

## Chapter 3

### The moment problem for finitely additive probabilities

Enrique Miranda, Gert de Cooman and Erik Quaeghebeur\*  
*Rey Juan Carlos University, Madrid, Spain (E. Miranda)*  
*SYSTeMS Research Group, Ghent University, Ghent, Belgium*  
*(G. de Cooman and E. Quaeghebeur)*

We study the moment problem for finitely additive probabilities and show that the information provided by the moments is equivalent to the one given by the associated lower and upper distribution functions.

**Keywords:** Moment problem, coherent lower previsions, lower and upper distribution functions.

#### 3.1. Introduction

The moment problem in probability theory refers to the existence and uniqueness of a  $\sigma$ -additive probability measure with a particular sequence of moments  $m$ , or equivalently  $m_k$ ,  $k \geq 0$ , i.e., of a measure  $\mu$  such that  $\int x^k d\mu = m_k$  for all  $k \geq 0$ . There are three classical moment problems: the Hamburger moment problem, where the support of  $\mu$  is the real line; the Stieltjes moment problem, where it is  $[0, +\infty)$ ; and the Hausdorff moment problem, where it is the closed interval  $[0, 1]$ . This last problem is the one we shall consider in this paper.

Hausdorff [4, 5] has proved that there is a solution to the moment problem for a moment sequence  $m$  if and only if this sequence is completely monotone, meaning that  $(-1)^n \Delta^n m_k \geq 0$  for all  $m, k \geq 0$ , with  $\Delta^n m_k = \Delta^{n-1} m_{k+1} - \Delta^{n-1} m_k$  for  $n \geq 1$  and  $\Delta^0 m_k = m_k$ . In this case, the existence of a probability measure  $\mu$  with a sequence of moments implies its uniqueness, by virtue of the Riesz Representation Theorem.

In this paper, we study the Hausdorff moment problem for finitely

---

\*Extended versions of parts this paper, with proofs, are available [6, 7]. E-mail addresses: [enrique.miranda@urjc.es](mailto:enrique.miranda@urjc.es), [Gert.deCooman@UGent.be](mailto:Gert.deCooman@UGent.be), [Erik.Quaeghebeur@UGent.be](mailto:Erik.Quaeghebeur@UGent.be).

additive probabilities. We consider a sequence  $m$  of real numbers and study whether there is a unique finitely additive probability (or *probability charge*) with this sequence of moments. We shall see that in this case the question of the existence of such a probability charge is a fairly trivial one, but the study of its unicity becomes much more involved as soon as we let go of the countable additivity (or continuity) axiom. Hence, it will be important to study for which functions (and in particular on which events) the expectation with respect to such probability charges  $\mu$  is uniquely determined by the moments. It turns out that studying and solving this problem can be done quite efficiently using the language of Walley's behavioural theory of coherent lower previsions [8].

### 3.2. A short introduction to lower previsions

Let us give a short introduction to those concepts from the theory of coherent lower previsions that we shall use in this paper. We refer to [8] for their behavioural interpretation, and for a much more complete introduction and treatment. Consider a non-empty set  $\Omega$ . Then a *gamble* on  $\Omega$  is a bounded real-valued function on  $\Omega$ . We denote the set of all gambles on  $\Omega$  by  $\mathcal{L}(\Omega)$ .

A *lower prevision*  $\underline{P}$  is a real-valued map defined on some subset  $\mathcal{K}$  of  $\mathcal{L}(\Omega)$ . If the domain  $\mathcal{K}$  of  $\underline{P}$  only contains (indicators of) events  $A$ , then  $\underline{P}$  is also called a *lower probability*. We also write  $\underline{P}(I_A)$  as  $\underline{P}(A)$ , the lower probability of the event  $A$ . The *conjugate upper prevision*  $\overline{P}$  of  $\underline{P}$  is defined on  $-\mathcal{K}$  by  $\overline{P}(f) := -\underline{P}(-f)$  for every  $-f$  in the domain of  $\underline{P}$ . If the domain of  $\overline{P}$  contains indicators only, then  $\overline{P}$  is also called an *upper probability*.

A lower prevision  $\underline{P}$  defined on  $\mathcal{L}(\Omega)$  is called *coherent* if it is super-additive:  $\underline{P}(f + g) \geq \underline{P}(f) + \underline{P}(g)$ , positively homogeneous:  $\underline{P}(\lambda f) = \lambda \underline{P}(f)$  for  $\lambda \geq 0$ , and positive:  $\underline{P}(h) \geq \inf h$ . A lower prevision  $\underline{P}$  on a general domain is then called *coherent* if it can be extended to some coherent lower prevision on all gambles. This is the case if and only if  $\sup[\sum_{i=1}^n f_i - m f_0] \geq \sum_{i=1}^n \underline{P}(f_i) - m \underline{P}(f_0)$  for any  $n, m \geq 0$ , and  $f_0, f_1, \dots, f_n$  in the domain of  $\underline{P}$ .

