On the moments and distribution of discrete Choquet integrals from continuous distributions

Ivan Kojadinovic [a], Jean-Luc Marichal [b]

[a] *Department of Statistics, The University of Auckland, Private Bag 92019, Auckland 1142, New Zealand.*

[b] *Mathematics Research Unit, University of Luxembourg, 162A, avenue de la Faïencerie, L-1511 Luxembourg, G.D. Luxembourg.*

Abstract

We study the moments and the distribution of the discrete Choquet integral when regarded as a real function of a random sample drawn from a continuous distribution. Since the discrete Choquet integral includes weighted arithmetic means, ordered weighted averaging functions, and lattice polynomial functions as particular cases, our results encompass the corresponding results for these aggregation functions. After detailing the results obtained in [1] in the uniform case, we present results for the standard exponential case, show how approximations of the moments can be obtained for other continuous distributions such as the standard normal, and elaborate on the asymptotic distribution of the Choquet integral. The results presented in this work can be used to improve the interpretation of discrete Choquet integrals when employed as aggregation functions.

Key words: Discrete Choquet integral; Lovász extension; order statistic; B-Spline; divided difference; asymptotic distribution.
1991 MSC: 60A10, 65C50 (Primary) 28B15, 62G30, 65D07 (Secondary).

1 Introduction

Aggregation functions are of central importance in many fields such as statistics, information fusion, risk analysis, or decision theory. In this paper, the primary object of interest is a natural extension of the weighted arithmetic mean

Email addresses: `ivan@stat.auckland.ac.nz` (Ivan Kojadinovic), `jean-luc.marichal@uni.lu` (Jean-Luc Marichal).

known as the (discrete) *Choquet integral* [2–4]. Also known in discrete mathematics as the *Lovász extension* of pseudo-Boolean functions [5], the Choquet integral is a very flexible aggregation function that includes weighted arithmetic means, ordered weighted averaging functions [6], and lattice polynomial functions as special cases [7,1].

Although the Choquet integral has been extensively employed as an aggregation function (see e.g. [8] for an overview), its moments and its distribution seem to have never been thoroughly studied from a theoretical perspective. The aim of this work is to attempt to fill this gap in the case when the Choquet integral is regarded as a real function of a random sample drawn from a continuous distribution.

The starting point of our study is a natural distributional relationship between linear combinations of order statistics and the Choquet integral, which merely results from the piecewise linear decomposition of the latter. As a consequence, exact formulations of the moments and the distribution of the Choquet integral can be provided whenever exact formulations are known for linear combinations of order statistics. Likewise, approximation and asymptotic results can be provided whenever available for linear combinations of order statistics.

The paper is organized as follows. In the second section, we recall the definition of the discrete Choquet integral. The third section is devoted to the expression of the distribution (resp. the moments) of the Choquet integral in terms of the distribution (resp. the moments) of linear combinations of order statistics. The case of standard uniform input variables is treated in Section 4. More precisely, the results obtained in [1] are detailed, and algorithms for computing the probability density function (p.d.f.) and the cumulative distribution function (c.d.f.) of the Choquet integral are provided. The fifth section deals with the standard exponential case, while the sixth one shows how approximations of moments can be obtained for other continuous distribution such as the standard normal. In the last section, we discuss conditions under which the asymptotic distribution of the Choquet integral is a mixture of normals.

The results obtained in this work have numerous applications. The most immediate ones are related to the interpretation of the Choquet integral when seen as an aggregation function. In multicriteria decision aiding in particular, the presented results can be used to generalize the behavioral indices studied e.g. in [9,10]. In classifier fusion, they can enable a theoretical study of the so-called *fuzzy approach* to classifier combination (see e.g. [11]) in the spirit of that done in [12].

Note that most of the methods and algorithms discussed in this work have been implemented in the R package `kappalab` [13] available on the Comprehensive

R Archive Network (`http://CRAN.R-project.org`).

2 The discrete Choquet integral

Define $N := \{1, \dots, n\}$ as a set of attributes, criteria, or players, and denote by $\mathfrak{S}_n$ the set of permutations on N. A set function $\nu : 2^N \to [0,1]$ is said to be a *game* on N if it satisfies $\nu(\varnothing) = 0$.

Definition 1 *The discrete Choquet integral of $\boldsymbol{x} \in \mathbb{R}^n$ w.r.t. a game ν on N is defined by*

$$C_\nu(\boldsymbol{x}) := \sum_{i=1}^{n} p_i^{\nu,\sigma}\, x_{\sigma(i)}\,,$$

where $\sigma \in \mathfrak{S}_n$ is such that $x_{\sigma(1)} \geqslant \cdots \geqslant x_{\sigma(n)}$, where

$$p_i^{\nu,\sigma} := \nu_i^\sigma - \nu_{i-1}^\sigma, \qquad \forall\, i \in N,$$

and where $\nu_i^\sigma := \nu\Big(\{\sigma(1), \dots, \sigma(i)\}\Big)$ for any $i = 0, \dots, n$. In particular, $\nu_0^\sigma := 0$.

Note that the permutation σ in the defintion of the Choquet integral of $\mathbf{x}$ is traditionally taken such that $x_{\sigma(1)} \leqslant \cdots \leqslant x_{\sigma(n)}$. The reason for not adopting this convention in this work is due to the fact that it would have led to much more complicated expressions of the results to be presented in Section 4.

From the above definition, we see that the Choquet integral is a piecewise linear function that coincides with a weighted sum on each n-dimensional polyhedron

$$R_\sigma := \{\mathbf{x} \in \mathbb{R}^n \mid x_{\sigma(1)} \geqslant \cdots \geqslant x_{\sigma(n)}\}, \qquad \sigma \in \mathfrak{S}_n, \tag{1}$$

whose union covers $\mathbb{R}^n$. It can additionally be immediately verified that it is a continuous function.

When defined as above, the Choquet integral coincides with the *Lovász extension* [5] of the unique pseudo-Boolean function that can be associated with ν [14] and can be alternatively regarded as a linear combination of lattice polynomial functions (see e.g. [1]).

In aggregation theory, it is natural to additionally require that the game ν is monotone w.r.t. inclusion and satisfies $\nu(N) = 1$, in which case it is called a *capacity* [2]. The resulting aggregation function C_ν is then nondecreasing in each variable and coincides with a weighted arithmetic mean on each of the n-dimensional polyhedra defined by (1). Furthermore, in this case, for any

$T \subseteq N$, the coefficient $\nu(T)$ can be naturally interpreted as the *weight* or the *importance* of the subset T of attributes [4].

The Choquet integral w.r.t. a capacity satisfies very appealing properties for aggregation. For instance, it is comprised between the minimum and the maximum, stable under the same transformations of interval scales in the sense of the theory of measurement, and coincides with a weighted arithmetic mean whenever the capacity is additive. An axiomatic characterization is provided in [4]. Moreover, the Choquet integral w.r.t. a capacity includes weighted arithmetic means, ordered weighted averaging functions [6], and lattice polynomial functions as particular cases [7,1].

