Random means generated by random variables: expectation and limit theorems

Mátyás Barczy*,◇, Pál Burai**

* MTA-SZTE Analysis and Stochastics Research Group, Bolyai Institute, University of Szeged, Aradi vértanúk tere 1, H–6720 Szeged, Hungary.

** Department of Differential Equations, Institute of Mathematics, Faculty of Natural Sciences, Budapest University of Technology and Economics, Egry József utca 1, H–1111 Budapest, Hungary.

e-mail: barczy@math.u-szeged.hu (M. Barczy), buraip@math.bme.hu (P. Burai).

◇ Corresponding author.

Abstract

We introduce the notion of a random mean generated by a random variable and give a construction of its expected value. We derive some sufficient conditions under which strong laws of large numbers and some limit theorems hold for random means generated by the elements of a sequence of independent and identically distributed random variables.

1 Introduction

The theory of means (also called averages) is an important, rich and growing field of mathematics, and it has several applications in practice as well. For a recent monograph on averaging functions and their applications, see Beliakov et al. [3]. Random-valued mappings (functions) also appear in various fields of mathematics such as in transportation theory (see, e.g., Panaretos and Zemel [12]), in iterative functional equations (see, e.g., Baron and Jarczyk [2], Baron [1], and Jarczyk and Jarczyk [7]) or in theory of random measures (see, e.g., Kallenberg [9, Chapter 1]).

In the present paper we introduce random (valued) means generated by a random variable and we give a construction of their expectations as well, which turn out to be usual (non-random) means. Further, we derive some sufficient conditions under which strong laws of large numbers and some limit theorems hold for random means generated by the elements of a sequence of independent and identically distributed random variables.

Concerning the notion of a random mean generated by a random variable given in Definition 2.2, there are (at least) two related notions in the literature, namely, the (continuous) family of

2020 Mathematics Subject Classifications: 60F05, 26E60, 46G12

Key words and phrases: random means, Bochner integral, random Hölder means, expectation, strong law of large numbers, limit theorem.

means in the sense of Páles and Zakaria [11, page 794] and the (continuous) random mean in the sense of Jarczyk and Jarczyk [8, page 6838]. In Remark 2.3 we compare our Definition 2.2 with these two related concepts. Here we only note that in their definitions there is no random variable a priori involved, and in their setups a random mean is a real valued mapping, while our random mean generated by a random variable maps into the set of continuous (non-random) means on a nondegenerate, compact interval of the real numbers. For historical fidelity, we mention that our Definition 2.2 was motivated by the definition of a random mean due to Jarczyk and Jarczyk [7, page 6838].

In Section 2 we introduce a notion of a random (valued) mean generated by a random variable. Roughly speaking, given an $\mathbb{R}^d$-valued random variable ξ, a p-variable random mean generated by ξ is a measurable function M from a probability space to the space consisting of all (continuous) means on a nondegenerate, compact interval I of the real numbers $\mathbb{R}$ such that there exists an auxiliary Borel measurable mapping $M_\xi : I^p \times \mathbb{R}^d \to I$ in a way that the mappings $M_\xi(\cdot, \xi(\omega))$ and $(M(\omega))(\cdot)$ coincides for almost every ω. For a precise definition, see Definition 2.2. We illustrate this definition by presenting a general method for constructing such random means (see Theorem 2.4), and we also give some examples such as discrete random means generated by discrete random variables and random Hölder means (see Examples 2.5 and 2.6).

The concept of Bochner integral (integral of maps defined on a measure space with values in a Banach space) allows us to define the expectation of a random mean generated by a random variable, see Definition 2.9. It turns out that the expectation in question is a usual (non-random) mean, see Theorem 2.10. In Examples 2.13 and 2.14 we calculate the expectation of some random means generated by random variables given in Examples 2.5 and 2.6. We derive that the expectation of a 2-variable random Hölder mean with weights governed by a uniform distribution on the interval $(0, 1)$ is nothing else but a Cauchy mean corresponding to some power functions or a logarithmic mean, see Example 2.14. In Remark 2.15, motivated by Examples 2.6 and 2.14, we initiate some possible future research directions.

Concerning the expectation of random means generated by a random variable, in probability theory there exist (at least) two somewhat related notions, namely, the expectation (also called barycenter) of a random probability measure on a compact metric space, see, e.g., Borsato et al. [4, Appendix A.2], and the Fréchet mean of a random measure with values in the 2-Wasserstein space on $\mathbb{R}^d$, see, e.g., Panaretos and Zemel [12, Section 3.2]. In Remark 2.12 we recall both notions in order to see the similarities and differences with the expectation of a random mean generated by a random variable given in Definition 2.9.

In Section 3 we derive some sufficient conditions under which strong laws of large numbers and some limit theorems hold for random means generated by the elements of a sequence of independent and identically distributed random variables. More precisely, given a sequence of random means generated by the elements of a sequence of independent and identically distributed random variables, we consider the supremum norm of the difference of the arithmetic mean of the first n random means in question and the common expectation of the random means, and we investigate the asymptotic behaviour of this random quantity as $n \to \infty$. In Theorem 3.1, the underlying sequence consists of independent and identically distributed discrete random variables, but the range of the random means in question is arbitrary in the sense that we do not suppose any special form of the usual (non-random) means in the range. In Corollary 3.2 we consider a special case of Theorem 3.1, namely, when the underlying sequence consists of independent, identically and Bernoulli distributed random variables and the range of the random means in question is the set consisting of the arithmetic and geometric means in $[0,1]$. In Theorem 3.3 we establish some limit theorems for p-variable randomly weighted arithmetic means, when the underlying independent and identically distributed random variables are not necessarily discrete, so this result is out of scope of Theorem 3.1. In case of $p = 2$ we also formulate a corollary of Theorem 3.1 by simplifying the limit distribution, see Corollary 3.4. Finally, we provide limit theorems for randomly weighted power means (which can be also called random Hölder means), and in this case instead of the arithmetic mean of the first n random means in question we consider their geometric mean, and so the limit theorems have somewhat different forms compared to the previous ones. For a comparison of our limit theorems for random means with the Wasserstein law of large numbers for Fréchet means, see Remark 3.6.

Section 4 is devoted to the proofs of Section 3. The main tools are the Kolmogorov's strong law of large numbers and the multidimensional central limit theorem together with the continuous mapping theorem.

We close the paper with an appendix, where we recall and prove a result on the continuity of the supremum for a two-variable continuous real-valued function by taking the supremum in one of its variables, see Theorem A.1. This result is used in the applications of the continuous mapping theorem in some of the proofs in Section 4.

2 Random means and their expectation

Let $\mathbb{N}$, $\mathbb{Z}_+$, $\mathbb{R}$ and $\mathbb{R}_+$ denote the sets of positive integers, non-negative integers, real numbers and non-negative real numbers, respectively. An interval $I \subset \mathbb{R}$ is called nondegenerate if it contains at least two distinct points. We denote by $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ the Euclidean inner product of $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$, where $d \in \mathbb{N}$. The Borel σ-algebra on $\mathbb{R}^d$ is denoted by $\mathcal{B}(\mathbb{R}^d)$. Convergence almost surely, convergence in distribution and equality in distribution will be denoted by $\xrightarrow{\text{a.s.}}$, $\xrightarrow{\mathcal{D}}$ and $\stackrel{\mathcal{D}}{=}$, respectively. For any $d \in \mathbb{N}$, $\mathcal{N}_d(\mathbf{0}, \boldsymbol{Q})$ denotes a d-dimensional normal distribution on $\mathbb{R}^d$ with mean vector $\mathbf{0} \in \mathbb{R}^d$ and covariance matrix $\boldsymbol{Q} \in \mathbb{R}^{d \times d}$. In case of $d = 1$, instead of $\mathcal{N}_1$ we simply write $\mathcal{N}$, and $\mathcal{N}(0,0)$ denotes the Dirac distribution concentrated at 0, i.e., for each $A \in \mathcal{B}(\mathbb{R})$, $(\mathcal{N}(0,0))(A) := 1$ if $0 \in A$, and $(\mathcal{N}(0,0))(A) := 0$ if $0 \notin A$.

Let I be a nondegenerate, compact interval of $\mathbb{R}$, $p \in \mathbb{N}$ be a positive integer, and let us denote by $\mathcal{C}(I^p)$ the vector space of real-valued, continuous functions defined on I^p, which becomes a Banach space with the usual supremum norm given by $\|u\| := \sup_{x \in I^p} |u(x)|$ for

$u \in \mathcal{C}(I^p)$. The Borel σ-algebra on $\mathcal{C}(I^p)$ is denoted by $\mathcal{B}(\mathcal{C}(I^p))$. Given a function $f : I^p \times \mathbb{R}^d \to I$, for any $y \in \mathbb{R}^d$, we will denote by $f(\cdot, y)$ the function $I^p \ni x \mapsto f(x, y)$.

An $m \in \mathcal{C}(I^p)$ is said to be a p-variable, continuous mean on I if

$$\min(x) \leqslant m(x) \leqslant \max(x), \qquad x \in I^p,$$

where $\min(x)$ and $\max(x)$ denotes the minimum and the maximum of the coordinates of $x \in I^p$, respectively. If the above inequalities are strict whenever x has at least two different coordinates, then m is called a strict mean.

From now on, we just simply use the terminology mean instead of continuous mean on I if there is no ambiguity. If $p = 1$, then the only 1-variable mean m is $m(x) = x$, $x \in I$.

The subset $\mathcal{M}_p \subset \mathcal{C}(I^p)$ denotes the class of p-variable means.

2.1 Proposition. *The set $\mathcal{M}_p \subset \mathcal{C}(I^p)$ is convex, bounded and closed with respect to the supremum norm.*

Proof. The convexity of $\mathcal{M}_p$ follows easily. Concerning boundedness, if $m \in \mathcal{M}_p$, then using that $m(x_1, \ldots, x_p) = x_1$ whenever $x_1 = \cdots = x_p \in I$, we have

$$\|m\| = \sup_{x \in I^p} |m(x)| = \max_{t \in I} |t| < \infty,$$

since I is compact. This means that the elements of $\mathcal{M}_p$ have the same (supremum) norm.

For proving closedness, let us assume that $m_n \in \mathcal{M}_p$, $n \in \mathbb{N}$, is a convergent sequence in $\mathcal{C}(I^p)$, and let us denote by $u \in \mathcal{C}(I^p)$ its limit. We intend to prove that

$$\min(x) \leqslant u(x) \leqslant \max(x), \qquad x \in I^p. \tag{2.1}$$

Because convergence in norm implies pointwise convergence, we have

$$m_n(x) \to u(x) \qquad \text{as } n \to \infty, \qquad x \in I^p. \tag{2.2}$$

Since $m_n \in \mathcal{M}_p$, $n \in \mathbb{N}$, we have $\min(x) \leqslant m_n(x) \leqslant \max(x)$, $x \in I^p$, $n \in \mathbb{N}$. So using (2.2), we have (2.1), as desired. □

For $m_1, m_2 \in \mathcal{M}_p$, let

$$\varrho(m_1, m_2) := \sup_{x \in I^p} |m_1(x) - m_2(x)|. \tag{2.3}$$

Note that $\mathcal{M}_p$ is metric space furnished with ϱ as a metric, but not a linear space.

2.2 Definition. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $\xi \colon \Omega \to \mathbb{R}^d$ be a d-dimensional random variable, where $d \in \mathbb{N}$. The map $M \colon \Omega \to \mathcal{M}_p$ is called a (p-variable, continuous) random mean (in I) generated by ξ if the following conditions are fulfilled:*

(i) M is $(\mathcal{A}, \mathcal{B}(\mathcal{C}(I^p)))$-measurable, that is to say, M is an $\mathcal{M}_p$-valued random variable,

(ii) *there is a* $(\mathcal{B}(I^p)\times\mathcal{B}(\mathbb{R}^d),\mathcal{B}(I))$*-measurable map* $M_\xi\colon I^p\times\mathbb{R}^d\to I$ *such that*

(a) *for every* $\omega\in\Omega$ *the map* $I^p\ni x\mapsto M_\xi(x,\xi(\omega))$ *is in* $\mathcal{C}(I^p)$, *and*

(b) $\mathbb{P}(\{\omega\in\Omega:(M(\omega))(x)=M_\xi(x,\xi(\omega))$ *for every* $x\in I^p\})=1$.

