. *Let $\tau \in]0, 1[$ and $\mathbf{P}$ be a law on $\mathbb{R}$. Let Φ defined by (6). Its right derivative at $\bar{q}$ is ≥ 0 if and only if $\mathbf{P}(]-\infty, \bar{q}])$ is $\geq \tau$. The derivative of Φ at $\bar{q}$ in direction -1 is ≥ 0 if and only if $\mathbf{P}(]-\infty, \bar{q}[)$ is $\leq \tau$.*

Proof. Let

$$\varphi_x(q) := \rho_{\tau}(x - q) - \rho_{\tau}(x)$$

so that $\Phi(q) = \int_{\mathbb{R}} \varphi_x(q) d\mathbf{P}(x)$.

a) We are going to check (right derivation under the integral sign)

$$\Phi'(\bar{q}; 1) = \int_{\mathbb{R}} (\varphi_x)'(\bar{q}; 1) d\mathbf{P}(x).$$

Let $(h_n)_n$ a sequence in $]0, +\infty[$ decreasing to 0. There holds

$$\Phi'(\bar{q}; 1) = \lim_n \frac{\Phi(\bar{q} + h_n) - \Phi(\bar{q})}{h_n} = \lim_n \int_{\mathbb{R}} \frac{\varphi_x(\bar{q} + h_n) - \varphi_x(\bar{q})}{h_n} d\mathbf{P}(x).$$

The last integrand decreases to $(\varphi_x)'(\bar{q}; 1)$ and Lebesgue's theorem applies⁶ thanks to the inequalities

$$-[\varphi_x(\bar{q} - 1) - \varphi_x(\bar{q})] \leq \frac{\varphi_x(\bar{q} + h_n) - \varphi_x(\bar{q})}{h_n} \leq \frac{\varphi_x(\bar{q} + h_0) - \varphi_x(\bar{q})}{h_0}.$$

Now it is easy to calculate

$$(\varphi_x)'(\bar{q}; 1) = \begin{cases} -\tau & \text{if } x \in]\bar{q}, +\infty[, \\ -(\tau - 1) & \text{if } x \in]-\infty, \bar{q}]. \end{cases}$$

⁶ The monotone convergence theorem applies, but due to the non-increasing property of the sequence, one should observe that the whole sequence is bounded from above by the first term, say the one with $n = 0$. See [18, Section 2 page 439] for quoting monotone convergence.

Hence

$$\begin{aligned}\Phi'(\bar{q}; 1) &= (1 - \tau) \mathbf{P}([-\infty, \bar{q}]) + (-\tau) \mathbf{P}([\bar{q}, +\infty[) \\ &= \mathbf{P}([-\infty, \bar{q}]) - \tau.\end{aligned}$$

Thus the right derivative is ≥ 0 if and only if $\mathbf{P}([-\infty, \bar{q}])$ is $\geq \tau$.

b) Similarly⁷ the directional derivative in the direction -1 is ≥ 0 if and only if $\mathbf{P}([-\infty, \bar{q}[)$ is $\leq \tau$. $\square$

6. Mean or center of mass

This Section continues our discussion about hypotheses. These are obvious facts. If $\mathbf{P}$ is of order 2 the mean $\mathbb{E}(X)$ is the unique m in $\mathbb{R}$ minimizing

$$\mathbb{E}(|X - m|^2) = \int_{\mathbb{R}} |x - m|^2 d\mathbf{P}(dx).$$

This comes from the Huygens formula

$$\mathbb{E}(|X - m|^2) = |m - \mathbb{E}(X)|^2 + \mathbf{v}(X)$$

(in Mechanics this is a formula about moments) where $\mathbf{v}(X)$ is the *variance*: in $L^2(\Omega)$ the random variables $X - \mathbb{E}(X)$ and the constant $(\mathbb{E}(X) - m) \mathbf{1}_{\Omega}$ are orthogonal.

As in Section 3 assuming “ $\mathbf{P}$ is of order 2” may seem too much since the mean exists as soon as “ $\mathbf{P}$ is of order 1”. But the very elementary following holds. If $\mathbf{P}$ is only of order 1 the function $x \mapsto |x - m|^2 - x^2$ is $\mathbf{P}$ -integrable and it remains that m is the mean if and only if it minimizes

$$\int_{\mathbb{R}} (|x - m|^2 - x^2) d\mathbf{P}(dx).$$

Indeed $|x - m|^2 - x^2 = m^2 - 2mx$ and so

$$\int_{\mathbb{R}} (|x - m|^2 - x^2) d\mathbf{P}(dx) = m^2 - 2m \mathbb{E}(X) = (m - \mathbb{E}(X))^2 - \mathbb{E}(X)^2.$$

For other means in metric spaces, see Fréchet [11]. Instead of subtracting a function independent of m , Fréchet uses limits.

