On measures resulting from the Choquet integration*

Zuzana Ontkovičová[1,2][0000−0001−8443−5543] and
Vicenç Torra[1][0000−0002−0368−8037]

[1] Department of Computing Science, Umeå University, Sweden
[2] Institute of Information Engineering, Automation and Mathematics,
Faculty of Chemical and Food Technology, Slovak University of Technology in
Bratislava, Slovakia
{zuzanao, vtorra}@cs.umu.se

Abstract. In this paper, we study measures arising as a result of the Choquet integration with respect to a particular class of measures. The initial insight is provided with several classes of additive and fuzzy measures. It can be seen that some classes are closed regarding the integration, e.g. probabilities, while some are not, such as distorted Lebesgue measures. Knowing both an integration measure and a resulting measure after the Choquet integration directly leads to a fuzzy analogue of the Radon-Nikodym derivatives. For them, a completely different possible approach to their existence is presented for a specific pair of measures.

Keywords: Choquet integral · Fuzzy measures · Closed classes of measures regarding the integration · Choquet-Radon-Nikodym derivatives.

1 Introduction

In some real-life situations, a description of sets through additive measures is not sufficient. To fully cover connections between sets, fuzzy (also nonadditive) measures are used instead. The Choquet integral is proposed for integration with respect to fuzzy measures, generalising the additive Lebesgue integral. These fuzzy concepts have a broad scope of applications, such as decision-making, finance, game theory and artificial intelligence, to list a few.

Although both mentioned notions are well-studied, this article aims to look at them from a different perspective. It is possible to consider the result after the Choquet integration as a measure, and thus we want to know more, preferably its classification. So the prime question here is:

"What type of measure results after the Choquet integration with respect to a particular measure?"

* This work was partially supported by the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation. Partial support from the KEMPE foundation is also gratefully acknowledged.

With a knowledge of the specific type/class of the resulting measure, we directly obtain its basic properties. Moreover, the measure type is also interesting for the Choquet-Radon-Nikodym derivatives as a generalisation of the Radon-Nikodym derivatives for fuzzy measures. There, an objective is to find the integrated function (derivative) when knowing the integration measure and the resulting measure after integration. These derivatives are the crucial part of the definitions for divergences (e.g. Kullback-Leibler divergence, Total variation or Hellinger distance) and also differential entropy.

Throughout the paper, we assume a measurable space $(X, \mathcal{F})$ with X being a non-empty sample space and $\mathcal{F}$ a σ-algebra of all possible subsets of X. Then, measure $\mu : \mathcal{F} \to [0, \infty)$ is a set function assigning every subset of X a nonnegative real number with additional condition $\mu(\emptyset) = 0$. A measure is called additive if for any $A, B \in \mathcal{F}, A \cap B = \emptyset$ the additivity property holds $\mu(A \cup B) = \mu(A) + \mu(B)$. When additivity is weakened to only monotonicity, so if for all $A, B \in \mathcal{F}, A \subseteq B$ it is satisfied that $\mu(A) \leq \mu(B)$, then the measure is said to be fuzzy or nonadditive. Particular cases of fuzzy measures are submodular and supermodular measures, where for all $A, B \in \mathcal{F}$ it holds $\mu(A) + \mu(B) \geq \mu(A \cup B) + \mu(A \cap B)$ or $\mu(A) + \mu(B) \leq \mu(A \cup B) + \mu(A \cap B)$, respectively. Functions are said to be measurable if for every $t \in \mathbb{R}$ it is satisfied that $\{x \in X : f(x) < t\} \in \mathcal{F}$.