Concerning the notion of a random mean generated by a random variable given in Definition 2.2, there are (at least) two related notions in the literature, namely, the (continuous) family of means in the sense of Páles and Zakaria [11, page 794] and the (continuous) random mean in the sense of Jarczyk and Jarczyk [8, page 6838]. In the next remark we compare Definition 2.2 with these two related concepts.

2.3 Remark. (i) If M is a random mean generated by ξ in the sense of Definition 2.2 such that the map $M_\xi:I^p\times\mathbb{R}^d\to I$ (appearing in part (ii) of Definition 2.2) satisfies that for every $y\in\mathbb{R}^d$ the map $I^p\ni x\mapsto M_\xi(x,y)$ is in $\mathcal{M}_p$, then for each non-empty open interval J of I, we have M_ξ restricted to $J^p\times\mathbb{R}^d$ is a continuous family of p-variable means on J corresponding to the measurable space $(\mathbb{R}^d,\mathcal{B}(\mathbb{R}^d))$ in the sense of Páles and Zakaria [11, page 794].

(ii) If M is a random mean generated by ξ in the sense of Definition 2.2 such that $M_\xi:I^p\times\mathbb{R}^d\to I$ (appearing in part (ii) of Definition 2.2) satisfies that for every $y\in\mathbb{R}^d$ the map $I^p\ni x\mapsto M_\xi(x,y)$ is in $\mathcal{M}_p$, then M_ξ is a continuous random mean on I in the sense of Jarczyk and Jarczyk [8, page 6838] corresponding to the probability space $(\mathbb{R}^d,\mathcal{B}(\mathbb{R}^d),\mathbb{P}_\xi)$, where $\mathbb{P}_\xi$ denotes the distribution of ξ (i.e., $\mathbb{P}_\xi(A):=\mathbb{P}(\xi\in A)$, $A\in\mathcal{B}(\mathbb{R}^d)$). Indeed, M_ξ is $(\mathcal{B}(I^p)\times\mathcal{B}(\mathbb{R}^d),\mathcal{B}(I))$-measurable, $\{y\in\mathbb{R}^d:M_\xi(\cdot,y)\in\mathcal{M}_p\}=\mathbb{R}^d$ and hence

$$\mathbb{P}_\xi(\{y\in\mathbb{R}^d:M_\xi(\cdot,y)\in\mathcal{M}_p\})=\mathbb{P}_\xi(\mathbb{R}^d)=1.$$

We point out to the facts that both in the definition of a (continuous) family of p-variable means on J due to Páles and Zakaria [11, page 794] and in the definition of a (continuous) random mean on I due to Jarczyk and Jarczyk [8, page 6838] there is no random variable a priori involved.

Further, a (continuous) family of means on J and a random mean on I is a J-valued and an I-valued mapping, respectively, while our random mean generated by a random variable is an $\mathcal{M}_p$-valued mapping. □

Next, we illustrate the definition of a random mean generated by a random variable by presenting a general method for constructing such random means, and we also give some examples.

2.4 Theorem. *Let* $(\Omega,\mathcal{A},\mathbb{P})$ *be a probability space,* $\xi:\Omega\to\mathbb{R}^d$ *be a d-dimensional random variable and* $f\colon I^p\times\mathbb{R}^d\to I$ *be a* $(\mathcal{B}(I^p)\times\mathcal{B}(\mathbb{R}^d),\mathcal{B}(I))$*-measurable map such that for each* $y\in\mathbb{R}^d$, *the map* $I^p\ni x\mapsto f(x,y)$ *is in* $\mathcal{M}_p$. *Then the map* $M:\Omega\to\mathcal{M}_p$ *given by*

$$(M(\omega))(x):=f(x,\xi(\omega)),\qquad \omega\in\Omega,\ \ x\in I^p,$$

is a random mean generated by ξ.

Proof. By the construction, $M(\omega) \in \mathcal{M}_p$ for all $\omega \in \Omega$, and part (ii) of Definition 2.2 holds with the choice $M_\xi := f$.

Further, M can be written as $M = \varphi \circ \xi$, where $\varphi : \mathbb{R}^d \to \mathcal{M}_p$, $\varphi(y) := f(\cdot, y)$, $y \in \mathbb{R}^d$. Here ξ is $(\mathcal{A}, \mathcal{B}(\mathbb{R}^d))$-measurable and we check that φ is $(\mathcal{B}(\mathbb{R}^d), \mathcal{B}(\mathcal{C}(I^p))$-measurable. It is known that $\mathcal{B}(\mathcal{C}(I^p))$ coincides with the σ-algebra generated by $\mathcal{C}$, where $\mathcal{C}$ is the set of so-called cylinder sets of $\mathcal{C}(I^p)$ having the form

$$\big\{g \in \mathcal{C}(I^p) : (g(t_1), ..., g(t_n)) \in B_1 \times \cdots \times B_n\big\}, \qquad n \in \mathbb{N},\ t_1, \ldots, t_n \in I^p,\ B_1, \ldots, B_n \in \mathcal{B}(\mathbb{R}),$$

see, e.g., Kuo [10, Chapter I, Theorem 4.2]. So it is enough to check that for all $n \in \mathbb{N}$, $t_1, \ldots, t_n \in I^p$ and $B_1, \ldots, B_n \in \mathcal{B}(\mathbb{R})$, we have

$$\varphi^{-1}\Big(\big\{g \in \mathcal{C}(I^p) : (g(t_1), ..., g(t_n)) \in B_1 \times \cdots \times B_n\big\}\Big) \in \mathcal{B}(\mathbb{R}^d).$$

Here

$$\begin{aligned}
&\varphi^{-1}\Big(\big\{g \in \mathcal{C}(I^p) : (g(t_1), ..., g(t_n)) \in B_1 \times \cdots \times B_n\big\}\Big) \\
&= \Big\{y \in \mathbb{R}^d : f(\cdot, y) \in \mathcal{C}(I^p) \text{ and } (f(t_1, y), \ldots, f(t_n, y)) \in B_1 \times \cdots \times B_n\Big\} \\
&= \Big\{y \in \mathbb{R}^d : f(\cdot, y) \in \mathcal{C}(I^p)\Big\} \cap \bigcap_{i=1}^n f_{t_i}^{-1}(B_i),
\end{aligned}$$

where for each $i = 1, \ldots, n$, the function $f_{t_i} : \mathbb{R}^d \to I$, $f_{t_i}(y) := f(t_i, y)$, $y \in \mathbb{R}^d$, is a section of f on $\mathbb{R}^d$. By our assumptions,

$$\big\{y \in \mathbb{R}^d : f(\cdot, y) \in \mathcal{C}(I^p)\big\} = \big\{y \in \mathbb{R}^d : f(\cdot, y) \in \mathcal{M}_p\big\} = \mathbb{R}^d,$$

and f_{t_i}, $i = 1, \ldots, n$, is $(\mathcal{B}(\mathbb{R}^d), \mathcal{B}(I))$-measurable (see, e.g., Cohn [5, Lemma 5.1.2]) yielding that $f_{t_i}^{-1}(B_i) \in \mathcal{B}(\mathbb{R}^d)$, $i = 1, \ldots, n$. As a consequence, $\varphi^{-1}(C) \in \mathcal{B}(\mathbb{R}^d)$ for all $C \in \mathcal{C}$, as desired. Consequently, M is $(\mathcal{A}, \mathcal{B}(\mathcal{C}(I^p)))$-measurable yielding part (i) of Definition 2.2. $\Box$

2.5 Example. (Discrete random mean generated by a discrete random variable) Let ξ be a (one-dimensional) discrete random variable. In this example without loss of generality, we can and do assume that the range of ξ is $\mathbb{Z}_+$. Let us consider a sequence m_i, $i \in \mathbb{Z}_+$, in $\mathcal{M}_p$, and let $M : \Omega \to \mathcal{M}_p$,

$$(M(\omega))(x) := \sum_{i=0}^{\infty} m_i(x) \mathbb{1}_{\{i\}}(\xi(\omega)) = m_{\xi(\omega)}(x), \qquad \omega \in \Omega,\ x \in I^p.$$

We check that M is a random mean generated by ξ. First, note that $M(\omega) \in \mathcal{M}_p$ for all $\omega \in \Omega$, since ξ is $\mathbb{Z}_+$-valued and if $\omega \in \Omega$ is such that $\xi(\omega) = i$, where $i \in \mathbb{Z}_+$, then $(M(\omega))(x) = m_i(x)$, $x \in I^p$, and $m_i \in \mathcal{M}_p$.

Next, we check that M is $(\mathcal{A}, \mathcal{B}(\mathcal{C}(I^p)))$-measurable. For each $n \in \mathbb{N}$, let

$$\left(M^{(n)}(\omega)\right)(x) := \sum_{i=0}^{n} m_i(x) \mathbb{1}_{\{i\}}(\xi(\omega)) = \sum_{i=0}^{n} m_i(x) \mathbb{1}_{\xi^{-1}(\{i\})}(\omega), \qquad \omega \in \Omega.$$

Then $M^{(n)}$ is a simple $\mathcal{M}_p$-valued random variable (i.e., it has only finitely many values) being $(\mathcal{A}, \mathcal{B}(\mathcal{C}(I^p)))$-measurable, and for each $\omega \in \Omega$, we have

$$\sup_{x\in I^p} |(M^{(n)}(\omega))(x) - (M(\omega))(x)| = \sup_{x\in I^p} \left| \sum_{i=n+1}^{\infty} m_i(x) \mathbb{1}_{\{i\}}(\xi(\omega)) \right|$$

$$= \sup_{x\in I^p} \left| m_{\xi(\omega)}(x) \mathbb{1}_{[n+1,\infty)}(\xi(\omega)) \right| = \mathbb{1}_{[n+1,\infty)}(\xi(\omega)) \sup_{x\in I^p} |m_{\xi(\omega)}(x)| \to 0 \qquad \text{as } n \to \infty,$$

since $\sup_{x\in I^p} |m_{\xi(\omega)}(x)| < \infty$ (as we have seen in the proof of Proposition 2.1). This yields that M is a pointwise limit of $M^{(n)}$ as $n \to \infty$, so M is $(\mathcal{A}, \mathcal{B}(\mathcal{C}(I^p)))$-measurable as well. So part (i) of Definition 2.2 holds.

Let $M_\xi : I^p \times \mathbb{R} \to I$ be given by

$$M_\xi(x, y) := \sum_{i=0}^{\infty} m_i(x) \mathbb{1}_{\{i\}}(y), \qquad x \in I^p, \ y \in \mathbb{R}.$$

For each $n \in \mathbb{N}$, let $M_\xi^{(n)} : I^p \times \mathbb{R} \to I$ be given by

$$M_\xi^{(n)}(x, y) := \sum_{i=0}^{n} m_i(x) \mathbb{1}_{\{i\}}(y), \qquad x \in I^p, \ y \in \mathbb{R}.$$

Then $M_\xi^{(n)}$ is $(\mathcal{B}(I^p)\times\mathcal{B}(\mathbb{R}), \mathcal{B}(I))$-measurable, since m_i is continuous and $\{i\} \in \mathcal{B}(\mathbb{R})$ for each $i \in \mathbb{Z}_+$. Further, for all $x \in I^p$ and $y \in \mathbb{R}\setminus\mathbb{Z}_+$, we have $|M_\xi^{(n)}(x, y) - M_\xi(x, y)| = |0-0| = 0$, and for all $x \in I^p$ and $y \in \mathbb{Z}_+$, we have

$$|M_\xi^{(n)}(x, y) - M_\xi(x, y)| = \left| \sum_{i=n+1}^{\infty} m_i(x) \mathbb{1}_{\{i\}}(y) \right| = |m_y(x)| \mathbb{1}_{[n+1,\infty)}(y) \to 0 \qquad \text{as } n \to \infty.$$

So M_ξ is a pointwise limit of $M_\xi^{(n)}$ as $n \to \infty$, yielding that M_ξ is $(\mathcal{B}(I^p) \times \mathcal{B}(\mathbb{R}), \mathcal{B}(I))$-measurable. Moreover, for all $\omega \in \Omega$,

$$M_\xi(x, \xi(\omega)) = \sum_{i=0}^{\infty} m_i(x) \mathbb{1}_{\{i\}}(\xi(\omega)) = (M(\omega))(x), \qquad x \in I^p.$$

At the beginning of the example we have seen that $M(\omega) \in \mathcal{M}_p$ for all $\omega \in \Omega$. So, we get that part (ii) of Definition 2.2 holds as well.