3 Distributional relationships with linear combinations of order statistics

In the present section, we investigate the moments and the distribution of the Choquet integral when considered as a function of n continuous i.i.d. random variables. Our main theoretical results, stated in the following proposition and its corollary, yield expressions of the moments and the distribution of the Choquet integral in terms of the moments and the distribution of linear combinations of order statistics.

Let $X_1, \ldots, X_n$ be a random sample drawn from a continuous c.d.f. $F : \mathbb{R} \to \mathbb{R}$ with associated p.d.f. $f : \mathbb{R} \to \mathbb{R}$, and let $X_{1:n} \leqslant \cdots \leqslant X_{n:n}$ denote the corresponding order statistics. Furthermore, let

$$\begin{aligned} Y_\nu &:= C_\nu(X_1, \ldots, X_n), \\ Y_\nu^\sigma &:= \sum_{i=1}^n p_i^{\nu,\sigma} X_{n-i+1:n}, \qquad \sigma \in \mathfrak{S}_n. \end{aligned}$$

Let also $F_\nu(y)$ and $F_\nu^\sigma(y)$ be the c.d.f.s of Y_ν and Y_ν^σ, respectively. Finally, let $h : \mathbb{R} \to \mathbb{R}$ be any measurable function.

Proposition 2 *For any game ν on N, we have*

$$\mathbf{E}[h(Y_\nu)] = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \mathbf{E}[h(Y_\nu^\sigma)].$$

Proof. By definition, we have

$$\begin{aligned}\mathbf{E}[h(Y_\nu)] &= \int_{\mathbb{R}^n} h(C_\nu(x_1,\ldots,x_n)) \prod_{i=1}^n f(x_i)\,\mathrm{d}x_i \\ &= \sum_{\sigma\in\mathfrak{S}_n} \int_{R_\sigma} h\Big(\sum_{i=1}^n p_i^{\nu,\sigma}\, x_{\sigma(i)}\Big) \prod_{i=1}^n f(x_i)\,\mathrm{d}x_i.\end{aligned}$$

Using the well-known fact (see e.g. [15, §2.2]) that the joint p.d.f. of $X_{1:n} \leqslant \cdots \leqslant X_{n:n}$ is

$$n!\prod_{i=1}^n f(x_i), \qquad x_1 \leqslant \cdots \leqslant x_n,$$

we obtain

$$\mathbf{E}[h(Y_\nu)] = \frac{1}{n!} \sum_{\sigma\in\mathfrak{S}_n} \mathbf{E}\Big[h\Big(\sum_{i=1}^n p_i^{\nu,\sigma}\, X_{n-i+1:n}\Big)\Big],$$

which completes the proof. □

Before going through the main corollary, recall that the *plus* (resp. *minus*) *truncated power function* x_+^n (resp. x_-^n) is defined to be x^n if $x > 0$ (resp. $x < 0$) and zero otherwise.

Corollary 3 *For any game ν on N, we have*

$$F_\nu(y) = \frac{1}{n!} \sum_{\sigma\in\mathfrak{S}_n} F_\nu^\sigma(y).$$

Proof. Define $h_y(x) := (x-y)_-^0$. Then, from Proposition 2, for any $y \in \mathbb{R}$, we have

$$F_\nu(y) = \mathbf{E}[h_y(Y_\nu)] = \frac{1}{n!} \sum_{\sigma\in\mathfrak{S}_n} \mathbf{E}[h_y(Y_\nu^\sigma)] = \frac{1}{n!} \sum_{\sigma\in\mathfrak{S}_n} F_\nu^\sigma(y).$$

□

The results stated in Proposition 2 and Corollary 3 are not very surprising. From Definition 1, it is clear that the Choquet integral is a linear combination of order statistics whose coefficients depend on the ordering of the arguments. The different possible orderings merely lead to a division of the integration domain $\mathbb{R}^n$ into the subdomains R_σ ($\sigma \in \mathfrak{S}_n$) defined in (1), and the difficult part still lies in the evaluation of the moments and the distribution of linear combinations of order statistics.

The relationship for the raw moments is obtained by considering the special case $h(x) = x^r$, which may still lead to tedious computations. From Proposition 2, we obtain

$$\mathbf{E}[Y_\nu] = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \sum_{k=1}^{n} p_k^{\nu,\sigma}\, \mathbf{E}[X_{n-k+1:n}],$$

and more generally,

$$\mathbf{E}[Y_\nu^r] = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \sum_{k_1,\ldots,k_r=1}^{n} \Big(\prod_{i=1}^{r} p_{k_i}^{\nu,\sigma} \Big)\, \mathbf{E}\Big[\prod_{i=1}^{r} X_{n-k_i+1:n} \Big].$$

Unfortunately, this latter formula involves a huge number of terms, namely $n!\, n^r$. The following result (see [1, Prop. 3] for the uniform case) yields the rth raw moment as a sum of $(r+1)^n$ terms, each of which is a product of coefficients $\nu(T)$.

Proposition 4 *For any integer $r \geqslant 1$ and any game ν on N, setting $T_{r+1} := N$ and $X_{0:n} := 0$, we have*

$$\mathbf{E}[Y_\nu^r] = \sum_{T_1 \subseteq \cdots \subseteq T_r \subseteq N} \frac{r!}{[T]_0! \cdots [T]_n!} \Big(\prod_{i=1}^{r} \frac{\nu(T_i)}{\binom{|T_{i+1}|}{|T_i|}} \Big)\, \mathbf{E}\Big[\prod_{i=1}^{r} (X_{n-|T_i|+1:n} - X_{n-|T_i|:n}) \Big],$$

where $[T]_j$ represents the number of "j" among $|T_1|, \ldots, |T_r|$.

Proof. Fix $\sigma \in \mathfrak{S}_n$. Rewriting Y_ν^σ as

$$Y_\nu^\sigma = \sum_{i=0}^{n} \nu_i^\sigma (X_{n-i+1:n} - X_{n-i:n}),$$

and then using the multinomial theorem, we obtain

$$\begin{aligned}
(Y_\nu^\sigma)^r &= \sum_{\substack{r_1,\ldots,r_n \geqslant 0 \\ r_1+\cdots+r_n = r}} \frac{r!}{r_1! \cdots r_n!} \prod_{i=0}^{n} (\nu_i^\sigma)^{r_i} \, (X_{n-i+1:n} - X_{n-i:n})^{r_i} \\
&= \sum_{0 \leqslant i_1 \leqslant \cdots \leqslant i_r \leqslant n} \frac{r!}{[i]_0! \cdots [i]_n!} \prod_{k=1}^{r} \nu_{i_k}^\sigma \, (X_{n-i_k+1:n} - X_{n-i_k:n}),
\end{aligned}$$

where $[i]_j$ represents the number of "j" among $i_1, \ldots, i_r$. Now, using Proposition 2 with $h(x) = x^r$, we immediately obtain

$$\mathbf{E}[Y_\nu^r] = \sum_{0 \leqslant i_1 \leqslant \cdots \leqslant i_r \leqslant n} \frac{r!}{[i]_0! \cdots [i]_n!}\, \mathbf{E}\Big[\prod_{k=1}^{r} (X_{n-i_k+1:n} - X_{n-i_k:n}) \Big]\, \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \prod_{k=1}^{r} \nu_{i_k}^\sigma .$$