A *linear prevision*  $P$  on  $\mathcal{L}(\Omega)$  is a coherent lower prevision that is self-conjugate:  $P(-f) = -P(f)$ , or in other words, a linear functional that is positive and normalised:  $P(1) = 1$  (1 also used as a constant function). A functional defined on an arbitrary subset  $\mathcal{K}$  of  $\mathcal{L}(\Omega)$  is called a *linear prevision* if it can be extended to a linear prevision on  $\mathcal{L}(\Omega)$ . This is the case if and only if  $\sup[\sum_{i=1}^n f_i - \sum_{j=1}^m g_j] \geq \sum_{i=1}^n P(f_i) - \sum_{j=1}^m P(g_j)$  for any  $n, m \geq 0$ , and  $f_1, \dots, f_n, g_1, \dots, g_m$  in the domain of  $P$ . The restriction

$Q$  of a linear prevision  $P$  on  $\mathcal{L}(\Omega)$  to the set  $\wp(\Omega)$  of all events is a finitely additive probability measure (*probability charge*). Moreover, it holds that  $P(h) = (D) \int h dQ$  for any gamble  $h$  (Dunford integral). Hence, linear previsions are completely determined by the values they assume on events, and are simply expectations with respect to finitely additive probabilities.

The *natural extension*  $\underline{E}_P$  to  $\mathcal{L}(\Omega)$  of a coherent lower prevision  $\underline{P}$  defined on  $\mathcal{K}$ , is the point-wise smallest coherent lower prevision that extends  $\underline{P}$  to all gambles. It is equal to the lower envelope of the set  $\mathcal{M}(\underline{P})$  of all linear previsions that point-wise dominate  $\underline{P}$  on its domain  $\mathcal{K}$ :  $\underline{E}_P(f) = \min_{Q \in \mathcal{M}(\underline{P})} Q(f)$  for any gamble  $f$  in  $\mathcal{L}(\Omega)$ . Moreover,  $\mathcal{M}(\underline{E}_P) = \mathcal{M}(\underline{P})$ .

A coherent lower prevision defined on a *lattice* of gambles  $\mathcal{K}$  (a set of gambles closed under pointwise minima and maxima) is called *n-monotone* if  $\sum_{I \subseteq \{1, \dots, p\}} (-1)^{|I|} \underline{P}(f \wedge \bigwedge_{i \in I} f_i) \geq 0$  for all  $1 \leq p \leq n$ , and all  $f, f_1, \dots, f_p$  in  $\mathcal{K}$ . A coherent lower prevision is *completely monotone* when it is *n-monotone* for any  $n \geq 1$ . A thorough study of *n-monotone* coherent lower previsions, and their properties, can be found in [1, 2].

### 3.3. Formulation and initial solution of the problem

We are now ready to formulate the moment problem using the language established in the previous section. Consider a moment sequence  $m$  and the subset  $\mathcal{V}_p([0, 1])$  of the set of all polynomials on the unit interval given by  $\mathcal{V}_p([0, 1]) := \{p^k : k \geq 0\}$ , where  $p^k(x) = x^k$ . We define a functional  $P_m$  on this set by letting  $P_m(p^k) := m_k$ . This functional can be *uniquely* extended to a linear functional  $\hat{P}_m$  on the set of all polynomials as follows:  $\hat{P}_m(\sum_{k=0}^n a_k p^k) = \sum_{k=0}^n a_k m_k$ . The following theorem summarises a number of results from the literature. It tells us under what conditions there exists a *linear prevision* on the set  $\mathcal{L}([0, 1])$  of all gambles on  $[0, 1]$ , or equivalently a *finitely additive probability* on the set  $\wp([0, 1])$  of all subsets of  $[0, 1]$ , for which the moments are given by  $m$ .

**Theorem 3.1.** *The following are equivalent.*

- (1) *The functional  $P_m$  can be extended uniquely to a linear prevision on the set  $\mathcal{C}([0, 1])$  of all continuous gambles on  $[0, 1]$ . We shall denote this extension by  $\hat{P}_m$ .*
- (2) *For all polynomials  $q = \sum_{k=0}^n a_k p^k$ ,  $\min q \leq \hat{P}_m(q) \leq \max q$ .*
- (3) *The moment sequence  $m$  satisfies the Hausdorff moment condition [4, 5]:  $m_0 = 1$  and  $m$  is completely monotone, meaning that  $(-1)^n \Delta^n m_k \geq 0$  for all  $k \geq 0$  and  $n \geq 0$ .*

So we see that the Hausdorff moment problem has a solution (as a linear prevision) if and only if the Hausdorff moment condition is satisfied, and in that case the solution is uniquely determined on the set  $\mathcal{C}([0, 1])$  of continuous functions.