Regarding integration, the Choquet integral is a fluent transition from additive measures with the Lebesgue integral to fuzzy measures. Its definition for any measurable nonnegative function on an arbitrary set $A \subseteq X$ is given as

$$(C) \int_A g \, \mathrm{d}\mu = \sum_{i=1}^{n} g_{\sigma(i)} \, \mathbb{1}_A \, \left(\mu[\sigma(i), \dots, \sigma(n)] - \mu[\sigma(i+1), \dots, \sigma(n)]\right)$$

in the discrete case with $X = [n] = \{1, \dots, n\}$, where $\sigma : [n] \to [n]$ is a permutation of function values taken as $g_{\sigma(i)} \leq g_{\sigma(i+1)}$ for all $i \in [n]$ with the convention $\sigma(n+1) = 0$ and $\mathbb{1}_A$ is the characteristic function of the set A, or as

$$(C) \int_A g \, \mathrm{d}\mu = \int_0^\infty \mu\left(\{x \in X : g(x) \geq t\} \cap A\right) \mathrm{d}t$$

for the continuous case, as first proposed in [1] and later summarised e.g. in [2].

After the motivation and necessary preliminaries above, the paper consists of two sections, structured as follows. In Section 2, we study measures resulting from the Choquet integration with respect to particular types of measures. An issue in Section 3 is the connection between the resulting and integration type of measures and their possible closure. Also, a relation of this connection with the Choquet-Radon-Nikodym derivatives is shown.

2 A look at resulting measures

Let us consider measures built from measures through the Choquet integration. For simplification, we use notation

$$\nu_{g,\mu}(A) = (C)\int_A g\,\mathrm{d}\mu,$$

where $A \subseteq X$ is an integration set, $g : A \to \mathbb{R}_0^+$ a measurable nonnegative integrated function and $\mu : X \to \mathbb{R}_0^+$ an integration measure. Note that when μ is additive, the Choquet integral reduces to the Lebesgue integral.

As stated in [2], the Choquet integral is additive if and only if the integration measure is additive. In our context, it means that when taking additive integration measure μ_{AD}, the resulting measure after the Choquet integration is also additive, so the additivity property holds

$$(\forall A, B \in \mathcal{F},\ A \cap B = \emptyset) \qquad \nu_{g,\mu_{AD}}(A \cup B) = \nu_{g,\mu_{AD}}(A) + \nu_{g,\mu_{AD}}(B).$$

Moreover, we can say that with the integration measure being fuzzy μ_F, also the resulting measure is fuzzy. Then the additivity property is replaced by the weaker property of monotonicity

$$(\forall A, B \in \mathcal{F},\ A \subseteq B) \qquad \nu_{g,\mu_F}(A) \leq \nu_{g,\mu_F}(B).$$

Is it also possible to show something similar about the resulting measure after the integration for specific classes within additive or fuzzy measures? In other words, does integrating with respect to one measure always result in another measure from the same class? Let us find an answer in the following parts, separately for some particular types of measures.

We start with additive measures, where the Choquet integral naturally reduces to the additive Lebesgue integral. First, we take the probability measure μ_P. It is given as $\mu_P(A) = \sum_{x\in A} p(x)$ on the discrete space and as $\mu_P(A) = \int_A p\,\mathrm{d}\lambda$ on the continuous space, with λ being the Lebesgue measure, p the corresponding density function and an additional condition of normalisation $\mu_P(X) = 1$ needs to be satisfied.

Proposition 1. *The measure ν_{g,μ_P} in the discrete case corresponds to μ_{g*P} where $g * P$ stands for $(g * P)(A) = \sum_{x\in A} g(x)p(x)$.*

The result is easy to see since $\nu_{g,\mu_P}(A) = \sum_{i=1}^{n} g_{\sigma(i)}\, \mathbb{1}_A\, p(\sigma(i)) = \sum_{x\in A} g(x)\,p(x)$. The same result is obtained when taking the continuous case, only there $g * P$ stands for $(g * P)(A) = \int_A g\,p\,\mathrm{d}\lambda$. Hence, the summary can be done for the probability measure on the general space.

Theorem 1. *The measure ν_{g,μ_P} corresponds to μ_{g*P}.*

If an additional condition of $\sum_{x\in X} g(x)p(x) = 1$ or $\int_X g\,p\,\mathrm{d}\lambda = 1$ is satisfied in the corresponding setup, measure μ_{g*P} is a probability measure.

Now, we look at the additive counting measure $\mu_\#$ (equivalent to Lebesgue measure in the discrete case) defined as $\mu_\#(A) = |A|$ for all finite sets.