The final result then follows from the identity (see the proof of [1, Prop. 3])

$$\frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n}\prod_{k=1}^{r}\nu_{i_k}^{\sigma} = \sum_{\substack{T_1\subseteq\cdots\subseteq T_r\subseteq N\\ |T_1|=i_1,\ldots,|T_r|=i_r}}\prod_{i=1}^{r}\frac{\nu(T_i)}{\binom{|T_{i+1}|}{|T_i|}}.$$

□

For example, the first two raw moments are

$$\mathbf{E}[Y_\nu] = \sum_{T\subseteq N}\frac{\nu(T)}{\binom{n}{|T|}}\,\mathbf{E}[X_{n-|T|+1:n} - X_{n-|T|:n}] \tag{2}$$

and

$$\mathbf{E}[Y_\nu^2] = \sum_{T_1\subseteq T_2\subseteq N}\frac{2}{[T]_0!\cdots[T]_n!}\,\frac{\nu(T_1)\nu(T_2)}{\binom{|T_2|}{|T_1|}\binom{n}{|T_2|}}\,\mathbf{E}\Big[(X_{n-|T_1|+1:n}-X_{n-|T_1|:n})(X_{n-|T_2|+1:n}-X_{n-|T_2|:n})\Big],$$

that is,

$$\begin{aligned}\mathbf{E}[Y_\nu^2] = \sum_{T_1\subsetneq T_2\subseteq N} 2\,\frac{\nu(T_1)\nu(T_2)}{\binom{|T_2|}{|T_1|}\binom{n}{|T_2|}}\,\mathbf{E}\Big[(X_{n-|T_1|+1:n}-X_{n-|T_1|:n})(X_{n-|T_2|+1:n}-X_{n-|T_2|:n})\Big]\\ + \sum_{T\subseteq N}\frac{\nu(T)^2}{\binom{n}{|T|}}\,\mathbf{E}\Big[X_{n-|T|+1:n} - X_{n-|T|:n}\Big]^2.\end{aligned} \tag{3}$$

4 The uniform case

In this section, we focus on the moments and the distribution of Y_ν when the random sample $X_1,\ldots,X_n$ is drawn from the standard uniform distribution. To emphasize this last point, as classically done, we shall denote the random sample as $U_1,\ldots,U_n$ and the corresponding order statistics by $U_{1:n}\leqslant\cdots\leqslant U_{n:n}$.

Before detailing the results obtained in [1] and providing algorithms for computing the p.d.f. and the c.d.f. of the Choquet integral, we recall some basic material related to divided differences (see e.g. [16–18] for further details).

4.1 Divided differences

Let $\mathcal{A}^{(n)}$ be the set of $n-1$ times differentiable one-place functions g such that $g^{(n-1)}$ is absolutely continuous. The nth *divided difference* of a function

$g \in \mathcal{A}^{(n)}$ is the symmetric function of $n+1$ arguments defined inductively by $\Delta[g : a_0] := g(a_0)$ and

$$\Delta[g : a_0, \ldots, a_n] := \begin{cases} \dfrac{\Delta[g : a_1, \ldots, a_n] - \Delta[g : a_0, \ldots, a_{n-1}]}{a_n - a_0}, & \text{if } a_0 \neq a_n, \\ \dfrac{\partial}{\partial a_0} \Delta[g : a_0, \ldots, a_{n-1}], & \text{if } a_0 = a_n. \end{cases}$$

The *Peano representation* of the divided differences is given by

$$\Delta[g : a_0, \ldots, a_n] = \frac{1}{n!} \int_{\mathbb{R}} g^{(n)}(t)\, M(t \mid a_0, \ldots, a_n)\, \mathrm{d}t,$$

where $M(t \mid a_0, \ldots, a_n)$ is the *B-spline* of order n, with knots $\{a_0, \ldots, a_n\}$, defined as

$$M(t \mid a_0, \ldots, a_n) := n\, \Delta[(\boldsymbol{\cdot} - t)_+^{n-1} : a_0, \ldots, a_n]. \tag{4}$$

We also recall the *Hermite-Genocchi formula*: For any function $g \in \mathcal{A}^{(n)}$, we have

$$\Delta[g : a_0, \ldots, a_n] = \int_{R_{id} \cap [0,1]^n} g^{(n)}\Big[a_0 + \sum_{i=1}^{n} (a_i - a_{i-1}) x_i\Big]\, \mathrm{d}x, \tag{5}$$

where R_{id} is the region defined in (1) when σ is the identity permutation.

For distinct arguments $a_0, \ldots, a_n$, we also have the following formula, which can be verified by induction,

$$\Delta[g : a_0, \ldots, a_n] = \sum_{i=0}^{n} \frac{g(a_i)}{\prod_{j \neq i} (a_i - a_j)}. \tag{6}$$

4.2 Moments and distribution

Let $g \in \mathcal{A}^{(n)}$. From (5), we immediately have that

$$\mathbf{E}\Big[g^{(n)}\Big(\sum_{i=1}^{n} p_i^{\nu,\sigma}\, U_{n-i+1:n}\Big)\Big] = n!\, \Delta[g : \nu_0^{\sigma}, \ldots, \nu_n^{\sigma}] \tag{7}$$

since the joint p.d.f. of $U_{1:n} \leqslant \cdots \leqslant U_{n:n}$ is $1/n!$ on $R_{id} \cap [0,1]^n$ and zero elsewhere.

Now, combining (7) with Proposition 2, we obtain

$$\mathbf{E}[g^{(n)}(Y_\nu)] = \sum_{\sigma \in \mathfrak{S}_n} \Delta[g : \nu_0^{\sigma}, \ldots, \nu_n^{\sigma}]. \tag{8}$$

Eq. (8) provides the expectation $\mathbf{E}[g^{(n)}(Y_\nu)]$ in terms of the divided differences of g with arguments $\nu_0^\sigma, \ldots, \nu_n^\sigma$ ($\sigma \in \mathfrak{S}_n$). An explicit formula can be obtained by (6) whenever the arguments are distinct for every $\sigma \in \mathfrak{S}_n$.

Clearly, the special cases

$$g(x) = \frac{r!}{(n+r)!}\, x^{n+r}, \quad \frac{r!}{(n+r)!}\, [x - \mathbf{E}(Y_\nu)]^{n+r}, \text{ and } \frac{e^{tx}}{t^n}$$

give, respectively, the raw moments, the central moments, and the moment-generating function of Y_ν. As far as the raw moments are concerned, we have the following result [1, Prop. 3], which is a special case of Proposition 4.