If we invoke the Riesz Representation Theorem (see for instance [3, Section V.1]), we see that there is a *unique*  $\sigma$ -additive probability measure  $P_m^\sigma$  on the Borel sets of  $[0, 1]$ , and a *unique* (right-continuous) distribution function  $F_m^\sigma$  on  $[0, 1]$  such that for all continuous gambles  $h$ , the expectation  $E_m^\sigma(h)$  is equal to  $(L) \int h dP_m^\sigma = (LS) \int h dF_m^\sigma = \hat{P}_m(h)$ , where the first integral is the Lebesgue integral associated with the probability measure  $P_m^\sigma$ , and the second integral the Lebesgue–Stieltjes integral associated with the distribution function  $F_m^\sigma$ . Also,  $F_m^\sigma(x) = P_m^\sigma([0, x])$ . Note that, actually, the expectation operator  $E_m^\sigma$ , as well as both integrals are defined for all Borel-measurable functions on  $[0, 1]$ .

In this sense, the moments determine a unique  $\sigma$ -additive probability measure on the Borel sets. But the solution is not as clear-cut if we look for the finitely additive probabilities on all events (or equivalently the linear previsions on all gambles) that correspond to the given moments. These are given by the set  $\mathcal{M}(\hat{P}_m)$  of all linear previsions  $Q$  that dominate, or, equivalently, coincide with,  $\hat{P}_m$  on continuous gambles.

For any gamble  $h$  on  $[0, 1]$ , it follows that the linear previsions that solve the moment problem can assume a value in the real interval  $[\underline{E}_m(h), \overline{E}_m(h)]$ , where  $\underline{E}_m(h) = \inf\{Q(h) : Q \in \mathcal{M}(\hat{P}_m)\}$  and  $\overline{E}_m(h) = \sup\{Q(h) : Q \in \mathcal{M}(\hat{P}_m)\}$ . In fact, given any real number  $a$  in this interval, there will be a solution  $Q$  to the moment problem such that  $Q(h) = a$ . The functional  $\underline{E}_m$  on  $\mathcal{L}([0, 1])$  is the natural extension of  $\hat{P}_m$ , and it is the point-wise smallest coherent lower prevision that coincides with  $\hat{P}_m$  on  $\mathcal{C}([0, 1])$ . The functional  $\overline{E}_m$  is its conjugate upper prevision and satisfies  $\overline{E}_m(h) = -\underline{E}_m(-h)$  for all gambles  $h$  on  $[0, 1]$ .

$\underline{E}_m$  is the smallest coherent lower prevision that satisfies  $\underline{E}_m(p^k) = \overline{E}_m(p^k) = m_k$ ,  $k \geq 0$ . It is easy to see (use [8, Theorem 3.4.1]) that  $\mathcal{M}(\underline{E}_m) = \mathcal{M}(\hat{P}_m) = \mathcal{M}(P_m)$ . So we see that the lower prevision  $\underline{E}_m$  *completely determines* the solution to the Hausdorff moment problem for linear previsions. In particular, the gambles  $h$  on  $[0, 1]$  where the lower and upper natural extensions coincide, i.e.  $\underline{E}_m(h) = \overline{E}_m(h)$ , are precisely those gambles to which  $P_m$  has a unique extension. We shall call such gambles *m-integrable*. One of the goals in this paper is precisely to study these *m-integrable* gambles. Another, closely related goal, is to study the functional  $\underline{E}_m$ .

With  $\underline{E}_m$  and its conjugate  $\overline{E}_m$  we can associate a *lower distribution function*  $\underline{F}_m$  and an *upper distribution function*  $\overline{F}_m$  on  $[0, 1]$ ,

$$\underline{F}_m(x) = \underline{E}_m([0, x]), \quad \overline{F}_m(x) = \overline{E}_m([0, x]),$$

for  $x \in [0, 1]$ . We then ask: what are the properties of these distribution functions, what is their relationship to  $F_m^\sigma$ , and to what extent do they determine the functional  $\underline{E}_m$ , and therefore the solution to the Hausdorff moment problem?

### 3.4. The natural extension $\underline{E}_m$ and $m$ -integrable gambles

Since  $\mathcal{C}([0, 1])$  is a linear subspace of  $\mathcal{L}([0, 1])$  that contains all constant gambles, we may apply another known result [8, Corollary 3.1.8] from the theory of coherent lower previsions to obtain the following expressions: for any gamble  $h$  on  $[0, 1]$ ,  $\underline{E}_m(h) = \sup\{\hat{P}_m(g) : g \in \mathcal{C}([0, 1]), g \leq h\}$  and  $\overline{E}_m(h) = \inf\{\hat{P}_m(g) : g \in \mathcal{C}([0, 1]), h \leq g\}$ . We use these expressions to prove a number of interesting properties of  $\underline{E}_m$  and the lower and upper distribution functions  $\underline{F}_m$  and  $\overline{F}_m$ .

**Proposition 3.1.** *Consider a moment sequence  $m$  satisfying the Hausdorff moment condition.*

- (1)  $0 \leq \underline{F}_m \leq F_m^\sigma \leq \overline{F}_m \leq 1$ .
- (2)  $\underline{F}_m$  and  $\overline{F}_m$  are non-decreasing functions.
- (3)  $\underline{F}_m(0) = 0$  and  $\underline{F}_m(1) = \overline{F}_m(1) = 1$ .