We can call M a discrete random mean generated by the discrete random variable ξ in question, since the range of M contains countably many elements of $\mathcal{M}_p$. □

2.6 Example. (Random Hölder means) If $I := [a, b]$, where $0 < a < b < \infty$, then let $f : I^2 \times \mathbb{R} \times (0, 1) \to I$ be defined by

$$f(x_1, x_2, \alpha, \lambda) := \begin{cases} (\lambda x_1^\alpha + (1-\lambda)x_2^\alpha)^{\frac{1}{\alpha}}, & \text{if } \alpha \neq 0, \\ x_1^\lambda x_2^{1-\lambda}, & \text{if } \alpha = 0, \end{cases} \qquad (x_1, x_2, \alpha, \lambda) \in I^2 \times \mathbb{R} \times (0, 1),$$

and let $\xi\colon \Omega \to \mathbb{R} \times (0,1)$ be a random variable. If $I := [0,b]$, where $0 < b < \infty$, then let $f : I^2 \times \mathbb{R}_+ \times (0,1) \to I$ be defined by

$$f(x_1,x_2,\alpha,\lambda) := \begin{cases} (\lambda x_1^\alpha + (1-\lambda)x_2^\alpha)^{\frac{1}{\alpha}}, & \text{if } \alpha > 0, \\ x_1^\lambda x_2^{1-\lambda}, & \text{if } \alpha = 0, \end{cases} \qquad (x_1,x_2,\alpha,\lambda) \in I^2 \times \mathbb{R}_+ \times (0,1),$$

and let $\xi\colon \Omega \to \mathbb{R}_+ \times (0,1)$ be a random variable. Note that for each $(\alpha,\lambda) \in \mathbb{R} \times (0,1)$ in case of $I = [a,b]$ $(0 < a < b < \infty)$, and for each $(\alpha,\lambda) \in \mathbb{R}_+ \times (0,1)$ in case of $I = [0,b]$ $(0 < b < \infty)$, the map $I^2 \ni (x_1,x_2) \mapsto f(x_1,x_2,\alpha,\lambda)$ is a Hölder mean (also called weighted power mean), so it is in $\mathcal{M}_2$. Hence, using also that f is $(\mathcal{B}(I^2) \times \mathcal{B}(\mathbb{R}^2), \mathcal{B}(I))$-measurable, we can apply Theorem 2.4 and we have $M : \Omega \to \mathcal{M}_2$, $(M(\omega))(x_1,x_2) := f(x_1,x_2,\xi(\omega))$, $\omega \in \Omega$, $(x_1,x_2) \in I^2$, is a random mean generated by ξ, which can be called a random Hölder mean. One can define the p-variable version of this random mean in a similar way. □

The next proposition allows us to define the expected value of a random mean generated by a random variable ξ. The concept of Bochner integrability, integral of maps defined on a measure space with values in a Banach space, has a key role. We use the results and terminology of Cohn [5, Appendix E].

2.7 Proposition. *If $M\colon \Omega \to \mathcal{M}_p$ is $(\mathcal{A}, \mathcal{B}(\mathcal{C}(I^p)))$-measurable, then it is Bochner integrable.*

Proof. The function $\Omega \ni \omega \mapsto M(\omega)$ (considered as a function with values in $\mathcal{C}(I^p)$) is Bochner integrable if it is strongly measurable – i.e., M is $(\mathcal{A}, \mathcal{B}(\mathcal{C}(I^p)))$-measurable and has a separable range, where, by the range of M we mean the subset $M(\Omega)$ of $\mathcal{C}(I^p)$ – and the function $\Omega \ni \omega \mapsto \|M(\omega)\|$ is integrable with respect to $\mathbb{P}$, see Cohn [5, Appendix E].

Because of the Stone-Weierstrass approximation theorem, $\mathcal{C}(I^p)$ is separable, and since each subspace of a separable metric space is separable, we have the range of M is also separable. Hence the $(\mathcal{A}, \mathcal{B}(\mathcal{C}(I^p)))$-measurability of M implies that M is strongly measurable.

Moreover, using that $(M(\omega))(x_1,\ldots,x_p) = x_1$ whenever $x_1 = \cdots = x_p \in I$ and $\omega \in \Omega$ (due to $M(\omega) \in \mathcal{M}_p$, $\omega \in \Omega$), we have

$$\|M(\omega)\| = \sup_{x \in I^p} |(M(\omega))(x)| = \max_{t \in I} |t| < \infty, \qquad \omega \in \Omega, \tag{2.4}$$

since I is compact. So the function $\Omega \ni \omega \mapsto \|M(\omega)\|$ is the constant $\max_{t \in I} |t|$ function, and, using the fact that $\mathbb{P}(\Omega) = 1$, we have that it is integrable. □

2.8 Corollary. *If $M : \Omega \to \mathcal{M}_p$ is a random mean generated by a d-dimensional random variable ξ, then it is Bochner integrable.*

Proof. Since M is $(\mathcal{A}, \mathcal{B}(\mathcal{C}(I^p)))$-measurable (following from part (i) of Definition 2.2), Proposition 2.7 yields the statement. □

According to Corollary 2.8 the following definition does make sense.

2.9 Definition. *Let* $M : \Omega \to \mathcal{M}_p$ *be a random mean generated by a* d*-dimensional random variable* ξ *defined on a probability space* $(\Omega, \mathcal{A}, \mathbb{P})$*. Then the element of* $\mathcal{C}(I^p)$ *given by*

$$\mathbb{E}(M) := \int_\Omega M(\omega)\, \mathrm{d}\,\mathbb{P}(\omega)$$

is called the expected value or the expectation of M*.*

2.10 Theorem. *If* $M : \Omega \to \mathcal{M}_p$ *is a random mean generated by a* d*-dimensional random variable* ξ*, then its expected value* $\mathbb{E}(M)$ *is a non-random mean, that is to say,* $\mathbb{E}(M) \in \mathcal{M}_p$*. Further,*

$$(\mathbb{E}(M))(x) = \int_\Omega (M(\omega))(x)\, \mathrm{d}\,\mathbb{P}(\omega), \qquad x \in I^p.$$

Proof. By Corollary 2.8, $\mathbb{E}(M)$ exists, and, especially, $\mathbb{E}(M) \in \mathcal{C}(I^p)$. So, it remains to check that it is in $\mathcal{M}_p$.

It follows from Hytönen et al. [6, Proposition 1.2.12] that

$$\mathbb{E}(M) \in \overline{\mathrm{conv}\{M(\omega) : \omega \in \Omega\}},$$

where $\mathrm{conv}\{M(\omega) : \omega \in \Omega\}$ denotes the convex hull of $\{M(\omega) : \omega \in \Omega\}$, and $\overline{\mathrm{conv}\{M(\omega) : \omega \in \Omega\}}$ is its closure in $\mathcal{C}(I^p)$. Additionally, Proposition 2.1 implies

$$\overline{\mathrm{conv}\{M(\omega) : \omega \in \Omega\}} \subset \mathcal{M}_p.$$

These two gives that $\mathbb{E}(M) \in \mathcal{M}_p$.

Further, for each $x \in I^p$, let $\varphi_x : \mathcal{C}(I^p) \to \mathbb{R}$, $\varphi_x(h) := h(x)$, $h \in \mathcal{C}(I^p)$. Then for each $x \in I^p$, φ_x is a linear functional on $\mathcal{C}(I^p)$, and hence Proposition E.11 in Cohn [5] yields that

$$(\mathbb{E}(M))(x) = \varphi_x(\mathbb{E}(M)) = \int_\Omega \varphi_x(M(\omega))\, \mathrm{d}\,\mathbb{P}(\omega) = \int_\Omega (M(\omega))(x)\, \mathrm{d}\,\mathbb{P}(\omega), \tag{2.5}$$

as desired.

Using (2.5) one can give another (a more elementary) proof of the fact that $\mathbb{E}(M) \in \mathcal{M}_p$. Namely, since $M(\omega) \in \mathcal{M}_p, \omega \in \Omega$, we have $\min(x) \leqslant (M(\omega))(x) \leqslant \max(x), x \in I^p$, thus

$$\min(x) \leqslant \int_\Omega (M(\omega))(x)\, \mathrm{d}\,\mathbb{P}(\omega) \leqslant \max(x), \qquad x \in I^p,$$

and, by (2.5), we have $\min(x) \leqslant (\mathbb{E}(M))(x) \leqslant \max(x)$, $x \in I^p$, i.e., $\mathbb{E}(M) \in \mathcal{M}_p$. □

2.11 Remark. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, $\xi : \Omega \to \mathbb{R}^d$ be a d-dimensional random variable and $f : I^p \times \mathbb{R}^d \to I$ be a $(\mathcal{B}(I^p) \times \mathcal{B}(\mathbb{R}^d), \mathcal{B}(I))$-measurable map such that for each $y \in \mathbb{R}^d$, the map $I^p \ni x \mapsto f(x, y)$ is in $\mathcal{M}_p$. Then, by Theorem 2.4, $M : \Omega \to \mathcal{M}_p$ given by $(M(\omega))(x) := f(x, \xi(\omega))$, $\omega \in \Omega, x \in I^p$, is a random mean generated by ξ. Further, by Theorem 2.10,

$$(\mathbb{E}(M))(x) = \int_\Omega f(x, \xi(\omega))\, \mathrm{d}\,\mathbb{P}(\omega) = \int_{\mathbb{R}^d} f(x, y)\, \mathrm{d}\,\mathbb{P}_\xi(y), \qquad x \in I^p, \tag{2.6}$$

since the map $\mathbb{R}^d \ni y \mapsto f(x,y) \in I$ is $(\mathcal{B}(\mathbb{R}^d), \mathcal{B}(I))$-measurable for each fixed $x \in I^p$ and hence one can apply a result on integration with respect to an image measure (see, e.g., Cohn [5, Proposition 2.6.8]). Note also that in this case $\mathbb{E}(M)$ depends only on f and the distribution $\mathbb{P}_\xi$ of ξ. We do not know whether all the random means can be written in the form given in Theorem 2.4. □

Next we recall the notions of expectation (also called barycenter) of a random probability measure on a compact metric space (see, e.g., Borsato et al. [4, Appendix A.2]), and the Fréchet mean of a random measure with values in the 2-Wassertein space on $\mathbb{R}^d$ (see, e.g., Panaretos and Zemel [12, Section 3.2]) in order to see the similarities and differences compared to the expected value of a random mean generated by a random variable given in Definition 2.9.

2.12 Remark. First, we recall the expectation of a random probability measure on a compact metric space. Given a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and a compact metric space S endowed with its Borel σ-algebra $\mathcal{B}(S)$, a random probability measure on S is defined to be a Borel measurable map $\eta : \Omega \to \mathcal{P}_1(S)$, where $\mathcal{P}_1(S)$ denotes the set of probability measures on $(S, \mathcal{B}(S))$ and $\mathcal{P}_1(S)$ is endowed with the Borel σ-algebra corresponding to the topology of weak convergence according to which a sequence $(\mu_n)_{n\geqslant 1}$ in $\mathcal{P}_1(S)$ converges to a given $\mu \in \mathcal{P}_1(S)$ if $\int_S f(s)\,\mu_n(\mathrm{d}s) \to \int_S f(s)\,\mu(\mathrm{d}s)$ as $n \to \infty$ for each continuous (hence bounded) function $f : S \to \mathbb{R}$. Then, as a consequence of Riesz-Markov's representation theorem, there exists a unique element $\mathbb{E}(\eta)$ of $\mathcal{P}_1(S)$ such that the equality $\int_S f(s)\,(\mathbb{E}(\eta))(\mathrm{d}s) = \int_\Omega \int_S f(s)\,\eta^\omega(\mathrm{d}s)\,\mathbb{P}(\mathrm{d}\omega)$ holds for each continuous (hence bounded) function $f : S \to \mathbb{R}$, where η^ω denotes the value of the random measure η at the point $\omega \in \Omega$, see, e.g., Borsato et al. [4, Theorem A.6 and Definition A.7].