Proposition 5 *For any integer $r \geqslant 1$ and any game ν on N, setting $T_{r+1} := N$, we have*

$$\mathbf{E}[Y_\nu^r] = \frac{1}{\binom{n+r}{r}} \sum_{T_1 \subseteq \cdots \subseteq T_r \subseteq N} \prod_{i=1}^{r} \frac{\nu(T_i)}{\binom{|T_{i+1}|}{|T_i|}}.$$

Proposition 5 provides an explicit expression for the rth raw moment of Y_ν as a sum of $(r+1)^n$ terms. For instance, the first two moments are

$$\mathbf{E}[Y_\nu] = \frac{1}{n+1} \sum_{T \subseteq N} \frac{\nu(T)}{\binom{n}{|T|}}, \tag{9}$$

$$\mathbf{E}[Y_\nu^2] = \frac{2}{(n+1)(n+2)} \sum_{T_1 \subseteq T_2 \subseteq N} \frac{\nu(T_1)\nu(T_2)}{\binom{|T_2|}{|T_1|}\binom{n}{|T_2|}}. \tag{10}$$

By using (8) with $g(x) = \frac{1}{n!}(x-y)_-^n$, we also obtain the c.d.f. $F_\nu(y)$ of Y_ν [1].

Theorem 6 *There holds*

$$F_\nu(y) = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \Delta[(\cdot - y)_-^n : \nu_0^\sigma, \ldots, \nu_n^\sigma] = 1 - \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \Delta[(\cdot - y)_+^n : \nu_0^\sigma, \ldots, \nu_n^\sigma]. \tag{11}$$

It follows from (11) that the p.d.f. of Y_ν is simply given by

$$\begin{aligned} f_\nu(y) &= -\frac{1}{(n-1)!} \sum_{\sigma \in \mathfrak{S}_n} \Delta[(\cdot - y)_-^{n-1} : \nu_0^\sigma, \ldots, \nu_n^\sigma] \\ &= \frac{1}{(n-1)!} \sum_{\sigma \in \mathfrak{S}_n} \Delta[(\cdot - y)_+^{n-1} : \nu_0^\sigma, \ldots, \nu_n^\sigma], \end{aligned} \tag{12}$$

or, using the B-spline notation (4), by

$$f_\nu(y) = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} M(y \mid \nu_0^\sigma, \ldots, \nu_n^\sigma).$$

Remark:

(i) When the arguments $\nu_0^\sigma,\ldots,\nu_n^\sigma$ are distinct for every $\sigma\in\mathfrak{S}_n$, then combining (6) with (11) immediately yields the following explicit expressions

$$F_\nu(y)=\frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n}\sum_{i=0}^{n}\frac{(\nu_i^\sigma-y)_-^n}{\prod_{j\neq i}(\nu_i^\sigma-\nu_j^\sigma)}=1-\frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n}\sum_{i=0}^{n}\frac{(\nu_i^\sigma-y)_+^n}{\prod_{j\neq i}(\nu_i^\sigma-\nu_j^\sigma)}.$$

(ii) The case of linear combinations of order statistics, called *ordered weighted averaging operators* in aggregation theory (see e.g. [6]), is of particular interest. In this case, each ν_i^σ is independent of σ, so that we can write $\nu_i:=\nu_i^\sigma$. The main formulas then reduce to (see e.g. [19,20])

$$\begin{aligned}\mathbf{E}[g^{(n)}(Y_\nu)]&=n!\,\Delta[g:\nu_0,\ldots,\nu_n],\\ F_\nu(y)&=\Delta[(\cdot-y)_-^n:\nu_0,\ldots,\nu_n],\\ f_\nu(y)&=M(y\mid\nu_0,\ldots,\nu_n).\end{aligned}$$

Note also that the Hermite-Genocchi formula (5) provides nice geometric interpretations of $F_\nu(y)$ and $f_\nu(y)$ in terms of volumes of slices and sections of canonical simplices (see also [21,22]).

4.3 Algorithms

Both the functions F_ν and f_ν require the computation of divided differences of truncated power functions. On this issue, we recall a recurrence equation, due to de Boor [23] and rediscovered independently by Varsi [24] (see also [21]), which allows to compute $\Delta[(\cdot-y)_+^{n-1}:a_0,\ldots,a_n]$ in $O(n^2)$ operations.

Rename as $b_1,\ldots,b_r$ the elements a_i such that $a_i<y$ and as $c_1,\ldots,c_s$ the elements a_i such that $a_i\geqslant y$ so that $r+s=n+1$. Then, the unique solution of the recurrence equation

$$\alpha_{k,l}=\frac{(c_l-y)\alpha_{k-1,l}+(y-b_k)\alpha_{k,l-1}}{c_l-b_k},\qquad k\leqslant r,\ l\leqslant s,$$

with initial values $\alpha_{1,1}=(c_1-b_1)^{-1}$ and $\alpha_{0,l}=\alpha_{k,0}=0$ for all $l,k\geqslant 2$, is given by

$$\alpha_{k,l}:=\Delta[(\cdot-y)_+^{k+l-2}:b_1,\ldots,b_k,c_1,\ldots,c_l],\qquad k+l\geqslant 2.$$

In order to compute $\Delta[(\cdot-y)_+^{n-1}:a_0,\ldots,a_n]=\alpha_{r,s}$, it suffices therefore to compute the sequence $\alpha_{k,l}$ for $k+l\geqslant 2$, $k\leqslant r$, $l\leqslant s$, by means of two nested loops, one on k, the other on l. We detail this computation in Algorithm 1 (see also [21,24]).

Algorithm 1 Algorithm for the computation of $\Delta[(\cdot - y)_+^{n-1} : a_0, \ldots, a_n]$.

```
Require: n, a_0, ..., a_n, y
  S ← 0, R ← 0
  for i = 0, 1, ..., n do
    if x_i − y ⩾ 0 then
      S ← S + 1
      C_S ← x_i − y
    else
      R ← R + 1
      B_R ← x_i − y
    end if
  end for
  A_0 ← 0, A_1 ← 1/(C_1 − B_1) {Initialization of the unidimensional temporary
  array of size S + 1 necessary for the computation of the divided difference}
  for j = 2, ..., S do
    A_j ← −B_1 A_{j−1}/(C_j − B_1)
  end for
  for i = 2, ..., R do
    for j = 1, ..., S do
      A_j ← (C_j A_j − B_i A_{j−1})/(C_j − B_i)
    end for
  end for
  return A_R {Contains the value of Δ[(· − y)_+^{n−1} : a_0, ..., a_n].}
```

We can compute $\Delta[(\cdot - y)_-^n : a_0, \ldots, a_n]$ similarly. Indeed, the same recurrence equation applied to the initial values $\alpha_{0,l} = 0$ for all $l \geqslant 1$ and $\alpha_{k,0} = 1$ for all $k \geqslant 1$, produces the solution

$$\alpha_{k,l} := \Delta[(\cdot - y)_-^{k+l-1} : b_1, \ldots, b_k, c_1, \ldots, c_l], \qquad k + l \geqslant 1.$$