It follows from Proposition 3.1 that the left and right limits of  $\underline{F}_m$  and  $\overline{F}_m$  exist everywhere. Let us denote by  $\mathcal{D}_{\underline{F}_m} := \{x \in [0, 1] : \underline{F}_m(x+) \neq \underline{F}_m(x-)\}$  the set of all points of discontinuity of  $\underline{F}_m$ , and similarly by  $\mathcal{D}_{\overline{F}_m} = \{x \in [0, 1] : \overline{F}_m(x-) \neq \overline{F}_m(x+)\}$  the set of points where  $\overline{F}_m$  is not continuous. Let  $\mathcal{D}_m := \mathcal{D}_{\underline{F}_m} \cup \mathcal{D}_{\overline{F}_m}$  denote their union. It follows from the non-decreasing character of  $\underline{F}_m$  and  $\overline{F}_m$  that  $\mathcal{D}_{\underline{F}_m}$ ,  $\mathcal{D}_{\overline{F}_m}$  and  $\mathcal{D}_m$  are countable subsets of  $[0, 1]$ .

**Proposition 3.2.** *Consider a moment sequence  $m$  satisfying the Hausdorff moment condition.*

- (1) For any  $x \in [0, 1]$ ,  $\underline{F}_m(x+) = \overline{F}_m(x) = \overline{F}_m^\sigma(x) = \overline{F}_m(x+)$ .
- (2)  $\underline{F}_m(x-) = \underline{F}_m(x) = \overline{F}_m(x-) \forall x \in (0, 1)$ .
- (3)  $\underline{F}_m(0-) := \underline{F}_m(0) \leq \overline{F}_m(0-) := \overline{F}_m(0)$ .
- (4)  $\underline{F}_m(1-) = \overline{F}_m(1-) \leq \underline{F}_m(1) = \overline{F}_m(1)$ .

- (5)  $\mathcal{D}_{\underline{F}_m} \cap (0, 1] = \mathcal{D}_{\overline{F}_m} \cap (0, 1]$ .  
 (6) The following statements are equivalent for all  $x \in (0, 1)$ : (i)  $x \notin \mathcal{D}_m$ ;  
 (ii)  $\underline{F}_m(x) = \overline{F}_m(x)$ ; and (iii)  $F_m^\sigma$  is continuous in  $x$ .

Hence, if  $m$  is a sequence satisfying the Hausdorff moment condition, the distribution function  $F_m^\sigma$  of the unique  $\sigma$ -additive probability with these moments is equal to the upper distribution function  $\overline{F}_m$ .

**Example 3.1.** Consider the moment sequence  $m$  given by  $m_0 = 1$ ,  $m_k = 0$ ,  $k > 0$ . It is completely monotone, because the probability measure all of whose mass is concentrated in 0 has these moments. If we consider the  $\underline{F}_m$  and  $\overline{F}_m$  produced by this sequence, it is easy to check that the fifth and sixth statements of Proposition 3.2 do not hold for  $x = 0$ , and that the inequality in the third statement may be strict.

Let us now define, for any gamble  $h$  on  $[0, 1]$ , the gambles  $h^\downarrow$  and  $h^\uparrow$  on  $[0, 1]$  by  $h^\uparrow(x) = \sup\{g(x) : g \in \mathcal{C}([0, 1]), g \leq h\}$  and  $h^\downarrow(x) = \inf\{g(x) : g \in \mathcal{C}([0, 1]), h \leq g\}$  for all  $x$  in  $[0, 1]$ . Then  $h^\uparrow$  is the pointwise greatest lower semi-continuous gamble that is dominated by  $h$ , and  $h^\downarrow$  is the pointwise smallest upper semi-continuous gamble that dominates  $h$ . Observe also that for any  $A \subseteq [0, 1]$ ,  $(I_A)^\uparrow = I_{\text{int}(A)}$  and  $(I_A)^\downarrow = I_{\text{cl}(A)}$ , where  $\text{int}(A)$  is the topological interior of  $A$ , and  $\text{cl}(A)$  its topological closure.

**Proposition 3.3.** Consider a moment sequence  $m$  satisfying the Hausdorff moment condition. Then for any gamble  $h$  on  $[0, 1]$  we have that  $\underline{E}_m(h) = \underline{E}_m(h^\uparrow)$  and  $\overline{E}_m(h) = \overline{E}_m(h^\downarrow)$ . In particular, for any  $A \subseteq [0, 1]$ ,  $\underline{E}_m(\text{int}(A)) = \underline{E}_m(A)$  and  $\overline{E}_m(\text{cl}(A)) = \overline{E}_m(A)$ .

Now consider, for any set  $A$  its interior  $\text{int}(A)$ . It is easy to check that  $\text{int}(A)$  is a countable union of disjoint open intervals. The following important result now tells us that it even suffices to know the values of  $\underline{E}_m$  on open intervals.