Next, we recall the Fréchet mean of a random measure with values in the 2-Wassertein space on $\mathbb{R}^d$. Given a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and $d \in \mathbb{N}$, the 2-Wasserstein space on $\mathbb{R}^d$ is defined by

$$\mathcal{W}_2(\mathbb{R}^d) := \Big\{\mu \in \mathcal{P}_1(\mathbb{R}^d) : \int_{\mathbb{R}^d} \|\boldsymbol{x}\|^2 \mu(\mathrm{d}\boldsymbol{x}) < \infty\Big\},$$

where $\mathcal{P}_1(\mathbb{R}^d)$ denotes the set of probability measures on $\mathbb{R}^d$. For $\mu, \nu \in \mathcal{P}_1(\mathbb{R}^d)$, let $\Pi(\mu,\nu)$ be the set of probability measures $\pi \in \mathcal{P}_1(\mathbb{R}^d \times \mathbb{R}^d)$ such that $\pi(A \times \mathbb{R}^d) = \mu(A)$, $A \in \mathcal{B}(\mathbb{R}^d)$, and $\pi(\mathbb{R}^d \times B) = \nu(B)$, $B \in \mathcal{B}(\mathbb{R}^d)$, i.e., μ and ν are the marginals of π. The 2-Wasserstein distance between μ and ν is defined as

$$W_2(\mu,\nu) := \left(\inf_{\pi \in \Pi(\mu,\nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} \|\boldsymbol{x}_1 - \boldsymbol{x}_2\|^2 \,\mathrm{d}\pi(\boldsymbol{x}_1, \boldsymbol{x}_2)\right)^{\frac{1}{2}}, \qquad \mu, \nu \in \mathcal{P}_1(\mathbb{R}^d).$$

Then W_2 is a metric on $\mathcal{W}_2(\mathbb{R}^d)$, see Villani [13, Chapter 7]. By a random measure with values in $\mathcal{W}_2(\mathbb{R}^d)$, we mean a measurable map $\Lambda : \Omega \to \mathcal{W}_2(\mathbb{R}^d)$, where $\mathcal{W}_2(\mathbb{R}^d)$ is endowed with its Borel σ-algebra (corresponding to the metric W_2). By the Fréchet mean (expectation) of a random measure Λ with values in $\mathcal{W}_2(\mathbb{R}^d)$, we mean the minimizer (if it is unique) of the Fréchet functional

$$F(\gamma) := \frac{1}{2}\,\mathbb{E}((W_2(\gamma, \Lambda))^2), \qquad \gamma \in \mathcal{W}_2(\mathbb{R}^d),$$

see, e.g., Definition 3.2.1 in Panaretos and Zemel [12]. We note that the Fréchet functional associated with any random measure Λ with values in $\mathcal{W}_2(\mathbb{R}^d)$ admits a minimizer (see, e.g., Panaretos and Zemel [12, Proposition 3.2.3]), and for a result on the uniqueness of Fréchet means, see, e.g., Proposition 3.2.7 in Panaretos and Zemel [12]. Further, see Remark 3.6 for a comparison of our forthcoming limit theorems for random means generated by random variables with the Wasserstein law of large numbers for Fréchet means (Panaretos and Zemel [12, Corollary 3.2.10]). □

Next, we determine the expectation of the random means given in Examples 2.5 and 2.6 (in case of Example 2.6 with special choices of ξ).

2.13 Example. The expectation of the random mean generated by a discrete random variable ξ having range in $\mathbb{Z}_+$ given in Example 2.5 takes the form

$$\mathbb{E}(M) = \int_\Omega M(\omega)\,\mathrm{d}\,\mathbb{P}(\omega) = \lim_{n\to\infty}\int_\Omega M^{(n)}(\omega)\,\mathrm{d}\,\mathbb{P}(\omega) = \lim_{n\to\infty}\sum_{i=0}^{n} m_i\,\mathbb{P}(\xi=i) = \sum_{i=0}^{\infty}\mathbb{P}(\xi=i)m_i,$$

where $M^{(n)}$ is introduced in Example 2.5, the series above converges in $\mathcal{C}(I^p)$, and for the second equality we used the construction of Bochner integral (see, e.g., Cohn [5, Appendix E]). □

2.14 Example. (Expectation of some random Hölder means) Let us consider the random Hölder mean M given in Example 2.6 generated by a random variable ξ. First, let us suppose that the distribution $\mathbb{P}_\xi$ of ξ takes the form $\mathbb{P}_\xi = \delta_{\alpha_0}\otimes\mathbb{P}_U$, where $\alpha_0\in\mathbb{R}_+$, δ_{α_0} denotes the Dirac measure concentrated at α_0, and U is a uniformly distributed random variable in the interval $(0,1)$. In Example 2.6, let us choose $I := [0,b]$, where $b>0$.
In case of $\alpha_0\in(0,\infty)$, for the expectation $\mathbb{E}(M)\in\mathcal{M}_2$ of M, we have

$$\begin{aligned}(\mathbb{E}(M))(x_1,x_2) &= \int_{\mathbb{R}^2} f(x_1,x_2,\alpha,\lambda)\;\mathbb{P}_\xi(\mathrm{d}\alpha,\mathrm{d}\lambda) = \int_0^1 (\lambda x_1^{\alpha_0} + (1-\lambda)x_2^{\alpha_0})^{\frac{1}{\alpha_0}}\,\mathrm{d}\lambda\\ &= \int_0^1 \big((x_1^{\alpha_0}-x_2^{\alpha_0})\lambda + x_2^{\alpha_0}\big)^{\frac{1}{\alpha_0}}\,\mathrm{d}\lambda\\ &= \begin{cases}\frac{x_1^{\alpha_0+1}-x_2^{\alpha_0+1}}{(\frac{1}{\alpha_0}+1)(x_1^{\alpha_0}-x_2^{\alpha_0})} & \text{if } x_1\neq x_2,\\ x_1 & \text{if } x_1 = x_2,\end{cases}\qquad x_1,x_2\in I,\end{aligned}$$

where the first equality follows by (2.6). In this case one can check that $\mathbb{E}(M)$ is nothing else but a Cauchy mean corresponding to the power functions x^{α_0+1}, $x\in I$, and x^{α_0}, $x\in I$, see, e.g., Beliakov et al. [3, Definition 2.50] or Jarczyk and Jarczyk [7, Section 5.1].

In case of $\alpha_0=0$, for the expectation $\mathbb{E}(M)\in\mathcal{M}_2$ of M, by (2.6), we have

$$(\mathbb{E}(M))(x_1,x_2) = \int_0^1 x_1^\lambda x_2^{1-\lambda}\,\mathrm{d}\lambda = \begin{cases}\frac{x_1-x_2}{\ln(x_1)-\ln(x_2)} & \text{if } x_1\neq x_2,\ x_1,x_2\in I\cap(0,\infty),\\ x_1 & \text{if } x_1=x_2\in I\cap(0,\infty),\\ 0 & \text{if } x_1=0 \text{ or } x_2=0.\end{cases}$$

In this case, $\mathbb{E}(M)$ restricted to $(0,\infty)\times(0,\infty)$ is nothing else but the logarithmic mean, see, e.g., Beliakov et al. [3, Definition 2.45] or Jarczyk and Jarczyk [7, Section 5.1].

Next, let us suppose that the distribution $\mathbb{P}_\xi$ of ξ takes the form $\mathbb{P}_\xi = \delta_{\alpha_0}\otimes\mathbb{P}_V$, where $\alpha_0\in\mathbb{R}_+$ and V is a random variable with density function $f_V(\lambda) = 2\lambda\mathbf{1}_{(0,1)}(\lambda)$, $\lambda\in\mathbb{R}$. As before, in Example 2.6, let us choose $I := [0,b]$, where $b>0$.
In case of $\alpha_0\in(0,\infty)$, for the expectation $\mathbb{E}(M)\in\mathcal{M}_2$ of M, by (2.6), we have

$$
\begin{aligned}
(\mathbb{E}(M))(x_1,x_2) &= \int_0^1 \left((x_1^{\alpha_0}-x_2^{\alpha_0})\lambda + x_2^{\alpha_0}\right)^{\frac{1}{\alpha_0}} 2\lambda\,\mathrm{d}\lambda \\
&= \frac{2}{x_1^{\alpha_0}-x_2^{\alpha_0}}\int_0^1 \left((x_1^{\alpha_0}-x_2^{\alpha_0})\lambda + x_2^{\alpha_0}\right)^{\frac{1}{\alpha_0}+1}\mathrm{d}\lambda - \frac{2x_2^{\alpha_0}}{x_1^{\alpha_0}-x_2^{\alpha_0}}\int_0^1 \left((x_1^{\alpha_0}-x_2^{\alpha_0})\lambda + x_2^{\alpha_0}\right)^{\frac{1}{\alpha_0}}\mathrm{d}\lambda \\
&= \frac{2}{\frac{1}{\alpha_0}+2}\frac{(x_1^{2\alpha_0+1}-x_2^{2\alpha_0+1})}{(x_1^{\alpha_0}-x_2^{\alpha_0})^2} - \frac{2}{\frac{1}{\alpha_0}+1}\frac{x_2^{\alpha_0}(x_1^{\alpha_0+1}-x_2^{\alpha_0+1})}{(x_1^{\alpha_0}-x_2^{\alpha_0})^2} \\
&= \frac{2}{\left(\frac{1}{\alpha_0}+1\right)\left(\frac{1}{\alpha_0}+2\right)}\cdot\frac{\left(\frac{1}{\alpha_0}+1\right)x_1^{2\alpha_0+1} - \left(\frac{1}{\alpha_0}+2\right)x_1^{\alpha_0+1}x_2^{\alpha_0} + x_2^{2\alpha_0+1}}{(x_1^{\alpha_0}-x_2^{\alpha_0})^2}
\end{aligned}
$$

for $x_1\neq x_2$, $x_1,x_2\in I$, and $(\mathbb{E}(M))(x_1,x_2) = x_1$ for $x_1=x_2\in I$.

In case of $\alpha_0=0$, for the expectation $\mathbb{E}(M)\in\mathcal{M}_2$ of M, by (2.6) and partial integration, we have

$$
(\mathbb{E}(M))(x_1,x_2) = \int_0^1 x_1^\lambda x_2^{1-\lambda}2\lambda\,\mathrm{d}\lambda = 2x_2\int_0^1 \lambda\left(\frac{x_1}{x_2}\right)^\lambda\mathrm{d}\lambda = \frac{2(x_1\ln(x_1)-x_1-x_1\ln(x_2)+x_2)}{(\ln(x_1)-\ln(x_2))^2}
$$

for $x_1\neq x_2$, $x_1,x_2\in I\cap(0,\infty)$, and $(\mathbb{E}(M))(x_1,x_2)=x_1$ for $x_1=x_2\in I\cap(0,\infty)$. If $x_1=0$ or $x_2=0$, then $(\mathbb{E}(M))(x_1,x_2)=0$.

Next, let us suppose that the distribution $\mathbb{P}_\xi$ of ξ takes the form $\mathbb{P}_\xi=\delta_0\otimes\mathbb{P}_W$, where W is a random variable with density function $f_W(\lambda)=\frac{\mathrm{e}}{\mathrm{e}-1}\mathrm{e}^{-\lambda}\mathbf{1}_{(0,1)}(\lambda)$, $\lambda\in\mathbb{R}$. In Example 2.6, let us choose $I:=[0,b]$, where $b>0$. For the expectation $\mathbb{E}(M)\in\mathcal{M}_2$ of M, by (2.6), we have

$$
(\mathbb{E}(M))(x_1,x_2) = \frac{\mathrm{e}}{\mathrm{e}-1}\int_0^1 x_1^\lambda x_2^{1-\lambda}\mathrm{e}^{-\lambda}\,\mathrm{d}\lambda = \frac{\mathrm{e}}{\mathrm{e}-1}x_2\int_0^1\left(\frac{x_1}{\mathrm{e}x_2}\right)^\lambda\mathrm{d}\lambda = \frac{1}{\mathrm{e}-1}\cdot\frac{x_1-\mathrm{e}x_2}{\ln(x_1)-\ln(\mathrm{e}x_2)}
$$

for $x_1\neq \mathrm{e}x_2$, $x_1,x_2\in I\cap(0,\infty)$, and $(\mathbb{E}(M))(x_1,x_2)=\frac{1}{\mathrm{e}-1}x_1$ for $x_1=\mathrm{e}x_2$, $x_1,x_2\in I\cap(0,\infty)$. If $x_1=0$ or $x_2=0$, then $(\mathbb{E}(M))(x_1,x_2)=0$. Note that the restriction of $\mathbb{E}(M)$ onto $(0,\infty)\times(0,\infty)$ can be considered as a variant of the logarithmic mean. Namely,

$$
(\mathbb{E}(M))(x_1,x_2) = \mathcal{L}\left(\frac{x_1}{\mathrm{e}-1},\frac{\mathrm{e}x_2}{\mathrm{e}-1}\right),\qquad x_1,x_2\in I\cap(0,\infty)=(0,b],
$$

where $\mathcal{L}$ denotes the logarithmic mean. However, note also that the mapping $I^2\ni(x_1,x_2)\mapsto(\mathbb{E}(M))(x_1,x_2)$ is a mean on its own right, following from Theorem 2.10, or it can be also checked directly. Indeed, using that the function $(0,1] \ni v \mapsto (\frac{\mathrm{e}}{v}-1)/\ln(\frac{\mathrm{e}}{v})$ is monotone decreasing, $(0,1] \ni v \mapsto (v-\mathrm{e})/\ln(\frac{v}{\mathrm{e}})$ is monotone increasing, and that their value at 1 is $\mathrm{e}-1$, we have

$$1 \leqslant \frac{1}{\mathrm{e}-1} \cdot \frac{\frac{\mathrm{e}}{v}-1}{\ln\left(\frac{\mathrm{e}}{v}\right)}, \qquad v \in (0,1) \qquad \text{and} \qquad \frac{1}{\mathrm{e}-1} \cdot \frac{v-\mathrm{e}}{\ln\left(\frac{v}{\mathrm{e}}\right)} \leqslant 1, \qquad v \in (0,1),$$

and hence in case of $0 < x_1 < x_2 \leqslant b$, by choosing $v := \frac{x_1}{x_2}$, we have

$$x_1 \leqslant \frac{1}{\mathrm{e}-1} \cdot \frac{x_1 - \mathrm{e}x_2}{\ln(x_1) - \ln(\mathrm{e}x_2)} \leqslant x_2,$$

as desired. Further, if $x_1 = \mathrm{e}x_2$, $x_1, x_2 \in (0, b]$, then

$$x_2 = \min(x_1, x_2) \leqslant \frac{1}{\mathrm{e}-1} x_1 \leqslant x_1 = \max(x_1, x_2),$$

as desired.