Example 1 *The Choquet integral is frequently used in multicriteria decision aiding, non-additive expected utility theory, or complexity analysis (see for instance [8] for an overview). For instance, when such an operator is used as an aggregation function in a given decision making problem, it is very informative for the decision maker to know its distribution. In that context, one of the most natural* a priori *p.d.f.s on* $[0,1]^n$ *is the standard uniform, which makes the results presented in this section of particular interest. Let* ν *be the capacity on* $N = \{1, 2, 3\}$ *defined by* $\nu(\{1\}) = 0.1$, $\nu(\{2\}) = 0.2$, $\nu(\{3\}) = 0.55$, $\nu(\{1, 2\}) = 0.7$, $\nu(\{1, 3\}) = 0.8$, $\nu(\{2, 3\}) = 0.6$, *and* $\nu(\{1, 2, 3\}) = 1$. *The p.d.f. of the Choquet integral w.r.t.* ν, *which can be computed through (12) and by means of Algorithm 1, is represented in Figure 1 (left) by the solid line. The dotted line represents the p.d.f. estimated by the kernel method from 10 000 randomly generated realizations of* U_1, U_2, U_3 *using the R statistical system [25]. The expectation and the standard deviation can also be calculated*

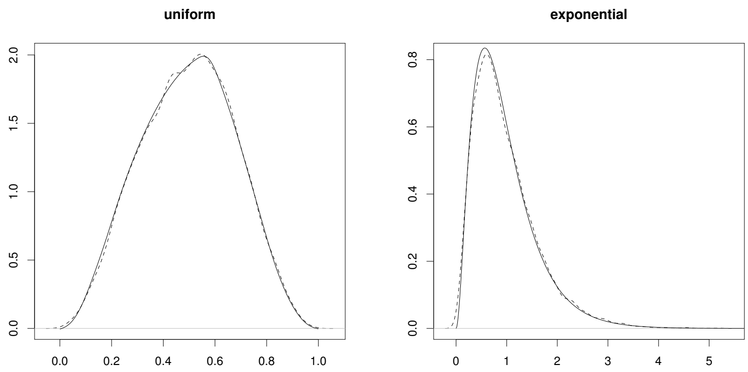

Fig. 1. P.d.f.s of discrete Choquet integral (solid lines) in the standard uniform and standard exponential cases. The dotted lines represent the corresponding p.d.f.s estimated by the kernel method from 10 000 randomly generated realizations.

through (9) and (10). We have

$$\mathbf{E}[Y_\nu] \approx 0.495 \quad \textit{and} \quad \sqrt{\mathbf{E}[Y_\nu^2] - \mathbf{E}[Y_\nu]^2} \approx 0.183.$$

The sample mean and the variance of the above mentioned 10 000 realizations of the Choquet integral are

$$\bar{y}_\nu \approx 0.497 \quad \textit{and} \quad s_{y_\nu} \approx 0.183.$$

5 The standard exponential case

In the standard exponential case, i.e., when $F(x) = 1 - e^{-x}$, $x \geqslant 0$, the exact distribution of the Choquet integral can be obtained if the numbers $\{\nu_i^\sigma\}_{i\in N, \sigma\in\mathfrak{S}_n}$ satisfy certain regularity conditions. The result is based on the following proposition (see [15, §6.5] and the references therein).

Proposition 7 *Let $a_1, \ldots, a_n \in \mathbb{R}$ and let $X_1, \ldots, X_n$ be a random sample drawn from the standard exponential distribution. For any $i \in N$, define*

$$c_i = \frac{1}{n-i+1} \sum_{j=i}^{n} a_j.$$

Then, if $c_i \neq c_k$ whenever $i \neq k$, and $c_i > 0$ for all $i \in N$, the p.d.f. of $T = \sum_{i=1}^n a_i X_{i:n}$ is given by

$$f_T(y) = \sum_{i=1}^{n} \frac{c_i^{n-2}}{\prod_{k\neq i}(c_i - c_k)} \exp\left(-\frac{y}{c_i}\right).$$

The p.d.f. $f_\nu(y)$ of the Choquet integral then results from Corollary 3 and Proposition 7.

Corollary 8 *Assume that, for any $\sigma \in \mathfrak{S}_n$, $\nu_i^\sigma/i \neq \nu_k^\sigma/k$ whenever $i \neq k$, and that $\nu_i^\sigma/i > 0$ for all $i \in N$. Then*

$$f_\nu(y) = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \sum_{i=1}^{n} \frac{(\nu_i^\sigma/i)^{n-2}}{\prod_{k \neq i}(\nu_i^\sigma/i - \nu_k^\sigma/k)} \exp\left(-\frac{y}{(\nu_i^\sigma/i)}\right). \tag{13}$$

Proof. The result is a direct consequence of Corollary 3, Proposition 7, and the fact that, for any $\sigma \in \mathfrak{S}_n$,

$$\frac{1}{n-i+1} \sum_{j=i}^{n} p_{n-j+1}^{\nu,\sigma} = \frac{\nu_{n-i+1}^\sigma}{n-i+1}, \qquad i \in N. \qquad \square$$

$\square$

The first two moments of the order statistics in the standard exponential case are given (see e.g. [15, p. 52]) by

$$\mathbf{E}[X_{i:n}] = \sum_{k=n-i+1}^{n} \frac{1}{k}, \tag{14}$$

and, if $i < j$,

$$\mathbf{E}[X_{i:n}X_{j:n}] - \mathbf{E}[X_{i:n}]\mathbf{E}[X_{j:n}] = \mathbf{E}[X_{i:n}^2] - \mathbf{E}[X_{i:n}]^2 = \sum_{k=n-i+1}^{n} \frac{1}{k^2}. \tag{15}$$

Used in combination with (2) and (3), these expressions enable us to obtain the first two raw moments of the Choquet integral.

Example 2 *Consider again the capacity given in Example 1 and assume now that X_1, X_2, X_3 is a random sample from the standard exponential distribution. The p.d.f. of the Choquet integral w.r.t. ν, which can be computed by means of (13), is represented in Figure 1 (right) by the solid line. The dotted line represents the p.d.f. estimated by the kernel method from 10 000 randomly generated realizations.*

Combining (14) and (15) with (2) and (3), we obtain the following values:

$$\mathbf{E}[Y_\nu] \approx 0.963 \quad \text{and} \quad \sqrt{\mathbf{E}[Y_\nu^2] - \mathbf{E}[Y_\nu]^2} \approx 0.624.$$

The sample mean and the variance of the above mentioned 10 000 realizations of the Choquet integral are

$$\bar{y}_\nu \approx 0.964 \quad \text{and} \quad s_{y_\nu} \approx 0.630.$$

6 Approximations of the moments

When F is neither the standard uniform, nor the standard exponential c.d.f., but F^{-1} and its derivatives can be easily computed, one can obtain approximations of the moments of order statistics, and therefore of those of the Choquet integral, using the approach initially proposed by David and Johnson [26].