**Proposition 3.4.** Consider a moment sequence  $m$  satisfying the Hausdorff moment condition. Let  $B$  be a countable union of disjoint open intervals  $B_n$ ,  $n \geq 0$ . Then  $\underline{E}_m(B) = \sup_{n \geq 0} \sum_{k=0}^n \underline{E}_m(B_k)$ .

Summarising,  $\underline{E}_m$  is completely determined on events if we know its values on all open intervals. The following proposition establishes that these values are determined by the lower and upper distribution functions  $\underline{F}_m$

and  $\overline{F}_m$ .

**Proposition 3.5.** *Consider a moment sequence  $m$  satisfying the Hausdorff moment condition.*

- (1)  $\underline{E}_m([0, x]) = \underline{F}_m(x-)$  for all  $x \in [0, 1]$ .
- (2)  $\underline{E}_m((x, 1]) = 1 - \overline{F}_m(x)$  for all  $x \in [0, 1]$ .
- (3)  $\underline{E}_m((x, y)) = \underline{F}_m(y-) - \overline{F}_m(x)$  for all  $0 \leq x < y \leq 1$ .
- (4) For all  $0 \leq x < y \leq 1$ , the interval  $(x, y)$  is  $m$ -integrable if and only if  $x$  and  $y$  do not belong to  $\mathcal{D}_m$ . For  $x \in [0, 1]$ ,  $[0, x)$  and  $(x, 1]$  are  $m$ -integrable if and only if  $x$  does not belong to  $\mathcal{D}_m$ .

We can also deduce from these results that there is *never* a unique linear prevision that solves the Hausdorff moment problem.

**Remark 3.1.** Consider the set  $\mathbb{Q} \cap [0, 1]$  of all rational numbers between zero and one, then  $\text{int}(\mathbb{Q} \cap [0, 1]) = \emptyset$  and  $\text{cl}(\mathbb{Q} \cap [0, 1]) = [0, 1]$ , so we infer from Proposition 3.3 that  $\underline{E}_m(\mathbb{Q} \cap [0, 1]) = \underline{E}_m(\emptyset) = 0$  and  $\overline{E}_m(\mathbb{Q} \cap [0, 1]) = \overline{E}_m([0, 1]) = 1$ . This shows that there is always either none (when  $m$  is not completely monotone) or an uncountable infinity of linear previsions that produce a given sequence of moments  $m_k$ ,  $k \geq 0$ .

There are two further questions we would still like to look at in this section. First of all, are the values of  $\underline{E}_m$  on events also completely determined by  $\underline{E}_m$  and  $\overline{F}_m$  in their points of continuity, or in other words, by  $F_m^\sigma$  in its points of continuity? By virtue of Proposition 3.5, this comes down to  $\underline{E}_m$  being determined by its values on  $m$ -integrable open intervals. And secondly, can we say something similar about the values that  $\underline{E}_m$  assumes on gambles, and not just events? We shall answer both questions in the positive in Theorem 3.3 further on.

But before we can address these issues, we need to prepare ourselves a bit better. In order to answer the first question, it will help us to consider the set of all  $m$ -integrable open intervals. By Proposition 3.5 this is the set  $\{[0, 1], [0, x), (x, y), (y, 1] : x, y \notin \mathcal{D}_m\}$ . This set is closed under intersections, so the lattice of events  $\mathcal{O}_m$  generated by all  $m$ -integrable open intervals is the set made up of all finite unions of  $m$ -integrable disjoint open intervals. For the second question, let  $\mathcal{L}_m$  denote the class of  $m$ -integrable gambles,  $\mathcal{L}_m := \{h \in \mathcal{L}([0, 1]) : \underline{E}_m(h) = \overline{E}_m(h)\}$ , and let  $\mathcal{F}_m$  denote the class of  $m$ -integrable events, i.e., those events with  $m$ -integrable indicators. Then

we have the following [1, 2]:

**Theorem 3.2.** *Consider a moment sequence  $m$  satisfying the Hausdorff moment condition.*

- (1)  $\underline{E}_m$  is the natural extension of its restriction to events, which is a completely monotone coherent lower probability.
- (2)  $\underline{E}_m$  is completely monotone on  $\mathcal{L}([0, 1])$ , and  $\underline{E}_m(h) = (C) \int h d\underline{E}_m$  for any  $h \in \mathcal{L}([0, 1])$ , where the integral is the Choquet integral of  $h$  with respect to  $\underline{E}_m$ , equal to (using the Riemann integral)  $\inf h + (R) \int_{\inf h}^{\sup h} \underline{E}_m(\{h > t\}) dt$ .
- (3)  $\mathcal{L}_m$  is a uniformly closed linear lattice that contains all constant gambles.
- (4)  $\mathcal{F}_m$  is a field of subsets of  $[0, 1]$  that includes  $\mathcal{O}_m$ .
- (5) A gamble  $f$  is  $m$ -integrable if and only if its cut sets  $\{f \geq t\} := \{x \in [0, 1] : f(x) \geq t\}$  are  $m$ -integrable for all but a countable number of  $t$  in  $[0, 1]$ .