Proposition 2. *The measure $\nu_{g,\mu_\#}$ corresponds to μ_P, where $p \equiv g$.*

The result is obvious because $\nu_{g,\mu_\#}(A) = \sum_{i=1}^{n} g_{\sigma(i)}\,\mathbb{1}_A(i+1-i) = \sum_{x\in A} g(x)$. In the continuous case, we assume Lebesgue measure λ as an equivalent to the counting measure and get the same result.

Proposition 3. *The measure $\nu_{g,\lambda}$ corresponds to μ_P with $p \equiv g$.*

Again, with an additional condition $\sum_{x\in X} g(x) = 1$ or $\int_X g\,\mathrm{d}\lambda = 1$, the resulting measure is a probability measure.

From now on, we focus on fuzzy measures. From properties of the Choquet integral for nonnegative functions in [2], it is easy to see that taking a submodular measure μ_{submod} results in measure $\nu_{g,\mu_{submod}}$ being submodular (also subadditive and convex). Similarly, with a supermodular measure $\mu_{supermod}$, measure $\nu_{g,\mu_{supermod}}$ is supermodular (also superadditive and concave).

The first class of fuzzy measures we assume is unanimity measures in a generalised fashion. We denote them as μ_{A_0,k_0} for a fixed constant $k_0 \in (0,1]$ (case with $k = 0$ leads to zero measure) and a fixed set $A_0 \subseteq X$, and define as $\mu_{A_0,k_0}(A) = \begin{cases} k_0, & A \supseteq A_0 \\ 0, & A \subset A_0 \end{cases}$. If $k_0 = 1$, shorter notation μ_{A_0} is used. Notice that minimal fuzzy measure can be seen as a special case with the formula $\mu_{min}(A) \equiv \mu_X(A) = \begin{cases} 1, & A = X \\ 0, & A \subset X \end{cases}$. Inductively going from particular examples to the whole class, the following propositions study the integration result.

Proposition 4. *The fuzzy measure ν_{g,μ_X} corresponds to μ_{X,k_0} for a certain k_0.*

Proof. The idea of the proof can be simply outlined in the discrete case. It is necessary to realise that measure μ_X is non-zero only when all the elements of X are included. That corresponds to the term with $\sigma(1)$-st function value, so the smallest one. Thus more generally, if $A \neq X$ then $\nu_{g,\mu_X}(A) = 0$, and if $A = X$ then $\nu_{g,\mu_X}(X) = (C)\int_X g\,\mathrm{d}\mu_X = \min_{x\in X} g(x)$. Therefore for $k_0 = \min_{x\in X} g(x)$ we have $\nu_{g,\mu_X} = \mu_{X,k_0}$.

Similarly, it can also be done for μ_{A_0}. The measure is zero for all smaller sets than A_0, so if $A \subset A_0$ then $\nu_{g,\mu_{A_0}}(A) = 0$. The case with $A = A_0$ corresponds to $A = X$ in the previous proposition. Hence, the result after integration is the smallest value of a function on A_0, resulting in $\nu_{g,\mu_{A_0}}(A_0) = \min_{x \in A_0} g(x)$. For bigger sets than A_0, the measure is equal to 1. In the discrete case, which is more illustrative here, it means that the difference of measures in the definition is zero everywhere but in the one term corresponding to the smallest function value on the set A_0. So, for any $A \supset A_0$ it holds $\nu_{g,\mu_{A_0}}(A) = \min_{x \in A_0} g(x)$. The final result is summarised in the next proposition.

Proposition 5. *The fuzzy measure $\nu_{g,\mu_{A_0}}$ corresponds to μ_{A_0,k_0} when taking $k_0 = \min_{x \in A_0} g(x)$.*

The most general case of this class of measures can be derived similarly as in the last proposition. The only difference is a height of the jump (difference between corresponding measures in the discrete definition), which is not equal to 1 but to k.