Next, let us suppose that the distribution $\mathbb{P}_\xi$ of ξ takes the form $\mathbb{P}_\xi = \delta_0 \otimes \mathbb{P}_X$, where X is a random variable with density function $f_X(\lambda) = \frac{1}{1-\cos(1)} \sin(\lambda) \mathbf{1}_{(0,1)}(\lambda)$, $\lambda \in \mathbb{R}$. In Example 2.6, let us choose $I := [0, b]$, where $b > 0$. For the expectation $\mathbb{E}(M) \in \mathcal{M}_2$ of M, by (2.6) and partial integration, we have

$$\begin{aligned}
(\mathbb{E}(M))(x_1, x_2) &= \frac{1}{1-\cos(1)} \int_0^1 x_1^\lambda x_2^{1-\lambda} \sin(\lambda)\,\mathrm{d}\lambda = \frac{x_2}{1-\cos(1)} \int_0^1 \left(\frac{x_1}{x_2}\right)^\lambda \sin(\lambda)\,\mathrm{d}\lambda \\
&= \frac{x_2}{1-\cos(1)} \left(1 - \cos(1)\frac{x_1}{x_2} + \ln\left(\frac{x_1}{x_2}\right) \int_0^1 \left(\frac{x_1}{x_2}\right)^\lambda \cos(\lambda)\,\mathrm{d}\lambda\right) \\
&= \frac{1}{1-\cos(1)} \left(x_2 - \cos(1)x_1 + \sin(1)x_1 \ln\left(\frac{x_1}{x_2}\right) - x_2 \left(\ln\left(\frac{x_1}{x_2}\right)\right)^2 \int_0^1 \left(\frac{x_1}{x_2}\right)^\lambda \sin(\lambda)\,\mathrm{d}\lambda\right)
\end{aligned}$$

for $x_1, x_2 \in I \cap (0, \infty)$. Consequently, for $x_1, x_2 \in I \cap (0, \infty)$, we have

$$(\mathbb{E}(M))(x_1, x_2) = \frac{1}{1-\cos(1)} \frac{x_2 - \cos(1)x_1 + \sin(1)x_1(\ln(x_1) - \ln(x_2))}{1 + (\ln(x_1) - \ln(x_2))^2}.$$

If $x_1 = 0$ or $x_2 = 0$, then $(\mathbb{E}(M))(x_1, x_2) = 0$. □

Motivated by Examples 2.6 and 2.14, in the next remark we initiate some possible future research directions.

2.15 Remark. (i) Is it possible to give a set of random variables with values in $\mathbb{R} \times (0,1)$ such that the set of expectations of the corresponding random Hölder means given in Example 2.6 coincide with $\mathcal{M}_p$? If the answer is yes, then characterize such a set of random variables. If the answer is no, then characterize the largest subset of $\mathcal{M}_p$, which can be achieved in this way. One can pose a similar question concerning any other random mean.

(ii) Moreover, given a usual (non-random) mean $m \in \mathcal{M}_p$ on I, let us characterize (possibly under some additional assumptions) those $p \times p$ matrices A with real entries such that the mapping $I^p \ni (x_1, \ldots, x_p) \mapsto m((x_1, \ldots, x_p)A)$ is a (usual) mean. Of course, if A is a $p \times p$ permutation matrix, then this property holds. Further, in Example 2.14, we have showed that the mapping $(0, b]^2 \ni (x_1, x_2) \mapsto \mathcal{L}((x_1, x_2)A)$ is a mean, where $b > 0$, $\mathcal{L}$ denotes the logarithmic mean and A is the 2×2 diagonal matrix with $(1,1)$-entry $\frac{1}{\mathrm{e}-1}$ and $(2,2)$-entry $\frac{\mathrm{e}}{\mathrm{e}-1}$ (not being a permutation matrix). □

3 Limit theorems for random means

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, I be a nondegenerate, compact interval of $\mathbb{R}$, and $d, p \in \mathbb{N}$. Let $(\xi_n)_{n\in\mathbb{N}}$ be a sequence of independent and identically distributed d-dimensional random variables, and for each $n \in \mathbb{N}$, let $M_n : \Omega \to \mathcal{M}_p$ be a random mean generated by ξ_n (in the sense of Definition 2.2). For each $n \in \mathbb{N}$, let $\overline{S}_n : \Omega \to \mathcal{M}_p$,

$$\overline{S}_n(\omega) := \frac{1}{n}\sum_{j=1}^{n} M_j(\omega), \qquad \omega \in \Omega.$$

Then, by Definition 2.2, for each $n \in \mathbb{N}$ and $\omega \in \Omega$, we have $\overline{S}_n(\omega) \in \mathcal{M}_p$, i.e., it is a (usual) p-variable mean, and for each $n \in \mathbb{N}$, the mapping $\Omega \ni \omega \mapsto \overline{S}_n(\omega)$ is an $\mathcal{M}_p$-valued random variable.

In what follows we are searching for sufficient conditions on the random means M_n, $n \in \mathbb{N}$, under which

$$\mathbb{P}\left(\left\{\omega \in \Omega : \lim_{n\to\infty} \varrho(\overline{S}_n(\omega), \mathbb{E}(M_1)) = 0\right\}\right) = 1$$

holds, where the metric ϱ is given in (2.3), and the law of the random variable

$$\Omega \ni \omega \mapsto \sqrt{n}\,\varrho(\overline{S}_n(\omega), \mathbb{E}(M_1)) =: \sqrt{n}\,\kappa_n(\omega) \tag{3.1}$$

converges in distribution to some normal distribution as $n \to \infty$. Here note that for each $n \in \mathbb{N}$, κ_n is indeed a real-valued random variable, since $\Omega \ni \omega \mapsto \overline{S}_n(\omega)$ is an $\mathcal{M}_p$-valued random variable and the metric ϱ is continuous.

3.1 Theorem. *Let $d, k \in \mathbb{N}$, and I be a nondegenerate, compact interval of $\mathbb{R}$. Let $(\xi_n)_{n\in\mathbb{N}}$ be a sequence of independent and identically distributed d-dimensional discrete random variables having finite range $\{a_1, \ldots, a_k\}$, where $a_1, \ldots, a_k \in \mathbb{R}^d$ are pairwise distinct. Let $q_i := \mathbb{P}(\xi_1 = a_i) \in (0, 1)$, $i = 1, \ldots, k$. Let $p \in \mathbb{N}$ and for each $n \in \mathbb{N}$ let $M_n : \Omega \to \mathcal{M}_p$,*

$$(M_n(\omega))(x_1, \ldots, x_p) := \sum_{i=1}^{k} m_i(x_1, \ldots, x_p)\mathbf{1}_{\{\xi_n(\omega)=a_i\}}, \qquad \omega \in \Omega,\ \ x_1, \ldots, x_p \in I,$$

where $m_i \in \mathcal{M}_p$, $i = 1, \ldots, k$. Then for each $n \in \mathbb{N}$, M_n is a random mean generated by ξ_n, and

$$\mathbb{P}\left(\left\{\omega \in \Omega : \lim_{n\to\infty} \kappa_n(\omega) = 0\right\}\right) = 1, \tag{3.2}$$

where κ_n *is given in* (3.1) *with* $(\mathbb{E}(M_1))(x_1,\ldots,x_p) = \sum_{i=1}^k m_i(x_1,\ldots,x_p)\,\mathbb{P}(\xi_1 = a_i)$ *for* $x_1,\ldots,x_p \in I$.

Further,

$$\sqrt{n}\,\kappa_n \stackrel{\mathcal{D}}{\longrightarrow} \sup_{x_1,\ldots,x_p\in I} \big|\langle \mathcal{N}_k(\mathbf{0},\boldsymbol{Q}), \boldsymbol{m}(x_1,\ldots,x_p)\rangle\big| \tag{3.3}$$

as $n\to\infty$, *where* $\boldsymbol{Q} := (q_{i,j})_{i,j=1}^k \in \mathbb{R}^{k\times k}$ *is the* $k\times k$ *matrix given by*

$$q_{i,j} := \begin{cases} q_i(1-q_i) & \text{if } i=j,\\ -q_iq_j & \text{if } i\neq j,\end{cases}$$

and $\boldsymbol{m}(x_1,\ldots,x_p) := (m_1(x_1,\ldots,x_p),\ldots,m_k(x_1,\ldots,x_p))^\top \in \mathbb{R}^k$.

Here $\boldsymbol{Q}$ is nothing else but the covariance matrix of $(\mathbb{1}_{\{\xi_1=a_1\}},\ldots,\mathbb{1}_{\{\xi_1=a_k\}})^\top$ having multinomial distribution with parameters 1 and $q_1,\ldots,q_k$.

Next, we consider a special case of Theorem 3.1, namely, when $p=2$, ξ_1 has a Bernoulli distribution and the range of M_1 is the set consisting of the arithmetic and geometric means in $[0,1]$.

3.2 Corollary. *Let* $I := [0,1]$, *and* $(\xi_n)_{n\in\mathbb{N}}$ *be a sequence of independent and identically distributed random variables such that* $\mathbb{P}(\xi_1 = 0) = q$ *and* $\mathbb{P}(\xi_1=1) = 1-q$, *where* $q\in(0,1)$, *i.e.,* ξ_1 *is Bernoulli distributed with parameter* q. *For each* $n\in\mathbb{N}$, *let* $M_n : \Omega \to \mathcal{M}_2$,

$$(M_n(\omega))(x_1,x_2) := m_0(x_1,x_2)\mathbf{1}_{\{\xi_n(\omega)=0\}} + m_1(x_1,x_2)\mathbf{1}_{\{\xi_n(\omega)=1\}}, \qquad \omega\in\Omega,\ \ x_1,x_2\in I,$$

where

$$m_0(x_1,x_2) := \frac{x_1+x_2}{2} \qquad \text{and} \qquad m_1(x_1,x_2) := \sqrt{x_1x_2}$$

for $x_1,x_2\in I$. *Then for each* $n\in\mathbb{N}$, M_n *is a random mean generated by* ξ_n, *and*

$$\mathbb{P}\left(\left\{\omega\in\Omega : \lim_{n\to\infty}\kappa_n(\omega) = 0\right\}\right) = 1, \tag{3.4}$$

where κ_n *is given in* (3.1) *with* $(\mathbb{E}(M_1))(x_1,x_2) = \frac{x_1+x_2}{2}\,q + \sqrt{x_1x_2}\,(1-q)$, $x_1,x_2\in I$. *Further,*

$$\sqrt{n}\,\kappa_n \stackrel{\mathcal{D}}{\longrightarrow} \sup_{x_1,x_2\in[0,1]} \left(\frac{x_1+x_2}{2} - \sqrt{x_1x_2}\right)\cdot|\mathcal{N}(0,q(1-q))| = \frac{1}{2}\,|\mathcal{N}(0,q(1-q))| \tag{3.5}$$

as $n\to\infty$.