Let us denote by  $\tilde{P}_m$  the restriction of  $\underline{E}_m$  to  $\mathcal{O}_m$ . Then we know by Proposition 3.4 that  $\tilde{P}_m$  is additive on this lattice of events. We show next that  $\underline{E}_m$  is completely determined by  $\tilde{P}_m$ .

**Theorem 3.3.** *Consider a moment sequence  $m$  satisfying the Hausdorff moment condition.*

- (1) The natural extension of  $\tilde{P}_m$  to all events is the inner set function  $\tilde{P}_{m,*}$  of  $\tilde{P}_m$ , where, for any  $A \subseteq [0, 1]$ ,  $\tilde{P}_{m,*}(A) = \sup_{B \in \mathcal{O}_m, B \subseteq A} P_m(B)$ .
- (2) For any  $A \subseteq [0, 1]$ ,  $\underline{E}_m(A) = \tilde{P}_{m,*}(A)$ .
- (3)  $\underline{E}_m(h) = (C) \int h d\tilde{P}_{m,*} \forall h \in \mathcal{L}([0, 1])$ .

### 3.5. The natural extension of lower and upper distribution functions

In the rest of this paper, we intend to show that  $\underline{E}_m$  is the natural extension of the lower and upper distribution functions  $\underline{F}_m$  and  $\overline{F}_m$ . But before we can do that, we must make a small digression, in order to explain exactly what we mean by the phrase “natural extension of lower and upper distribution functions” in a general context. This is the subject of the present section. In the next section, we take up the thread of the moment problem again.



### 3.5.1. A precise distribution function

Let us begin with the simplest problem. We call any non-decreasing function  $F: [0, 1] \rightarrow [0, 1]$  that satisfies the normalisation condition  $F(1) = 1$  a *distribution function* on  $[0, 1]$ . The interpretation of such a distribution function is as follows: for any  $x \in [0, 1]$ , the (lower and upper) probability  $P_F([0, x])$  of  $[0, x]$  is equal to  $F(x)$ . Consequently, the probability  $P_F((x, 1])$  of  $(x, 1]$  is equal to  $1 - F(x)$ . In other words, specifying a distribution function  $F$  is tantamount to specifying a set function  $P_F$  on the set of events  $\mathcal{H} := \{[0, x] : x \in [0, 1]\} \cup \{(x, 1] : x \in [0, 1]\}$ , and it is easy to check that this  $P_F$  is a linear prevision. It can be uniquely extended to a linear prevision on the lattice  $\mathcal{Q}$  generated by  $\mathcal{H}$ , where all elements have the form

$$[0, x_1] \cup (x_2, x_3] \cup \cdots \cup (x_{2n}, 1] \quad \text{or} \quad (x_2, x_3] \cup \cdots \cup (x_{2n}, 1]$$

where  $0 \leq x_1 \leq x_2 \leq x_3 \leq \cdots \leq x_{2n} \leq 1$ .

The natural extension  $\underline{E}_F$  of  $P_F$  is the lower envelope of the set  $\mathcal{M}(P_F)$  of all linear previsions  $Q$  for which  $Q([0, x]) = F(x)$ ,  $x \in [0, 1]$ . For any gamble  $h$  on  $[0, 1]$ ,  $[\underline{E}_F(h), \overline{E}_F(h)]$  is the range of the value  $Q(h)$  for all such linear previsions  $Q$ .

Using the results in [1, 2], we see that  $\underline{E}_F$  is a completely monotone coherent lower prevision, that the restriction of  $\underline{E}_F$  to events is the inner set function  $P_{F,*}$  of  $P_F$ , and that for all gambles  $h$  on  $[0, 1]$ ,  $\underline{E}_F(h) = (C) \int h d\underline{E}_F$ . If  $\mathcal{L}_F := \{h \in \mathcal{L}([0, 1]) : \underline{E}_F(h) = \overline{E}_F(h)\}$  is the set of all  $F$ -integrable gambles, then we also know that  $\mathcal{L}_F$  is a uniformly closed linear lattice containing all constant gambles, and that a gamble  $h$  is  $F$ -integrable if and only if its cut sets  $\{h \geq t\}$  are  $F$ -integrable for all but a countable number of  $t$ .

Interestingly, it can be checked that *any distribution function  $F$  produces precise moments*, i.e., the polynomials  $p^k$  are  $F$ -integrable for any distribution function  $F$ .