Theorem 2. *The fuzzy measure $\nu_{g,\mu_{A_0,k}}$ corresponds to μ_{A_0,k_0} when assuming $k_0 = k \min_{x \in A_0} g(x)$.*

With these results, we can say that the resulting measure after the Choquet integration with respect to μ_{A_0,k_0} is of the same type.

Similar to μ_{min} (minimal measure) introduced above is the maximal measure given as $\mu_{max}(A) = \begin{cases} 1, & A \supset \emptyset \\ 0, & A = \emptyset \end{cases}$. Strictly speaking, it is not a special case of the previous class μ_{A_0,k_0} in the form of $\mu_\emptyset$. The reason is that taking $A_0 = \emptyset$ means that the zero value is assigned to every set, so we should have the trivial zero measure, which is clearly not true. For the purpose of the following proposition, recall that μ_π is a possibility (also maxitive) measure with $\pi : X \to \mathbb{R}_0^+$ given as $\mu_\pi(A) = \max_{x \in A} \pi(x)$.

Proposition 6. *The fuzzy measure $\nu_{g,\mu_{max}}$ corresponds to μ_π with $\pi(x) \equiv g(x)$.*

Proof. First, we assume the discrete case for more illustrative insight. This measure changes its value just for the empty set, so the only non-zero term in the Choquet integral definition is obtained when subtracting the smallest and empty set. That corresponds to the $\sigma(n)$-th function value, hence the biggest one. Then more generally, if $A = \emptyset$ it is clear that $\nu_{g,\mu_{max}}(A) = 0$. For other sets A, we have $\nu_{g,\mu_{max}}(A) = \max_{x \in A} g(x)$. In particular, $\nu_{g,\mu_{max}}(\{x\}) = g(x)$.

Now, restricting ourselves to $X = \mathbb{R}$, let us take the class of distorted Lebesgue measures denoted as λ_m. With λ being the Lebesgue measure and

$m : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ a nondecreasing (sometimes taken as strictly increasing) distortion where $m(0) = 0$, the measure is defined as $\lambda_m([a, b]) = m(\lambda([a, b])) = m(b - a)$. For a simplifying approach, we focus on monotone functions as in [3], yet using computational formulas for a more general setup from [4]. So, assuming nonnegative monotone continuous functions and differentiable distortions, computationally convenient formulas for the Choquet integral on an arbitrary set $[s, \tau] \subseteq \mathbb{R}$, $\tau > s$ are given

- for nondecreasing function g as

$$\nu_{g,\lambda_m}([s, \tau]) = \int_s^\tau m'(\tau - \alpha)\, g(\alpha)\, \mathrm{d}\alpha \tag{1}$$

- for nonincreasing function g as

$$\nu_{g,\lambda_m}([s, \tau]) = \int_s^\tau m'(\alpha - s)\, g(\alpha)\, \mathrm{d}\alpha. \tag{2}$$

Remark 1. Even though we are restricted to only monotone functions here, our conclusions are not necessarily less general. The reason is a reordering method proposed in [5] and [6], which reorders a non-monotone function with respect to a distorted Lebesgue measure to a monotone one without changing its Choquet integral value. The method can be sketched as follows. Let us assume a non-monotone continuous function $g : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ on a set $[0, \tau]$, where $\hat{g} = \max_{0 \le x \le \tau} g(x)$. Defining $\lambda_g : [0, \hat{g}] \to [0, \tau]$ as $\lambda_g(\alpha) = t = \tau - \lambda(\{x : g(x) \ge \alpha\})$, a function $g^* : [0, \tau] \to [0, \hat{g}]$ given as $g^*(t) = \alpha = \lambda_g^{-1}(t)$ is called a rearrangement of g. Moreover $(C)\int_{[0,\tau]} g^*\, \mathrm{d}\lambda_m = (C)\int_{[0,\tau]} g\, \mathrm{d}\lambda_m$. Since g^* is a nondecreasing and continuous function with the same value of the Choquet integral as the original non-monotone function g, the computational formula (1) can be used. Even more general case of reordering is proposed in [4].