Next, we establish limit theorems for randomly weighted arithmetic means, where ξ_1 is not necessarily discrete, so our next result is out of scope of Theorem 3.1.

3.3 Theorem. *Let* $p\geqslant 2$, $p\in\mathbb{N}$, I *be a nondegenerate, compact interval of* $\mathbb{R}$, *and* $(\boldsymbol{\xi}_n := (\xi_n^{(1)},\ldots,\xi_n^{(p-1)}))_{n\in\mathbb{N}}$ *be a sequence of independent and identically distributed* $\mathbb{R}_+^{p-1}$-*valued random variables such that* $\mathbb{P}(\xi_1^{(1)}+\cdots+\xi_1^{(p-1)} \leqslant 1) = 1$. *For each* $n\in\mathbb{N}$, *let* $M_n:\Omega\to\mathcal{M}_p$,

$$(M_n(\omega))(x_1,\ldots,x_p) := \xi_n^{(1)}(\omega)x_1 + \cdots + \xi_n^{(p-1)}(\omega)x_{p-1} + \big(1-\xi_n^{(1)}(\omega) - \cdots - \xi_n^{(p-1)}(\omega)\big)x_p$$

for $\omega \in \Omega$ *and* $x_1, \ldots, x_p \in I$. *Then for each* $n \in \mathbb{N}$, M_n *is a random mean generated by* $\boldsymbol{\xi}_n$, *and*

$$\mathbb{P}\left(\left\{\omega \in \Omega : \lim_{n\to\infty} \kappa_n(\omega) = 0\right\}\right) = 1, \tag{3.6}$$

where κ_n *is given in* (3.1) *with*

$$(\mathbb{E}(M_1))(x_1, \ldots, x_p) = \sum_{i=1}^{p-1} x_i \, \mathbb{E}(\xi_1^{(i)}) + x_p \left(1 - \sum_{i=1}^{p-1} \mathbb{E}(\xi_1^{(i)})\right), \qquad x_1, \ldots, x_p \in I.$$

Further,

$$\sqrt{n}\, \kappa_n \stackrel{\mathcal{D}}{\longrightarrow} \sup_{x_1,\ldots,x_p \in I} \left| \left\langle \mathcal{N}_{p-1}\left(\mathbf{0}, \operatorname{Cov}(\boldsymbol{\xi}_n)\right), \begin{pmatrix} x_1 - x_p \\ \vdots \\ x_{p-1} - x_p \end{pmatrix} \right\rangle \right| \tag{3.7}$$

as $n \to \infty$, *where* $\operatorname{Cov}(\boldsymbol{\xi}_n)$ *denotes the covariance matrix of* $\boldsymbol{\xi}_n$.

Next, we formulate a corollary of Theorem 3.3 in case of $p = 2$ by simplifying the limit distribution in (3.7).

3.4 Corollary. *Let* I *be a nondegenerate, compact interval of* $\mathbb{R}$, *and* $(\xi_n)_{n\in\mathbb{N}}$ *be a sequence of independent and identically distributed random variables such that* $\mathbb{P}(\xi_1 \in [0,1]) = 1$. *For each* $n \in \mathbb{N}$, *let* $M_n : \Omega \to \mathcal{M}_2$, $(M_n(\omega))(x_1, x_2) := \xi_n(\omega)x_1 + (1 - \xi_n(\omega))x_2$, $\omega \in \Omega$, $x_1, x_2 \in I$. *Then for each* $n \in \mathbb{N}$, M_n *is a random mean generated by* ξ_n, *and*

$$\mathbb{P}\left(\left\{\omega \in \Omega : \lim_{n\to\infty} \kappa_n(\omega) = 0\right\}\right) = 1, \tag{3.8}$$

where κ_n *is given in* (3.1) *with* $(\mathbb{E}(M_1))(x_1, x_2) = x_1 \, \mathbb{E}(\xi_1) + x_2(1 - \mathbb{E}(\xi_1))$, $x_1, x_2 \in I$. *Further,*

$$\sqrt{n}\, \kappa_n \stackrel{\mathcal{D}}{\longrightarrow} \sup_{x_1,x_2 \in I} |x_1 - x_2| \cdot \left|\mathcal{N}\left(0, \mathbb{D}^2(\xi_1)\right)\right| \tag{3.9}$$

as $n \to \infty$.

Finally, we provide limit theorems for randomly weighted power means (which can be also called random Hölder means, see Example 2.6). We point out to the facts that in this case instead of the arithmetic mean of the given random means we consider their geometric mean, and so the limit theorems have somewhat different forms compared to the previous ones.

3.5 Theorem. *Let* I *be a nondegenerate, compact interval of* $(0, \infty)$, *and* $(\xi_n)_{n\in\mathbb{N}}$ *be a sequence of independent and identically distributed random variables such that* $\mathbb{P}(\xi_1 \in [0,1]) =$

1. *For each* $n \in \mathbb{N}$, *let* $M_n : \Omega \to \mathcal{M}_2$, $(M_n(\omega))(x_1, x_2) := x_1^{\xi_n(\omega)} x_2^{1-\xi_n(\omega)}$, $\omega \in \Omega$, $x_1, x_2 \in I$. *Then for each* $n \in \mathbb{N}$, M_n *is a random mean generated by* ξ_n, *and*

$$\mathbb{P}\left(\left\{\omega \in \Omega : \lim_{n\to\infty} \sup_{x_1,x_2\in I} \frac{\left(\prod_{j=1}^n (M_j(\omega))(x_1, x_2)\right)^{\frac{1}{n}}}{x_1^{\mathbb{E}(\xi_1)} x_2^{1-\mathbb{E}(\xi_1)}} = 1\right\}\right) = 1. \tag{3.10}$$

Further,

$$\left(\sup_{x_1,x_2\in I} \frac{\left(\prod_{j=1}^n (M_j(\cdot))(x_1, x_2)\right)^{\frac{1}{n}}}{x_1^{\mathbb{E}(\xi_1)} x_2^{1-\mathbb{E}(\xi_1)}}\right)^{\sqrt{n}} \xrightarrow{\mathcal{D}} \left(\frac{\max(I)}{\min(I)}\right)^{\left|\mathcal{N}\left(0,\mathbb{D}^2(\xi_1)\right)\right|} \tag{3.11}$$

as $n \to \infty$, *where for any* $j \in \mathbb{N}$ *and* $x_1, x_2 \in I$, $(M_j(\cdot))(x_1, x_2)$ *denotes the random variable* $\Omega \ni \omega \mapsto (M_j(\omega))(x_1, x_2)$, *and* $\max(I) := \max\{x : x \in I\}$ *and* $\min(I) := \min\{x : x \in I\}$.

3.6 Remark. Using the notations of the second part of Remark 2.12, we note that a Wasserstein law of large numbers holds for a sequence $(\Lambda_n)_{n\in\mathbb{N}}$ of independent and identically distributed random means with values in $\mathcal{W}_2(\mathbb{R}^d)$ having unique Fréchet means, namely, the so-called empirical Fréchet mean of $(\Lambda_1, \ldots, \Lambda_n)$ (see Panaretos and Zemel [12, Definition 3.1.1]) converges almost surely to the Fréchet mean of Λ_1 as $n \to \infty$, see Corollary 3.2.10 in Panaretos and Zemel [12]. Note that in present section, we derived different kinds of limit theorems for random means generated by a sequence of independent and identically distributed random variables, since our limit theorems are about the random means itself and not about their expectations. □

4 Proofs for Section 3

Proof of Theorem 3.1. By Example 2.5, for each $n \in \mathbb{N}$, M_n is a random mean generated by ξ_n. For all $x_1, \ldots, x_p \in I$, we have

$$(\overline{S}_n(\omega))(x_1, \ldots, x_p) = \frac{1}{n}\sum_{j=1}^n \mathbb{1}_{\{\xi_j(\omega)=a_1\}}\, m_1(x_1, \ldots, x_p) + \cdots + \frac{1}{n}\sum_{j=1}^n \mathbb{1}_{\{\xi_j(\omega)=a_k\}}\, m_k(x_1, \ldots, x_p),$$

and, by Example 2.13,

$$(\mathbb{E}(M_1))(x_1, \ldots, x_p) = \sum_{i=1}^k m_i(x_1, \ldots, x_p)\, \mathbb{P}(\xi_1 = a_i), \qquad x_1, \ldots, x_p \in I.$$

Hence

$$\kappa_n(\omega) = \varrho(\overline{S}_n(\omega), \mathbb{E}(M_1)) = \sup_{(x_1,\ldots,x_p)\in I^p} \left|(\overline{S}_n(\omega))(x_1, \ldots, x_p) - (\mathbb{E}(M_1))(x_1, \ldots, x_p)\right| =$$

$$= \sup_{(x_1,\dots,x_p)\in I^p} \Bigg| \left(\frac{1}{n}\sum_{j=1}^n \mathbb{1}_{\{\xi_j(\omega)=a_1\}} - \mathbb{P}(\xi_1 = a_1)\right) m_1(x_1,\dots,x_p)$$

$$+ \cdots + \left(\frac{1}{n}\sum_{j=1}^n \mathbb{1}_{\{\xi_j(\omega)=a_k\}} - \mathbb{P}(\xi_1 = a_k)\right) m_k(x_1,\dots,x_p)\Bigg|$$

$$\leqslant \left|\frac{1}{n}\sum_{j=1}^n \mathbb{1}_{\{\xi_j(\omega)=a_1\}} - \mathbb{P}(\xi_1 = a_1)\right| \sup_{(x_1,\dots,x_p)\in I^p} |m_1(x_1,\dots,x_p)|$$

$$+ \cdots + \left|\frac{1}{n}\sum_{j=1}^n \mathbb{1}_{\{\xi_j(\omega)=a_k\}} - \mathbb{P}(\xi_1 = a_k)\right| \sup_{(x_1,\dots,x_p)\in I^p} |m_k(x_1,\dots,x_p)|$$

$$\to 0 \qquad \text{as} \;\; n\to\infty$$

for $\mathbb{P}$-a.e. $\omega\in\Omega$, yielding (3.2), where the last step follows by the strong law of large numbers and

$$\sup_{(x_1,\dots,x_p)\in I^p} |m_i(x_1,\dots,x_p)| = \max_{t\in I}|t| < \infty, \qquad i=1,\dots,k,$$

where we used that $m_i(x_1,\dots,x_p) = x_1$ whenever $x_1 = \ldots = x_p \in I$ and that I is compact.

Now we turn to prove (3.3). For all $\omega\in\Omega$ and $x_1,\dots,x_p\in I$, we have

$$\sqrt{n}\,\big|(\overline{S}_n(\omega))(x_1,\dots,x_p) - (\mathbb{E}(M_1))(x_1,\dots,x_p)\big|$$

$$= \left|\left\langle \sqrt{n}\begin{pmatrix} \frac{1}{n}\sum_{j=1}^n \mathbb{1}_{\{\xi_j(\omega)=a_1\}} - \mathbb{P}(\xi_1=a_1) \\ \vdots \\ \frac{1}{n}\sum_{j=1}^n \mathbb{1}_{\{\xi_j(\omega)=a_k\}} - \mathbb{P}(\xi_1=a_k)\end{pmatrix}, \boldsymbol{m}(x_1,\dots,x_p)\right\rangle\right|,$$

and, by the multidimensional central limit theorem,

$$\sqrt{n}\begin{pmatrix} \frac{1}{n}\sum_{j=1}^n \mathbb{1}_{\{\xi_j=a_1\}} - \mathbb{P}(\xi_1=a_1) \\ \vdots \\ \frac{1}{n}\sum_{j=1}^n \mathbb{1}_{\{\xi_j=a_k\}} - \mathbb{P}(\xi_1=a_k)\end{pmatrix} \xrightarrow{\mathcal{D}} \mathcal{N}_k\Big(\mathbf{0}, (\operatorname{Cov}(\mathbb{1}_{\{\xi_1=a_i\}}, \mathbb{1}_{\{\xi_1=a_j\}}))_{i,j=1}^k\Big)$$

as $n\to\infty$, where

$$\operatorname{Cov}(\mathbb{1}_{\{\xi_1=a_i\}}, \mathbb{1}_{\{\xi_1=a_j\}}) = \begin{cases} \mathbb{P}(\xi_1=a_i) - \mathbb{P}(\xi_1=a_i)^2 & \text{if } \; i=j, \\ -\,\mathbb{P}(\xi_1=a_i)\,\mathbb{P}(\xi_1=a_j) & \text{if } \; i\neq j, \end{cases}$$

$$= q_{i,j}, \qquad i,j=1,\dots,k.$$

Further, since I is compact and $m_i\in\mathcal{M}_p$, $i=1,\dots,k$, we have the set $\boldsymbol{m}(I\times\cdots\times I) = I^k$ is a compact subset of $\mathbb{R}^k$, so, by Theorem A.1, the mapping

$$\mathbb{R}^k \ni \boldsymbol{q} \mapsto \sup_{\boldsymbol{y}\in\boldsymbol{m}(I,\dots,I)} |\langle \boldsymbol{q}, \boldsymbol{y}\rangle| = \sup_{(x_1,\dots,x_p)\in I^p} |\langle \boldsymbol{q}, \boldsymbol{m}(x_1,\dots,x_p)\rangle|$$

is well-defined and continuous. Consequently, the continuous mapping theorem yields (3.3). □