Let $U_1, \dots, U_n$ be a random sample from the standard uniform distribution. The product moments of the corresponding order statistics are then given by the following formula:

$$\mathbf{E}\Big[\prod_{j=1}^{l} U_{i_j:n}^{m_j}\Big] = \frac{n!}{\big(n + \sum_{j=1}^{l} m_j\big)!} \prod_{j=1}^{l} \frac{(i_j + m_1 + \cdots + m_j - 1)!}{(i_j + m_1 + \cdots + m_{j-1} - 1)!}, \tag{16}$$

where $1 \leqslant i_1 < \cdots < i_l \leqslant n$. Now, it is well known that the c.d.f. of $X_{i:n}$ is given by

$$\Pr[X_{i:n} \leqslant x] = \sum_{j=i}^{n} \binom{n}{j} F^j(x)[1 - F(x)]^{n-j}.$$

It immediately follows that

$$\Pr[F^{-1}(U_{i:n}) \leqslant x] = \Pr[U_{i:n} \leqslant F(x)] = \Pr[X_{i:n} \leqslant x],$$

i.e., that $F^{-1}(U_{i:n})$ and $X_{i:n}$ are equal in distribution.

Starting from this distributional equality, David and Johnson [26] expanded $F^{-1}(U_{i:n})$ in a Taylor series around the point $\mathbf{E}[U_{i:n}] = i/(n+1)$ in order to obtain approximations of product moments of non-uniform order statistics. Setting $r_i := i/(n+1)$, $G := F^{-1}$, $G_i := G(r_i)$, $G_i^{(1)} := G^{(1)}(r_i)$, etc., we have

$$X_{i:n} = G_i + (U_{i:n} - r_i)G_i^{(1)} + \frac{1}{2}(U_{i:n} - r_i)^2 G_i^{(2)} + \frac{1}{6}(U_{i:n} - r_i)^3 G_i^{(3)} + \dots$$

Setting $s_i := 1 - r_i$, taking the expectation of the previous expression and using (16), the following approximation for the expectation of $X_{i:n}$ can be obtained to order $(n+2)^{-2}$:

$$\mathbf{E}[X_{i:n}] \approx G_i + \frac{r_i s_i}{2(n+2)} G_i^{(2)} + \frac{r_i s_i}{(n+2)^2} \left[\frac{1}{3}(s_i - r_i)G_i^{(3)} + \frac{1}{8} r_i s_i G_i^{(4)}\right]. \tag{17}$$

Similarly, for the first product moment, we have

$$\begin{aligned}\mathbf{E}[X_{i:n}X_{j:n}] \approx G_iG_j &+ \frac{r_is_j}{n+2}G_i^{(1)}G_j^{(1)} + \frac{r_is_i}{2(n+2)}G_jG_i^{(2)} + \frac{r_js_j}{2(n+2)}G_iG_j^{(2)} \\ &+ \frac{r_is_j}{(n+2)^2}\Big[(s_i-r_i)G_i^{(2)}G_j^{(1)} + (s_j-r_j)G_i^{(1)}G_j^{(2)} + \frac{1}{2}r_is_iG_i^{(3)}G_j^{(1)} \\ &+ \frac{1}{2}r_js_jG_i^{(1)}G_j^{(3)} + \frac{1}{2}r_is_jG_i^{(2)}G_j^{(2)}\Big] + \frac{r_ir_js_is_j}{4(n+2)^2}G_i^{(2)}G_j^{(2)} \\ &+ \frac{r_is_iG_j}{(n+2)^2}\Big[\frac{1}{8}r_is_iG_i^{(4)} + \frac{1}{3}(s_i-r_i)G_i^{(2)}\Big] \\ &+ \frac{r_js_jG_i}{(n+2)^2}\Big[\frac{1}{8}r_js_jG_j^{(4)} + \frac{1}{3}(s_j-r_j)G_j^{(2)}\Big]. \qquad (18)\end{aligned}$$

The accuracy of the above approximations is discussed in [15, §4.6]. Note that Childs and Balakrishnan [27] have recently proposed MAPLE routines facilitating the computations and permitting the inclusion of higher order terms.

As already mentioned, the previous expressions are useful only if $G := F^{-1}$ and its derivatives can be easily computed. This is the case for instance when F is the standard normal c.d.f. Indeed, there exist algorithms that enable an accurate computation of F^{-1} and it can be verified (see e.g. [15, p 85]) that $G^{(1)} = (f \circ G)^{-1}$,

$$G^{(2)} = \frac{G}{f^2 \circ G}, \qquad G^{(3)} = \frac{1+2G^2}{f^3 \circ G} \qquad \text{and} \qquad G^{(4)} = \frac{G(7+6G^2)}{f^4 \circ G},$$

where $f := F^{(1)}$.

From a practical perspective, in order to obtain a better accuracy for $\mathbf{E}[X_{i:n}]$ and $\mathbf{E}[X_{i:n}X_{j:n}]$ in the standard normal case, one can use the expressions obtained to order $(n+2)^{-3}$ in [26] and recalled in [27]. We do not reproduce these expressions here as they are very long. We provide however the expressions of $G^{(5)}$ and $G^{(6)}$ required for computing them:

$$G^{(5)} = \frac{7+G^2(46+24G^2)}{f^5 \circ G} \qquad \text{and} \qquad G^{(6)} = \frac{G(127+326G^2+96G^4)}{f^6 \circ G}.$$

Example 3 *Consider again the capacity given in Example 1 and assume now that the decision maker wants the standard normal as* a priori *p.d.f. Combining (17) and (18) with (2) and (3), we obtain the following approximate values:*

$$\mathbf{E}[Y_\nu] \approx -0.014 \quad \textit{and} \quad \sqrt{\mathbf{E}[Y_\nu^2] - \mathbf{E}[Y_\nu]^2} \approx 0.615.$$

For comparison, the sample mean and the variance of 10 000 independent realizations of the corresponding Choquet integral are

$$\bar{y}_\nu \approx -0.013 \quad \textit{and} \quad s_{y_\nu} \approx 0.620.$$

7 Asymptotic distribution of the Choquet integral

Conditions under which a linear combination of order statistics is asymptotically normal have been extensively studied in the statistical literature. A good synthesis on the subject is given in [15, §11.4]. Provided some regularity conditions are satisfied, typically on ν and F in the context under consideration, the existing theoretical results, combined with Proposition 2, practically imply that, for large n, Y_ν is approximately distributed as a mixture of $n!$ normals $N(\mathbf{E}[Y_\nu^\sigma], \mathbf{V}[Y_\nu^\sigma])$, $\sigma \in \mathfrak{S}_n$, each weighted by $\frac{1}{n!}$.