### 3.5.2. Lower and upper distribution functions

Suppose now that we have two maps  $\underline{F}, \overline{F}: [0, 1] \rightarrow [0, 1]$ , which we interpret as a lower and an upper distribution function, respectively. This means that  $\underline{F}$  and  $\overline{F}$  determine a lower probability  $\underline{P}_{\underline{F}, \overline{F}}$  on the set  $\mathcal{H}$  by  $\underline{P}_{\underline{F}, \overline{F}}([0, x]) = \underline{F}(x)$  and  $\underline{P}_{\underline{F}, \overline{F}}((x, 1]) = 1 - \overline{F}(x)$  for all  $x \in [0, 1]$ . Walley [8, Section 4.6.6] has shown that  $\underline{P}_{\underline{F}, \overline{F}}$  is a coherent lower prevision if and only if  $\underline{F} \leq \overline{F}$  and both  $\underline{F}$  and  $\overline{F}$  are distribution functions, i.e., non-decreasing and normalised.

We shall assume in what follows that these conditions are satisfied.

The natural extension  $\underline{E}_{\underline{F}, \overline{F}}$  of the coherent lower probability  $\underline{P}_{\underline{F}, \overline{F}}$  to all gambles is the smallest coherent lower probability that coincides with  $\underline{P}_{\underline{F}, \overline{F}}$  on  $\mathcal{H}$ , or in other words, that has lower and upper distribution functions  $\underline{F}$  and  $\overline{F}$ . Denote by  $\mathcal{M}(\underline{F}, \overline{F})$  the set of all distribution functions (non-decreasing and normalised) on  $[0, 1]$  that lie between  $\underline{F}$  and  $\overline{F}$ .

**Theorem 3.4.**  $\underline{E}_{\underline{F}, \overline{F}}(h) = \inf\{\underline{E}_F(h) : F \in \mathcal{M}(\underline{F}, \overline{F})\}$  for all  $h \in \mathcal{L}([0, 1])$ .

### 3.6. The information given by the lower and the upper distribution functions

Let us now go back to the Hausdorff moment problem. As we have seen, if we consider a sequence  $m$  of moments  $m_k$ ,  $k \geq 0$  that satisfies Hausdorff's moment condition, we can consider the lower and upper envelopes  $\underline{E}_m$  and  $\overline{E}_m$  of all linear previsions with those moments. These lower and upper envelopes induce the lower and upper distribution functions  $\underline{F}_m$  and  $\overline{F}_m$ . We now proceed to show that these two functions already capture all the information that is present in the moments.

**Theorem 3.5.** *Consider a moment sequence  $m$  satisfying the Hausdorff moment condition.*

- (1) For all  $F$  in  $\mathcal{M}(\underline{F}_m, \overline{F}_m)$ , the restriction of  $\underline{E}_F$  to  $\mathcal{O}_m$  is equal to  $\tilde{P}_m$ .  
Hence,  $\underline{E}_F \geq \underline{E}_m$  and  $\mathcal{L}_m \subseteq \mathcal{L}_F$ .
- (2)  $\underline{E}_m = \underline{E}_{\underline{F}_m, \overline{F}_m}$ .

**Corollary 3.1.** *The following statements are equivalent.*

- (1)  $\underline{F}_m = \overline{F}_m$ ;
- (2)  $\underline{F}_m$ ,  $\overline{F}_m$  and  $F_m^\sigma$  are continuous on  $[0, 1]$ ;
- (3)  $\underline{E}_m = \underline{E}_F$  for some  $F \in \mathcal{M}(\underline{F}_m, \overline{F}_m)$ ;
- (4)  $\underline{E}_m = \underline{E}_F$  for all  $F \in \mathcal{M}(\underline{F}_m, \overline{F}_m)$ ;
- (5)  $\underline{E}_m = (RS)\int \cdot dF$  for all  $F \in \mathcal{M}(\underline{F}_m, \overline{F}_m)$ , where the integral is a lower Riemann-Stieltjes integral.

We can also prove the following:

**Theorem 3.6.** *Consider a moment sequence  $m$  satisfying the Hausdorff moment condition. Then for any  $F \in \mathcal{M}(\underline{F}_m, \overline{F}_m)$  such that  $F(0) = 0$  and any gamble  $h$  on  $[0, 1]$ ,  $\underline{E}_m(h) \leq (RS)\int h dF \leq (RS)\int h dF \leq \overline{E}_m(h)$ .*

Hence, if a gamble  $h$  is  $m$ -integrable, then the Riemann–Stieltjes integral of  $h$  with respect to  $F$  exists for any  $F \in \mathcal{M}(\underline{F}_m, \overline{F}_m)$ , and they all agree with  $\underline{E}_m(h) = \overline{E}_m(h)$ . It may happen nonetheless that a gamble  $h$  is Riemann–Stieltjes integrable with respect to some  $F \in \mathcal{M}(\underline{F}_m, \overline{F}_m)$  but not with respect to all of them.