First proof of Corollary 3.2. We can apply Theorem 3.1 with $p = 2$. Namely, using the notations of Theorem 3.1, we have

$$\mathcal{N}_2(\mathbf{0}, \boldsymbol{Q}) = \mathcal{N}_2\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix} q(1-q) & -q(1-q) \\ -q(1-q) & q(1-q)\end{pmatrix}\right) \stackrel{\mathcal{D}}{=} \begin{pmatrix}-\eta\\ \eta\end{pmatrix},$$

where η is a 1-dimensional random variable having distribution $\mathcal{N}(0, q(1-q))$, and

$$\boldsymbol{m}(x_1, x_2) = \begin{pmatrix} \frac{x_1+x_2}{2} \\ \sqrt{x_1x_2}\end{pmatrix}, \qquad x_1, x_2 \in [0,1].$$

Hence

$$\begin{aligned}\sup_{x_1,x_2\in[0,1]} \left|\langle \mathcal{N}_2(\mathbf{0}, \boldsymbol{Q}), \boldsymbol{m}(x_1,x_2)\rangle\right| &\stackrel{\mathcal{D}}{=} \sup_{x_1,x_2\in[0,1]} \left| -\eta \frac{x_1+x_2}{2} + \eta\sqrt{x_1x_2}\right| \\ &= |\eta| \sup_{x_1,x_2\in[0,1]} \left|\frac{x_1+x_2}{2} - \sqrt{x_1x_2}\right| \stackrel{\mathcal{D}}{=} \frac{1}{2}|\mathcal{N}(0, q(1-q))|,\end{aligned}$$

as desired, since

$$\begin{aligned}\sup_{x_1,x_2\in[0,1]} \left|\frac{x_1+x_2}{2} - \sqrt{x_1x_2}\right| &= \sup_{x_1,x_2\in[0,1]} \left(\frac{x_1+x_2}{2} - \sqrt{x_1x_2}\right) \\ &= \frac{1}{2} \sup_{x_1,x_2\in[0,1]} (\sqrt{x_1} - \sqrt{x_2})^2 = \frac{1}{2}\left(\sup_{x_1,x_2\in[0,1]} |\sqrt{x_1} - \sqrt{x_2}|\right)^2 = \frac{1}{2}.\end{aligned} \tag{4.1}$$

□

Second proof of Corollary 3.2. We give a direct proof as well, not refereeing to Theorem 3.1. By Example 2.5, for each $n \in \mathbb{N}$, M_n is a random mean generated by ξ_n. Using that $\sum_{j=1}^n \mathbb{1}_{\{\xi_j=1\}} = n - \sum_{j=1}^n \mathbb{1}_{\{\xi_j=0\}}$, $n \in \mathbb{N}$, for all $x_1, x_2 \in I$, we have

$$\begin{aligned}(\overline{S}_n(\omega))(x_1,x_2) &= \frac{1}{n}\sum_{j=1}^n \mathbb{1}_{\{\xi_j(\omega)=0\}} \frac{x_1+x_2}{2} + \frac{1}{n}\sum_{j=1}^n \mathbb{1}_{\{\xi_j(\omega)=1\}}\sqrt{x_1x_2} \\ &= \frac{1}{n}\sum_{j=1}^n \mathbb{1}_{\{\xi_j(\omega)=0\}} \left(\frac{x_1+x_2}{2} - \sqrt{x_1x_2}\right) + \sqrt{x_1x_2},\end{aligned}$$

and, by Example 2.13,

$$\begin{aligned}(\mathbb{E}(M_1))(x_1,x_2) &= \frac{x_1+x_2}{2}\,\mathbb{P}(\xi_1 = 0) + \sqrt{x_1x_2}\,\mathbb{P}(\xi_1 = 1) \\ &= \mathbb{P}(\xi_1 = 0)\left(\frac{x_1+x_2}{2} - \sqrt{x_1x_2}\right) + \sqrt{x_1x_2}, \qquad x_1, x_2 \in I.\end{aligned}$$

Hence

$$\kappa_n(\omega) = \varrho(\overline{S}_n(\omega), \mathbb{E}(M_1)) = \sup_{(x_1,x_2)\in I^2} \left|(\overline{S}_n(\omega))(x_1,x_2) - (\mathbb{E}(M_1))(x_1,x_2)\right|$$

$$= \sup_{(x_1,x_2)\in I^2} \left| \left(\frac{1}{n}\sum_{j=1}^{n} \mathbb{1}_{\{\xi_j(\omega)=0\}} - \mathbb{P}(\xi_1 = 0)\right) \left(\frac{x_1+x_2}{2} - \sqrt{x_1x_2}\right) \right|$$

$$= \left(\sup_{(x_1,x_2)\in I^2} \left(\frac{x_1+x_2}{2} - \sqrt{x_1x_2}\right)\right) \left| \frac{1}{n}\sum_{j=1}^{n} \mathbb{1}_{\{\xi_j(\omega)=0\}} - \mathbb{P}(\xi_1 = 0) \right|$$

for all $n \in \mathbb{N}$ and $\omega \in \Omega$. By the strong law of large numbers, we have (3.4). The central limit theorem together with (4.1) and the continuous mapping theorem applied to the function $\mathbb{R} \ni x \mapsto |x|$ yield (3.5). □

Proof of Theorem 3.3. First, we check that for each $n \in \mathbb{N}$, M_n is a random mean generated by $\boldsymbol{\xi}_n$. For each $n \in \mathbb{N}$ and $\omega \in \Omega$, $M_n(\omega)$ can be written in the form

$$(M_n(\omega))(x_1,\ldots,x_p) = f(x_1,\ldots,x_p,\boldsymbol{\xi}_n(\omega)), \qquad x_1,\ldots,x_p \in I,$$

where $f : I^p \times \mathbb{R}^{p-1} \to I$ is a $(\mathcal{B}(I^p)\times\mathcal{B}(\mathbb{R}^{p-1}), \mathcal{B}(I))$-measurable function satisfying

$$f(x_1,\ldots,x_p,y_1,\ldots,y_{p-1}) = y_1x_1 + \cdots + y_{p-1}x_{p-1} + (1 - y_1 - \cdots - y_{p-1})x_p$$

for $x_1,\ldots,x_p \in I$, $y_1,\ldots,y_{p-1} \in \mathbb{R}_+$ with $y_1+\cdots+y_{p-1} \leqslant 1$, and $f(\cdot,\ldots,\cdot,y_1,\ldots,y_{p-1})$ is a fixed (arbitrary) element of $\mathcal{M}_p$ for any $(y_1,\ldots,y_{p-1}) \in \mathbb{R}^{p-1} \setminus \{(y_1,\ldots,y_{p-1}) \in \mathbb{R}_+^p : y_1+\cdots+y_{p-1} \leqslant 1\}$. Hence, by Theorem 2.4, M_n is a random mean generated by $\boldsymbol{\xi}_n$ for each $n \in \mathbb{N}$. Further, for the expectation $\mathbb{E}(M_1) \in \mathcal{M}_p$ of M_1 we have

$$(\mathbb{E}(M_1))(x_1,\ldots,x_p) = \int_{\mathbb{R}^{p-1}} f(x_1,\ldots,x_p,y_1,\ldots,y_{p-1})\, \mathbb{P}_{\boldsymbol{\xi}_1}(\mathrm{d}y_1,\ldots,\mathrm{d}y_{p-1})$$

$$= \int_{\{y_1,\ldots,y_{p-1}\in\mathbb{R}_+ : y_1+\cdots+y_{p-1}\leqslant 1\}} \Big(y_1x_1 + \cdots + y_{p-1}x_{p-1} + (1-y_1-\cdots-y_{p-1})x_p\Big)\, \mathbb{P}_{\boldsymbol{\xi}_1}(\mathrm{d}y_1,\ldots,\mathrm{d}y_{p-1})$$

$$= \sum_{i=1}^{p-1} x_i\, \mathbb{E}(\xi_1^{(i)}) + x_p\left(1 - \sum_{i=1}^{p-1}\mathbb{E}(\xi_1^{(i)})\right), \qquad x_1,\ldots,x_p \in I,$$

where $\mathbb{E}(\xi_1^{(i)}) \in [0,1]$, $i = 1,\ldots,p-1$, $1 - \sum_{i=1}^{p-1}\mathbb{E}(\xi_1^{(i)}) \in [0,1]$, and the first equality follows by (2.6).

For all $n \in \mathbb{N}$, $\omega \in \Omega$ and $x_1,\ldots,x_p \in I$, we have

$$(\overline{S}_n(\omega))(x_1,\ldots,x_p) = \frac{1}{n}\sum_{j=1}^{n}\Big(\xi_j^{(1)}(\omega)x_1 + \cdots + \xi_j^{(p-1)}(\omega)x_{p-1} + (1 - \xi_j^{(1)}(\omega) - \cdots - \xi_j^{(p-1)}(\omega))x_p\Big) =$$

$$= \left(\frac{1}{n}\sum_{j=1}^{n}\xi_j^{(1)}(\omega)\right)x_1 + \cdots + \left(\frac{1}{n}\sum_{j=1}^{n}\xi_j^{(p-1)}(\omega)\right)x_{p-1}$$
$$+ \left(1 - \frac{1}{n}\sum_{j=1}^{n}\xi_j^{(1)}(\omega) - \cdots - \frac{1}{n}\sum_{j=1}^{n}\xi_j^{(p-1)}(\omega)\right)x_p.$$

Hence

$$\begin{aligned}
\kappa_n(\omega) &= \varrho(\overline{S}_n(\omega), \mathbb{E}(M_1)) \\
&= \sup_{x_1,\ldots,x_p \in I} \left|(\overline{S}_n(\omega))(x_1,\ldots,x_p) - (\mathbb{E}(M_1))(x_1,\ldots,x_p)\right| \\
&= \sup_{x_1,\ldots,x_p \in I} \Bigg| \left(\frac{1}{n}\sum_{j=1}^{n}\xi_j^{(1)}(\omega) - \mathbb{E}(\xi_1^{(1)})\right)x_1 + \cdots + \left(\frac{1}{n}\sum_{j=1}^{n}\xi_j^{(p-1)}(\omega) - \mathbb{E}(\xi_1^{(p-1)})\right)x_{p-1} \\
&\qquad - \left[\frac{1}{n}\sum_{j=1}^{n}\xi_j^{(1)}(\omega) - \mathbb{E}(\xi_1^{(1)}) + \cdots + \frac{1}{n}\sum_{j=1}^{n}\xi_j^{(p-1)}(\omega) - \mathbb{E}(\xi_1^{(p-1)})\right]x_p \Bigg| \\
&\leqslant 2\left(\max_{t\in I}|t|\right)\sum_{k=1}^{p-1}\left|\frac{1}{n}\sum_{j=1}^{n}\xi_j^{(k)}(\omega) - \mathbb{E}(\xi_1^{(k)})\right| \to 0 \qquad \text{as} \quad n\to\infty
\end{aligned}$$

for $\mathbb{P}$-a.e. $\omega \in \Omega$, yielding (3.6), where we used the strong law of large numbers and that $\max_{t\in I}|t| < \infty$, since I is compact.