From a practical perspective, the most useful result seems to be that of Stigler [28]. For any $\sigma \in \mathfrak{S}_n$, let $J^{\nu,\sigma}$ be a real function on $[0,1]$ such that $J^{\nu,\sigma}(i/n) = np_{n-i+1}^{\nu,\sigma}$. Then, Y_ν^σ can be rewritten as

$$Y_{\nu,n}^\sigma = \frac{1}{n}\sum_{i=1}^n J^{\nu,\sigma}\left(\frac{i}{n}\right) X_{i:n},$$

where the subscript n in $Y_{\nu,n}^\sigma$ is added to emphasize dependence on the sample. Furthermore, let

$$\alpha(J^{\nu,\sigma}, F) := \int_{-\infty}^{\infty} x J^{\nu,\sigma}[F(x)]dF(x) = \int_0^1 J^{\nu,\sigma}(u)F^{-1}(u)\mathrm{d}u,$$

and

$$\begin{aligned}\beta^2(J^{\nu,\sigma}, F) :&= 2\int_{-\infty<x<y<+\infty} J^{\nu,\sigma}(F(x))J^{\nu,\sigma}(F(y))F(x)(1-F(y))\mathrm{d}x\mathrm{d}y \\ &= 2\int_{0<u<v<1} J^{\nu,\sigma}(u)J^{\nu,\sigma}(v)u(1-v)\mathrm{d}F^{-1}(u)\mathrm{d}F^{-1}(v).\end{aligned}$$

Then, Stigler's results [28, Theorems 2 and 3] (see also [15, Theorem 11.4]) state that, if F has a finite variance and if $J^{\nu,\sigma}$ is bounded and continuous almost everywhere w.r.t. F^{-1}, one has

$$\lim_{n\to\infty} \mathbf{E}[Y_{\nu,n}^\sigma] = \alpha(J^{\nu,\sigma}, F), \qquad \lim_{n\to\infty} n\mathbf{V}[Y_{\nu,n}^\sigma] = \beta^2(J^{\nu,\sigma}, F),$$

and, if additionally $\beta^2(J^{\nu,\sigma}, F) > 0$,

$$\frac{Y_{\nu,n}^\sigma - \mathbf{E}[Y_{\nu,n}^\sigma]}{\sqrt{\mathbf{V}[Y_{\nu,n}^\sigma]}} \to_d N(0,1) \quad \text{as} \quad n \to \infty.$$

Example 4 *To illustrate the applicability of these results, consider the following game ν on N defined by*

$$\nu(S) = \sum_{j=1}^{|S|} \frac{1}{n}\left(\frac{n-j+1}{n}\right)^a, \qquad \forall S \subseteq N,$$

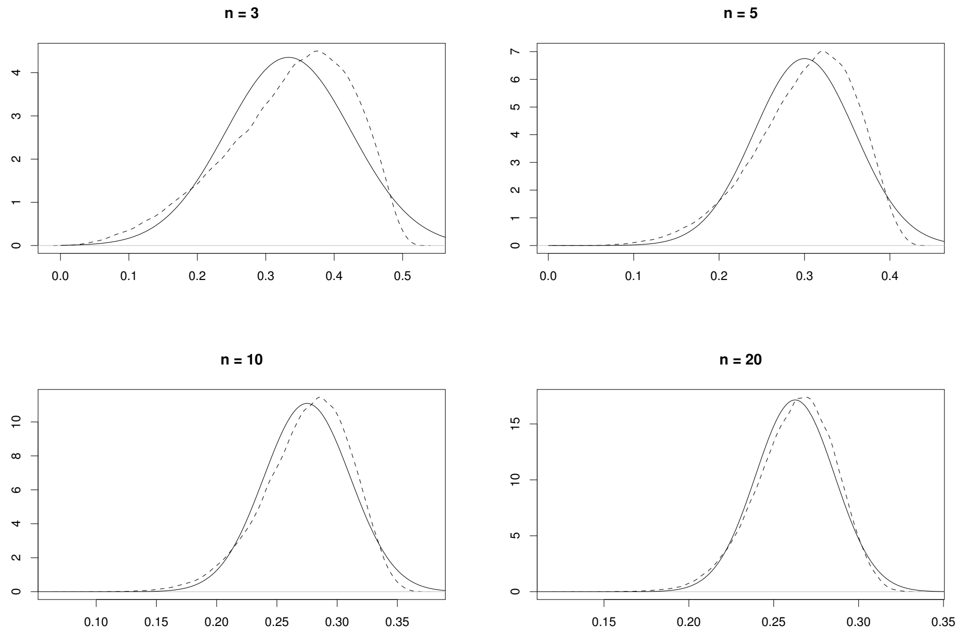

Fig. 2. Approximations of the p.d.f.s of discrete Choquet integral by mixtures of normals (solid lines) for $n = 3, 5, 10$ and 20. The dotted lines represent the corresponding p.d.f.s estimated by the kernel method from 10 000 randomly generated realizations.

where a is a strictly positive real number. We then have

$$p_i^{\nu,\sigma} = \frac{1}{n}\left(\frac{n-i+1}{n}\right)^a, \qquad \forall i \in N, \qquad \forall \sigma \in \mathfrak{S}_n.$$

As the coefficients $p_i^{\nu,\sigma}$ do not depend on σ, the corresponding Choquet integral is merely a linear combination of order statistics. Note however that the game ν is by no means additive. Next, define $J^{\nu,\sigma}(x) := x^a$, for all $x \in [0,1]$. Then, clearly, $J^{\nu,\sigma}(i/n) = np_{n-i+1}^{\nu,\sigma}$ for all $i \in N$.

In order to simplify the calculations, assume furthermore that F is the standard uniform c.d.f. and that $a = 2$. Then, $J^{\nu,\sigma}$ is clearly bounded and continuous almost everywhere w.r.t. F^{-1} and we have $\alpha(J^{\nu,\sigma}, F) = 1/4$ and $\beta^2(J^{\nu,\sigma}, F) = 1/112$.

The dotted lines in Figure 2 represent the p.d.f. of the Choquet integral w.r.t. ν estimated by the kernel method from 10 000 randomly generated realizations for $n = 3, 5, 10$ and 20. The solid lines represent the normal p.d.f.s $N(\mathbf{E}[Y_{\nu,n}^\sigma], \mathbf{V}[Y_{\nu,n}^\sigma])$, where $\mathbf{E}[Y_{\nu,n}^\sigma]$ and $\mathbf{V}[Y_{\nu,n}^\sigma]$ are computed by means of (9) and (10).

From the previous example, it clearly appears that one strong prerequisite before being able to apply the previous theoretical results is the knowledge of the expression of the game ν in terms of n. In practical applications of aggregation operators, this is rarely the case as ν is usually determined for

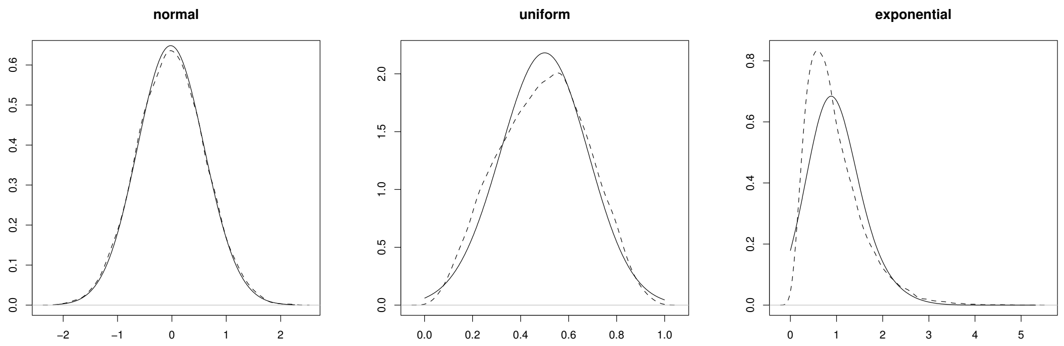

Fig. 3. Approximations of the p.d.f.s of discrete Choquet integrals by mixtures of normals (solid lines) in the standard normal, standard uniform and standard exponential cases. The dotted lines represent the corresponding p.d.f.s estimated by the kernel method from 10 000 randomly generated realizations.

some fixed n from learning data (see e.g. [29]). It follows that in such situations the above theoretical conditions cannot be rigorously verified.