⁷ Minus signs are always more perilous. In the multivariate case one works with linear vector space and this technicality disappears.

A. Appendix

In [26, Section 11] Moreau studied carefully the action of a bilinear form on BV functions. In particular, he provided a systematic development of more ancient results.

Here is one statement by J. J. Moreau. Essential for understanding is that x and y are BV (or even locally BV) functions defined on the real interval I , Φ is a bilinear mapping and d is for the *differential* (or Stieltjes) measure obtained from a BV function. (In [26] x , y and Φ are Banach-valued.)

Corollary 11.4. For every $[a, b] \subset I$, under Convention 3.5 if a or b happen to equal an end of I , one has

$$\int_{[a,b]} \Phi(dx, y^-) + \int_{[a,b]} \Phi(x^+, dy) = \Phi(x^+(b), y^+(b)) - \Phi(x^-(a), y^-(a)) \quad (11.9)$$

and analogous formulas corresponding to (11.2) and (11.3).

We are going to deduce Lemma 2 from (11.9): $I = \mathbb{R}$, the BV functions are real valued, Φ is the usual product of real numbers. More precisely $x = F$ (the cumulative function of $\mathbf{P}$) and $y = \text{id}_{\mathbb{R}}$. So (11.9) rewrites

$$\int_{[a,b]} y dF + \int_{[a,b]} F dy = b F(b) - a F^-(a).$$

Subtracting the integral $\int_{\{a\}} y dF$ from both sides one recovers (3).

For the curious connoisseurs the corollary above follows from Proposition 11.1 p. 39 and specially from formula (11.1) which asserts

$$d\Phi(x, y) = \Phi(dx, y^-) + \Phi(x^+, dy)$$

the other formulas (11.2) and (11.3) being

$$\begin{aligned} d\Phi(x, y) &= \Phi(dx, y^+) + \Phi(x^-, dy) \\ d\Phi(x, y) &= \Phi\left(dx, \frac{y^+ + y^-}{2}\right) + \Phi\left(\frac{x^+ + x^-}{2}, dy\right). \end{aligned}$$

These formulas are related to Integration by parts results. We could refer to the 1976 paper by Rockafellar [28, Prop. 1 pp. 161–162] which relies on the 1966 book [3, Prop. 8.5.5 pp. 374–375] (these two references are in [26]). Hildebrandt in 1963 gave a lot of Integration by parts results [13, pages 53, 86, 93, 127 and 342]. And the first who proved an integration by parts result for general functions was König [19] in 1897! About König see the remark on page 76 of R. M. Dudley & R. Norvaiša <http://www.maphysto.dk/publications/MPS-LN/1998/1.pdf> (title *An introduction to p-variation and Young integrals, With emphasis on sample functions of stochastic processes*).

Acknowledgements. Thanks to Manuel Monteiro Marques, Paul Raynaud de Fitte and Lionel Thibault for their comments. They are not responsible of possible weaknesses.