Thinking about the resulting measure after the Choquet integration with distorted Lebesgue measures, a hypothesis may be that ν_{g,λ_m} corresponds to another distorted Lebesgue measure with a different distortion, let us say λ_n. Consequently, if it is correct, the value of the Choquet interval for an arbitrary function and an arbitrary distortion should be the same for all intervals of the same length because it is the property of distorted Lebesgue measures. The following example shows a sketch of this idea in a particular setup.

Example 1. Let us take two real intervals $[0, \tau]$ and $[\tau, 2\tau]$, $\tau > 0$, where clearly $\lambda_m([0, \tau]) = \lambda_m([\tau, 2\tau])$. Then, assuming an increasing function $g_1(x) = e^x$ and a distortion $m(x) = x^2$, we use the formula (1) to compute the Choquet integral for both intervals. Thus, $\nu_{g_1,\lambda_m}([0, \tau]) = \int_0^\tau 2(\tau - \alpha)\, e^\alpha\, \mathrm{d}\alpha = 2e^\tau - 2\tau - 2$ and similarly $\nu_{g_1,\lambda_m}([\tau, 2\tau]) = \int_\tau^{2\tau} 2(2\tau - \alpha)\, e^\alpha\, \mathrm{d}\alpha = 2e^{2\tau} - 2\tau e^\tau - 2e^\tau$. Comparing

these two results, they are apparently not equal for any $\tau > 0$.
Analogously, the computation can be done for a decreasing function $g(x) = e^{-x}$ and the same distortion using the formula (2). Then, for both intervals we can compute the values as $\nu_{g_2,\lambda_m}([0,\tau]) = \int_0^{\tau} 2\alpha\, e^{-\alpha}\,\mathrm{d}\alpha = 2 - 2\tau e^{-\tau} - 2e^{-\tau}$ and $\nu_{g_2,\lambda_m}([\tau,2\tau]) = \int_{\tau}^{2\tau} 2(\alpha-\tau)\, e^{-\alpha}\,\mathrm{d}\alpha = 2e^{-\tau} - 2\tau e^{-2\tau} - 2e^{-2\tau}$. Again, these two results are not equal for any $\tau > 0$.

From Example 1, we can conclude that ν_{g,λ_m} does not correspond in general to any distorted Lebesgue measure since the results for two intervals of the same length are not equal. Although, based on particular examples and our intuition, it is possible to propose a hypothesis.

Hypothesis 1 *The fuzzy measure ν_{g,λ_m} corresponds to μ_{P_n} with P_n being a distorted probability with a corresponding distortion n (maybe except for the normalisation condition).*

3 Closure regarding the integration

It can be said that if μ as an integration measure and $\nu_{g,\mu}$ as a resulting measure after the integration are from the same class, then this class is closed regarding the Choquet integration. Table 1 summarises related results from the previous section, adding information on whether or not the particular class of measures is closed.

Integration measure μ	Resulting measure $\nu_{g,\mu}$	Closed?
μ_{AD}	μ_{AD}	✓
μ_P	μ_P	✓
$\mu_{\#}$ / λ	μ_P	✗
μ_F	μ_F	✓
μ_{submod}	μ_{submod} $(\mu_{subadd} / \mu_{convex})$	✓
$\mu_{supermod}$	$\mu_{supermod}$ $(\mu_{superadd} / \mu_{concave})$	✓
$\mu_{A_0,k}$	μ_{A_0,k_0}	✓
μ_{max}	μ_{π}	✗
λ_m	$?P_n?$	✗

Table 1. Pairs of measures connected through the Choquet integration.

At this stage, it is interesting to point out a connection between closed classes of measures after the Choquet integration and the Choquet-Radon-Nikodym

(CRN) derivatives. These derivatives were studied mostly in [3], [6], [7], and as the naming suggests, they generalise the additive Radon-Nikodym derivatives for fuzzy measures. The CRN derivative can be seen as an inverse approach to the Choquet integral computation. Using our notation for the Choquet integral and a notation for the CRN derivatives from [7], it can be written that

$$\nu_{g,\mu}(A) = (C)\int_A g \,\mathrm{d}\mu \quad \iff \quad g = \frac{\partial \nu_{g,\mu}}{\partial \mu}.$$

So, the basic idea behind the CRN derivatives is that we know both an integration measure μ and a resulting measure $\nu_{g,\mu}$ and want to find an integrated function g. However, even conditions for the existence of the CRN derivatives are either too complicated - the strong decomposition property in [8], or too strict - submodularity of both measures in [9].