In informal terms, Stigler [28] states that a linear combination of order statistics is likely to be asymptotically normally distributed if the extremal order statistics do not contribute “too much”, which is satisfied is the weights are “smooth” and “bounded”. When dealing with a Choquet integral, several numerical indices could be computed to assess whether the operator behaves in a too *conjunctive* (minimum-like) or too *disjunctive* (maximum-like) way. One such index is the degree of *orness* studied in [9,10].

Example 5 *Consider again the capacity given in Example 1. The degree of orness of this capacity, computed using the* `kappalab` *R package, is 0.49, which indicates a fairly neutral (slightly conjunctive) behavior. The solid lines in Figure 3 represent the mixtures of* $3! = 6$ *normals in the standard normal, standard uniform and standard exponential cases as possible approximations of the p.d.f. of the corresponding Choquet integral. As previously, the dotted lines represent the p.d.f.s estimated by the kernel method from 10 000 randomly generated realizations. As one can see, the approximation is very good in the standard normal case, may be considered as acceptable in the standard uniform case, and poor in the exponential case. Provided considering such a approximation is valid (which, as discussed above, cannot be verified), one could argue that the poor results in the exponential case are due to the too low value of* $n(= 3)$. *Although such low values for* n *make no sense in statistics, in multicriteria decision aiding for instance, they are quite common. In fact, in practical decision problems involving aggregation operators, the value of* n *is very rarely greater than 10.*

References

[1] J.-L. Marichal, I. Kojadinovic, Distribution functions of linear combinations of lattice polynomials from the uniform distribution, Statistics & Probability Letters 78 (2008) 985–991.

[2] G. Choquet, Theory of capacities, Annales de l'Institut Fourier 5 (1953) 131–295.

[3] M. Grabisch, Fuzzy integral in multicriteria decision making, Fuzzy Sets and Systems 69 (1995) 279–298.

[4] J.-L. Marichal, An axiomatic approach of the discrete Choquet integral as a tool to aggregate interacting criteria, IEEE Transactions on Fuzzy Systems 8 (6) (2000) 800–807.

[5] L. Lovász, Submodular function and convexity, in: A. Bachem, M. Grötschel, B. Korte (Eds.), Mathematical programming: The state of the art, Springer-Verlag, 1983, pp. 235–257.

[6] R. D. Yager, On ordered weighted averaging aggregation operators in multicriteria decision making, IEEE Trans. on Systems, Man and Cybernetics 18 (1988) 183–190.

[7] J.-L. Marichal, Cumulative distribution functions and moments of lattice polynomials, Statistics & Probability Letters 76 (12) (2006) 1273–1279.

[8] M. Grabisch, T. Murofushi, M. Sugeno (Eds.), Fuzzy measures and integrals, Vol. 40 of Studies in Fuzziness and Soft Computing, Physica-Verlag, Heidelberg, 2000, theory and applications.

[9] J.-L. Marichal, Behavioral analysis of aggregation in multicriteria decision aid, in: J. Fodor, B. D. Baets, P. Perny (Eds.), Preferences and Decisions under Incomplete Knowledge, Physica-Verlag, 2000, pp. 153–178.

[10] J.-L. Marichal, Tolerant or intolerant character of interacting criteria in aggregation by the Choquet integral, European Journal of Operational Research 155 (3) (2004) 771–791.

[11] L. Kuncheva, "Fuzzy" versus "nonfuzzy" in combining classifiers designed by boosting, IEEE Transactions on Fuzzy Systems 11 (6) (2003) 729–741.

[12] L. Kuncheva, A theoretical study on six classifier fusion strategies, IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (2) (2002) 281–286.

[13] M. Grabisch, I. Kojadinovic, P. Meyer, `kappalab`: Non additive measure and integral manipulation functions, R package version 0.4-3 (2008).

[14] J.-L. Marichal, On Choquet and Sugeno integrals as aggregation functions, in: M. Grabisch, T. Murofushi, M. Sugeno (Eds.), Fuzzy Measures and Integrals: Theory and Applications, Physica-Verlag, 2000, pp. 247–272.

[15] H. David, H. Nagaraja, Order statistics. 3rd ed., Wiley Series in Probability and Statistics. Chichester: John Wiley & Sons., 2003.

[16] P. J. Davis, Interpolation and approximation. 2nd ed., Dover Books on Advanced Mathematics. New York: Dover Publications, 1975.

[17] R. DeVore, G. Lorentz, Constructive approximation, Berlin: Springer-Verlag, 1993.

[18] M. Powell, Approximation theory and methods, Cambridge University Press, 1981.

[19] J. A. Adell, C. Sangüesa, Error bounds in divided difference expansions. A probabilistic perspective, J. Math. Anal. Appl. 318 (1) (2006) 352–364.

[20] G. Agarwal, R. Dalpatadu, A. Singh, Linear functions of uniform order statistics and B-splines, Commun. Stat., Theory Methods 31 (2) (2002) 181–192.

[21] M. M. Ali, Content of the frustum of a simplex, Pacific Journal of Mathematics 48 (2) (1973) 313–322.

[22] L. Gerber, The volume cut off a simplex by a half space, Pacific Journal of Mathematics 94 (2) (1981) 311–314.

[23] C. de Boor, On calculating with B-splines, Journal of Approximation Theory 6 (1972) 50–62.

[24] G. Varsi, The multidimensional content of the frustrum of the simplex, Pacific Journal of Mathematics 46 (1) (1973) 303–314.

[25] R Development Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0 (2008).
URL `http://www.R-project.org`

[26] F. David, N. Johnson, Statistical treatment of censored data, I. Fundamental formulae, Biometrika 41 (1954) 228–240.

[27] A. Childs, N. Balakrishnan, Series approximation for moments of order statistics using MAPLE, Computational Statistics and Data Analysis 38 (2002) 331–347.

[28] S. Stigler, Linear functions of order statistics with smooth weight functions, Ann. Statist. 2 (1974) 676–693, correction 7.

[29] M. Grabisch, I. Kojadinovic, P. Meyer, A review of methods for capacity identification in Choquet integral based multi-attribute utility theory: Applications of the Kappalab R package, European Journal of Operational Research 186 (2) (2008) 766–785.