References

- [1] M. Armatte: Maurice Fréchet statisticien, enquêteur et agitateur public, *Rev. Hist. Math.* 7 (2001) 7–65.
- [2] M. Armatte: Fréchet et la médiane: un moment dans une histoire de la robustesse, *J. Soc. Franç. Stat.* 147 (2006) 23–37.
- [3] E. Asplund, L. Bungart: *A First Course in Integration*, Holt, Rinehart and Winston, New York (1966).
- [4] N. Belili, H. Heinich: Median for metric spaces, *Appl. Math.* 28 (2001) 191–209.
- [5] P. Chaudhuri: On a geometric notion of quantiles for multivariate data, *J. Amer. Stat. Assoc.* 91 (1996) 862–872.
- [6] D. Dacunha-Castelle, M. Duflo: *Probabilités et Statistiques. 1. Problèmes à Temps Fixe*, Masson, Paris (1982).
- [7] D. Dacunha-Castelle, M. Duflo: *Exercices de Probabilités et Statistiques. 1. Problèmes à Temps Fixe*, Masson, Paris (1982).
- [8] T. S. Ferguson: *Mathematical Statistics: A Decision Theoretic Approach*, Probability and Mathematical Statistics 1, Academic Press, New York (1967).
- [9] T. S. Ferguson: Solutions to selected exercises, available at: <http://www.math.ucla.edu/~tom/MathematicalStatistics/Sec18.pdf>.
- [10] D. Foata, A. Fuchs: *Calcul des Probabilités*, 2nd Ed., Dunod, Paris (1998).
- [11] M. Fréchet: Les éléments aléatoires de nature quelconque dans un espace distancié, *Ann. Inst. H. Poincaré* 10 (1948) 215–310.
- [12] J. B. S. Haldane: Note on the median of a multivariate distribution, *Biometrika* 35 (1948) 414–415.
- [13] T. H. Hildebrandt: *Introduction to the Theory of Integration*, Academic Press, New York (1963).
- [14] A. D. Ioffe, V. M. Tihomirov: *Theory of Extremal Problems*, Studies in Mathematics and its Applications 6, North-Holland, Amsterdam (1979).
- [15] J. H. B. Kemperman: The median of a finite measure on a Banach space, *Statistical Data Analysis Based on the L_1 -Norm and Related Methods* (Neuchâtel, 1987), Y. Dodge (ed.), North-Holland, Amsterdam (1987) 217–230.
- [16] R. Koenker: *Quantile Regression*, Econometric Society Monographs 38, Cambridge University Press, Cambridge (2005).
- [17] R. Koenker, G. Bassett: Regression Quantiles, *Econometrica* 46 (1978) 33–50.
- [18] V. I. Koltchinskii: M -estimation, convexity and quantiles, *Ann. Stat.* 25 (1997) 435–477.
- [19] G. König: A határozott integrálok elméletéhez (On the theory of definite integrals), *Mathematikai és Természettudományi Érteztő* 15 (1897) 380–384.

- [20] P. Milasevic, G. R. Ducharme: Uniqueness of the spatial median, *Ann. Stat.* 15 (1987) 1332–1333.
- [21] M. D. P. Monteiro Marques: Regularization and graph approximation of a discontinuous evolution problem, *J. Differ. Equations* 67 (1987) 145–164.
- [22] M. D. P. Monteiro Marques: Differential Inclusions in Nonsmooth Mechanical Problems. Shocks and Dry Friction, *Progress in Nonlinear Differential Equations and their Applications* 9, Birkhäuser, Basel (1993).
- [23] J. J. Moreau: Proximité et dualité dans un espace hilbertien, *Bull. Soc. Math. France* 93 (1965) 273–299.
- [24] J. J. Moreau: Fonctionnelles convexes, Séminaire sur les Équations aux Dérivées Partielles (1966–1967) II, Collège de France, Paris (1967), 108 p.
- [25] J. J. Moreau: Approximation en graphe d’une évolution discontinue, *RAIRO, Anal. Numér.* 12 (1978) 75–84.
- [26] J. J. Moreau: Bounded variation in time, in: *Topics in Nonsmooth Mechanics*, J. J. Moreau et al. (ed.), Birkhäuser, Basel (1988) 1–74.
- [27] P. Raynaud de Fitte: Théorème ergodique ponctuel et lois fortes des grands nombres pour des points aléatoires d’un espace métrique à courbure négative, *Ann. Probab.* 25 (1997) 738–766.
- [28] R. T. Rockafellar: Dual problems of Lagrange for arcs of bounded variation, in: *Calculus of Variations and Control Theory* (Madison, 1975), D. L. Russell (ed.), Academic Press, New York (1976) 155–192.
- [29] R. T. Rockafellar, J. O. Royset: Random variables, monotone relations and convex analysis, *Math. Program.* 148 (2014) 297–331.
- [30] R. T. Rockafellar, R. J.-B. Wets: *Variational Analysis*, Grundlehren der Mathematischen Wissenschaften 317, Springer, Berlin (1998).
- [31] R. L. Schilling: *Measures, Integrals and Martingales*, Cambridge University Press, New York (2005).
- [32] R. L. Schilling: *Measures, Integrals & Martingales*, Solution Manual Chapter 13–24, available at: http://www.motapa.de/measures_integrals_and_martingales/solutions13-24.pdf.
- [33] M. Valadier: Sur l’intégration d’ensembles convexes compacts en dimension infinie, *C. R. Acad. Sci., Paris, Sér. A* 266 (1968) 14–16.
- [34] M. Valadier: Sous-différentiels d’une borne supérieure et d’une somme continue de fonctions convexes, *C. R. Acad. Sci., Paris, Sér. A* 268 (1969) 39–42.
- [35] M. Valadier: La multi-application médianes conditionnelles, *Z. Wahrscheinlichkeitstheor. Verw. Geb.* 67 (1984) 279–282.
- [36] M. Valadier: Quantiles as minimizers, arXiv:1411.2732 [math.ST] (2014).