This article opens a new possibility on how to look at the CRN derivatives from another perspective, which is connected directly with the better-studied concept of the Choquet integral and its computation. We can propose some preliminary results considering the corresponding pairs of integration and resulting measures obtained in this article. In some cases, such as probabilities in Theorem 1 or generalised unanimity measures in Theorem 2, we know that the CRN derivatives exist and have also found particular forms for the measures. On the other hand, although some classes of measures are closed under the Choquet integration, e.g. submodular or supermodular measures, nothing can be said even about the existence of the CRN derivatives in general because the class is too broad to propose any specific results. Possibly the most interesting classes of measures are those not closed regarding the Choquet integration, for instance distorted Lebesgue measures as shown in Example 1. What can be concluded about the CRN derivatives and their existence for these classes?

Hypothesis 2 *If a class of measures is not closed regarding the Choquet integration, then the CRN derivative corresponding to any two measures from this class equals a constant.*

An intuition behind this idea can be illustrated on the class of distorted Lebesgue measures. Since it is not closed with respect to the Choquet integration, the easiest (and probably the only) way to get a distorted Lebesgue measure also as a resulting measure is to integrate a (positive) constant, so then $(C)\int_A k \,\mathrm{d}\lambda_m = k\,\lambda_m(A) = \lambda_n(A)$ with $n(x) = k\,m(x)$. Notice, that this is not a violation of our previous Hypothesis 1 because distorted Lebesgue measures can be seen as a trivial case of distorted probabilities with $p(x) = x$ (only without the normalisation condition as suggested in the hypothesis).

4 Conclusion

As a continuation of this article, more classes of fuzzy measures could be studied. A key to a better understanding and further results of the concepts presented here is proving (or disproving) the Hypotheses 1 and 2. They represent interesting notions regarding the Choquet integration and its closure. Also, they provide a better insight into the Choquet-Radon-Nikodym derivatives, their conditions for existence and possible computational formulas.

References

1. Choquet, G.: Theory of Capacities. Annales de l'Institut Fourier **5**, 131–295 (1954).
2. Grabisch, M.: Set functions, games and capacities in decision making. Springer, Berlin, Vol. 46.(2016). https://doi.org/10.1007/978-3-319-30690-2
3. Sugeno, M.: A note on derivatives of functions with respect to fuzzy measures. Fuzzy Sets and Systems **222**, 1–17 (2013). https://doi.org/10.1016/j.fss.2012.11.003
4. Ontkovičová, Z., Torra, V.: Computation of Choquet integrals: Analytical approach for continuous functions. Manuscript submitted for publication (2024).
5. Narukawa, Y., Torra, V., Sugeno, M.: Choquet integral with respect to a symmetric fuzzy measure of a function on the real line. Annals of Operations Research **244**(2) 571-581 (2016). https://doi.org/10.1007/s10479-012-1166-6
6. Sugeno, M.: A way to Choquet calculus. IEEE Transactions on Fuzzy Systems **23**(5), 1439–1457 (2014). https://doi.org/10.1109/TFUZZ.2014.2362148
7. Ontkovičová, Z., Kiseľák, J.: A way to proper generalization of ϕ-divergence based on Choquet derivatives. Soft Computing, **26**(21), 11295–11314 (2022). https://doi.org/10.1007/s00500-022-07381-5
8. Graf, S.: A Radon-Nikodym theorem for capacities. Journal für die reine und angewandte Mathematik, **1980**(320), 192–214 (1980).
9. Huber, P.J., Strassen, V.: Minimax tests and the Neyman-Pearson lemma for capacities. The Annals of Statistics, **1**(2), 251–263 (1973).