There is a final result that summarises much of what has been said before in a concise manner. For this, note that any distribution function  $F$  in  $\mathcal{M}(\underline{F}_m, \overline{F}_m)$  can be written as a convex mixture  $F = \mu_m F_b + (1 - \mu_m) F_c$  of a continuous distribution function  $F_c$  and a ‘pure break function’  $F_b$ , which is a uniformly and absolutely convergent sum (convex mixture) of simple break functions. Explicitly, we have for all  $x \in [0, 1]$

$$\mu_m F_b(x) := F(x) - \underline{F}_m(x-) + \sum_{d \in \mathcal{D}_m, d < x} [\overline{F}_m(d) - \underline{F}_m(d-)]$$

and

$$(1 - \mu_m) F_c(x) := F(x) - \mu_m F_b(x) = \underline{F}_m(x) - \sum_{d \in \mathcal{D}_m, d < x} [\overline{F}_m(d) - \underline{F}_m(d-)],$$

where  $0 \leq \mu_m = \sum_{d \in \mathcal{D}_m} [\overline{F}_m(d) - \underline{F}_m(d-)] \leq 1$ , so the continuous part  $F_c$  is the same for all distributions  $F$  in  $\mathcal{M}(\underline{F}_m, \overline{F}_m)$ , and the pure break parts are identical in all the continuity points of  $\underline{F}_m$  and  $\overline{F}_m$ , and differ only by the values  $F(d)$  they assume in the countably many discontinuity points  $d \in \mathcal{D}_m$ . Define, for  $d \in [0, 1]$ , the functional  $\underline{osc}_d$  by  $\underline{osc}_d(h) := \sup_{z \in B_{\text{open}}} \inf_{z \in B} h(z)$  for all gambles  $h$  on  $[0, 1]$ .

**Theorem 3.7.** *Consider a moment sequence  $m$  satisfying the Hausdorff moment condition, and  $h \in \mathcal{L}([0, 1])$ . Then  $\underline{E}_m(h) = (1 - \mu_m) \underline{E}_{F_c}(h) + \sum_{d \in \mathcal{D}_m} [\overline{F}_m(d) - \underline{F}_m(d-)] \underline{osc}_d(h)$ , and the following are equivalent:*

- (1)  $h$  is  $m$ -integrable.
- (2)  $h$  is continuous in all the discontinuity points  $d \in \mathcal{D}_m$ , and Riemann–Stieltjes-integrable with respect to  $F_c$  if  $\mu_m < 1$ .

### 3.7. Conclusions

We have proven that the existence of a (finitely additive) probability charge with a given sequence of moments is equivalent to the existence of a ( $\sigma$ -additive) probability measure with these moments. Perhaps more surprisingly, and contrary to what happens in the case of  $\sigma$ -additive probabilities, a sequence of moments does not determine a unique probability charge.

If we consider the set of all probability charges with the given moment sequence, we can also induce lower and an upper distribution functions. We have proven that these two functions (actually any of the two) capture all the information given by the moments, i.e., we can deduce the lower and upper natural extensions from them. Moreover, we can gain more insight in the problem by considering these functions. First, the (unique)  $\sigma$ -additive probability measure with these moments induces the greatest distribution function of all the compatible probability charges; secondly, if there exists a continuous distribution function with these moments, then the results in Section 3.6 imply that *all* the distribution functions with those moments coincide and are, therefore, continuous. In particular, we deduce that the moments of a continuous distribution function can never be induced by a different, and discrete, distribution function on  $[0, 1]$ . Note however that even this does not imply the unicity of probability charges producing them! This is a consequence of Remark 3.1.

## Acknowledgements

Research supported by the projects MTM2004-01269, TSI2004-06801-C04-01 and by the research grant G.0139.01 of the Flemish Fund for Scientific Research (FWO). Erik Quaeghebeur's research is financed by a Ph.D. grant of the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT Vlaanderen).

## References

- [1] G. de Cooman, M. C. M. Troffaes, and E. Miranda.  $n$ -Monotone exact functionals. *Journal of Mathematical Analysis and Applications*, 347(1): 143–156, 2008.
- [2] G. de Cooman, M. C. M. Troffaes, and E. Miranda.  $n$ -Monotone lower previsions and lower integrals. In F. Cozman, R. Nau and T. Seidenfeld, editors, *Proceedings of ISIPTA'2005*, pages 145–154. 2005.
- [3] W. Feller. *An Introduction to Probability Theory and Its Applications*, volume II. John Wiley and Sons, New York, 1971.
- [4] F. Hausdorff. Summationmethoden und Momentfolgen I. *Mathematische Zeitschrift*, 9:74–109, 1921.
- [5] F. Hausdorff. Summationmethoden und Momentfolgen II. *Mathematische Zeitschrift*, 9:280–299, 1921.
- [6] E. Miranda, G. de Cooman and E. Quaeghebeur. The Hausdorff moment problem under finite additivity. *Journal of Theoretical Probability*, 20(3): 663–693, 2007.

- [7] E. Miranda, G. de Cooman and E. Quaeghebeur. Finitely additive extensions of distribution functions and moment sequences: the coherent lower prevision approach. 2007. *International Journal of Approximate Reasoning*, 48(1): 132–155, 2008.
- [8] P. Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1991.