Now we turn to prove (3.7). For all $\omega \in \Omega$ and $x_1,\ldots,x_p \in I$, we have

$$\begin{aligned}
&\sqrt{n}\left|(\overline{S}_n(\omega))(x_1,\ldots,x_p) - (\mathbb{E}(M_1))(x_1,\ldots,x_p)\right| \\
&= \left|\left\langle \sqrt{n}\begin{pmatrix} \frac{1}{n}\sum_{j=1}^{n}\xi_j^{(1)}(\omega) - \mathbb{E}(\xi_1^{(1)}) \\ \vdots \\ \frac{1}{n}\sum_{j=1}^{n}\xi_j^{(p-1)}(\omega) - \mathbb{E}(\xi_1^{(p-1)}) \end{pmatrix}, \begin{pmatrix} x_1 - x_p \\ \vdots \\ x_{p-1} - x_p \end{pmatrix} \right\rangle\right|,
\end{aligned}$$

and, by the multidimensional central limit theorem,

$$\sqrt{n}\begin{pmatrix} \frac{1}{n}\sum_{j=1}^{n}\xi_j^{(1)} - \mathbb{E}(\xi_1^{(1)}) \\ \vdots \\ \frac{1}{n}\sum_{j=1}^{n}\xi_j^{(p-1)} - \mathbb{E}(\xi_1^{(p-1)}) \end{pmatrix} \xrightarrow{\mathcal{D}} \mathcal{N}_{p-1}(\mathbf{0}, \operatorname{Cov}(\xi_1^{(1)},\ldots,\xi_1^{(p-1)}))$$

as $n \to \infty$. Since I is compact, the set

$$\left\{ \begin{pmatrix} x_1 - x_p \\ \vdots \\ x_{p-1} - x_p \end{pmatrix} : x_1,\ldots,x_p \in I \right\}$$

is compact as well. Indeed,

$$\begin{pmatrix} x_1 - x_p \\ \vdots \\ x_{p-1} - x_p \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 & -1 \\ 0 & 1 & \cdots & 0 & -1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}, \qquad (x_1, \ldots, x_p) \in I^p,$$

the set I^p is compact, and it is known that the image of a compact set of $\mathbb{R}^p$ by a continuous map is compact. Hence, by Theorem A.1, the mapping

$$\mathbb{R}^{p-1} \ni \boldsymbol{q} \mapsto \sup_{x_1,\ldots,x_p \in I} \left| \left\langle \boldsymbol{q}, \begin{pmatrix} x_1 - x_p \\ \vdots \\ x_{p-1} - x_p \end{pmatrix} \right\rangle \right|$$

is well-defined and continuous, so the continuous mapping theorem yields (3.7). □

Proof of Corollary 3.4. Theorem 3.3 yields Corollary 3.4, since

$$\sup_{x_1,x_2 \in I} \left| \langle \mathcal{N}(0, \mathbb{D}^2(\xi_1)), x_1 - x_2 \rangle \right| = \sup_{x_1,x_2 \in I} |x_1 - x_2| \cdot |\mathcal{N}(0, \mathbb{D}^2(\xi_1))|.$$

□

Proof of Theorem 3.5. First, we check that for each $n \in \mathbb{N}$, M_n is a random mean generated by ξ_n. For all $n \in \mathbb{N}$ and $\omega \in \Omega$, $M_n(\omega)$ can be written in the form

$$(M_n(\omega))(x_1, x_2) = f(x_1, x_2, \xi_n(\omega)), \qquad x_1, x_2 \in I,$$

where $f : I^2 \times \mathbb{R} \to I$ is a $(\mathcal{B}(I^2) \times \mathcal{B}(\mathbb{R}), \mathcal{B}(I))$-measurable function satisfying

$$f(x_1, x_2, y) = x_1^y x_2^{1-y}, \qquad x_1, x_2 \in I, \quad y \in [0, 1],$$

and $f(\cdot, \cdot, y)$ is a fixed (arbitrary) element of $\mathcal{M}_2$ for any $y \in \mathbb{R} \setminus [0, 1]$. Hence, by Theorem 2.4, M_n is a random mean generated by ξ_n for each $n \in \mathbb{N}$. We also have that for each $n \in \mathbb{N}$ and $x_1, x_2 \in I$, the mapping $\Omega \ni \omega \mapsto (M_n(\omega))(x_1, x_2)$ is $(\mathcal{A}, \mathcal{B}(\mathbb{R}))$-measurable, i.e., it is a random variable, since $(M_n(\omega))(x_1, x_2) = \varphi_{x_1,x_2}(M_n(\omega))$, $\omega \in \Omega$, M_n is $(\mathcal{A}, \mathcal{B}(\mathcal{C}(I^2)))$-measurable and $\varphi_{x_1,x_2}(h) := h(x_1, x_2)$, $h \in \mathcal{C}(I^2)$, is a linear functional.

By the assumptions, there exists $c, \widetilde{c} \in (0, \infty)$ such that $c < x < \widetilde{c}$ for all $x \in I$, so $0 < c \leqslant \min(I) \leqslant \max(I) \leqslant \widetilde{c} < \infty$. Further, note that $\mathbb{E}(\xi_1)$ and $\mathbb{D}^2(\xi_1)$ exist, and $\mathbb{E}(\xi_1) \in [c, \widetilde{c}]$. For all $n \in \mathbb{N}$ and $x_1, x_2 \in I$, we have

$$\frac{\left(\prod_{j=1}^n (M_j(\cdot))(x_1, x_2)\right)^{\frac{1}{n}}}{x_1^{\mathbb{E}(\xi_1)} x_2^{1-\mathbb{E}(\xi_1)}} = \frac{\left(\prod_{j=1}^n x_1^{\xi_j} x_2^{1-\xi_j}\right)^{\frac{1}{n}}}{x_1^{\mathbb{E}(\xi_1)} x_2^{1-\mathbb{E}(\xi_1)}} = x_1^{\frac{1}{n}\sum_{j=1}^n \xi_j - \mathbb{E}(\xi_1)} x_2^{\frac{1}{n}\sum_{j=1}^n (1-\xi_j) - (1-\mathbb{E}(\xi_1))}.$$

Hence, using that the functions $\exp$ and $\ln$ are strictly increasing, we have for each $n \in \mathbb{N}$,

$$
\begin{aligned}
&\sup_{x_1,x_2\in I} \frac{\left(\prod_{j=1}^n (M_j(\cdot))(x_1,x_2)\right)^{\frac{1}{n}}}{x_1^{\mathbb{E}(\xi_1)} x_2^{1-\mathbb{E}(\xi_1)}} = \sup_{x_1,x_2\in I} x_1^{\frac{1}{n}\sum_{j=1}^n \xi_j - \mathbb{E}(\xi_1)} x_2^{\frac{1}{n}\sum_{j=1}^n (1-\xi_j) - (1-\mathbb{E}(\xi_1))} \\
&\quad = \sup_{x_1,x_2\in I} \exp\left\{ \left(\frac{1}{n}\sum_{j=1}^n \xi_j - \mathbb{E}(\xi_1)\right)\ln(x_1) + \left(\frac{1}{n}\sum_{j=1}^n (1-\xi_j) - (1-\mathbb{E}(\xi_1))\right)\ln(x_2)\right\} \\
&\quad = \exp\left\{ \sup_{x_1,x_2\in I} \left(\frac{1}{n}\sum_{j=1}^n \xi_j - \mathbb{E}(\xi_1)\right)\ln\left(\frac{x_1}{x_2}\right)\right\} \\
&\quad = \exp\left\{ \left(\frac{1}{n}\sum_{j=1}^n \xi_j - \mathbb{E}(\xi_1)\right) \mathbb{1}_{\left\{\frac{1}{n}\sum_{j=1}^n \xi_j - \mathbb{E}(\xi_1)\geqslant 0\right\}} \sup_{x_1,x_2\in I} \ln\left(\frac{x_1}{x_2}\right)\right. \\
&\qquad\qquad \left. + \left(\frac{1}{n}\sum_{j=1}^n \xi_j - \mathbb{E}(\xi_1)\right) \mathbb{1}_{\left\{\frac{1}{n}\sum_{j=1}^n \xi_j - \mathbb{E}(\xi_1)< 0\right\}} \inf_{x_1,x_2\in I} \ln\left(\frac{x_1}{x_2}\right)\right\} \\
&\quad = \exp\left\{ \left|\frac{1}{n}\sum_{j=1}^n \xi_j - \mathbb{E}(\xi_1)\right| \ln\left(\frac{\max(I)}{\min(I)}\right)\right\} \\
&\quad = \left(\frac{\max(I)}{\min(I)}\right)^{\left|\frac{1}{n}\sum_{j=1}^n \xi_j - \mathbb{E}(\xi_1)\right|}.
\end{aligned}
$$

Using the strong law of large numbers and that $\frac{\max(I)}{\min(I)} \in (0,\infty)$, we have (3.10). The central limit theorem together with the continuous mapping theorem applied to the function $\mathbb{R} \ni x \mapsto \left(\frac{\max(I)}{\min(I)}\right)^{|x|}$ yield (3.11). □

5 Declarations

Funding. Mátyás Barczy is supported by grant NKFIH-1279-2/2020 of the Ministry for Innovation and Technology, Hungary.

Conflicts of interest/Competing interests. We do not have any conflicts of interest/competing interests.

Availability of data and material. Not applicable.

Code availability. Not applicable.

Appendix

A Continuity of the supremum

The following result is known, however, we could not address any book or article containing it, only an internet blog due to Wong [14]. Because it is used in the verifications of Theorems 3.1 and 3.3 we present a proof of it.

A.1 Theorem. *Let X and Y be topological spaces such that Y is compact, and let $f : X \times Y \to \mathbb{R}$ be a continuous function. Then the function $g : X \to \mathbb{R}$, $g(x) := \sup_{y \in Y} f(x, y)$, $x \in X$, is well-defined and continuous.*

Proof. Because of the continuity of f, the function $Y \ni y \mapsto f(x, y)$ is continuous for every fixed $x \in X$, and the compactness of Y implies that its supremum is finite and it is attained. So, $g(x)$, $x \in X$, is well-defined.

Let $r \in \mathbb{R}$ be arbitrarily fixed. We will prove that the inverse images $g^{-1}((-\infty, r))$ and $g^{-1}((r, \infty))$ are open in X.

First, we prove that $g^{-1}((r, \infty))$ is open in X. Let us denote by $\pi_X : X \times Y \to X$, $\pi_X(x, y) := x$, $(x, y) \in X \times Y$, the canonical projection onto X, which is known to be continuous and open (i.e., maps open sets to open sets). Moreover, for every $x_0 \in X$ there is at least one $y_0 \in Y$ such that

$$g(x_0) = \sup_{y \in Y} f(x_0, y) = f(x_0, y_0).$$

So, for every $r \in \mathbb{R}$ we can write

$$\begin{aligned} g^{-1}((r, \infty)) = \{x \in X \mid g(x) > r\} &= \Big\{x \in X \mid \sup_{y \in Y} f(x, y) > r\Big\} \\ &= \Big\{x \in X \mid f(x, y) > r \text{ for some } y \in Y\Big\} = \pi_X(f^{-1}((r, \infty))). \end{aligned}$$

Because f is continuous, $f^{-1}((r, \infty))$ is open in $X \times Y$. The canonical projection π_X is an open map, which entails that $\pi_X(f^{-1}((r, \infty)))$ is an open subset of X, so is $g^{-1}((r, \infty))$.

Next, we prove that $g^{-1}((-\infty, r))$ is open in X for every $r \in \mathbb{R}$. If $g(x) < r$ for some $x \in X$, then, by the definition of g, we have $f(x, y) < \tilde{r} < r$ for every $y \in Y$, where $\tilde{r}$ satisfies $g(x) < \tilde{r} < r$. In other words, if $x \in g^{-1}((-\infty, r))$, then $\{x\} \times Y \subset f^{-1}((-\infty, \tilde{r}))$. Because of the continuity of f, the set $f^{-1}((-\infty, \tilde{r}))$ is open in $X \times Y$. So, if $x \in g^{-1}((-\infty, r))$ and $y \in Y$, then there are open sets $U_{x,y} \subset X$ and $V_{x,y} \subset Y$ such that $U_{x,y} \times V_{x,y}$ is an open neighbourhood of $(x, y) \in X \times Y$ and it is contained in $f^{-1}((-\infty, \tilde{r}))$. For a fixed $x \in g^{-1}((-\infty, r))$, the sets $V_{x,y}$, $y \in Y$, give an open cover of Y, and, because of the compactness of Y, there exist $k(x) \in \mathbb{N}$ and $y_1, \ldots, y_{k(x)} \in Y$ such that $Y = \bigcup_{i=1}^{k(x)} V_{x,y_i}$. Using that $A \times (B \cup C) = (A \times B) \cup (A \times C)$ for any sets A, B, C, this entails that

$$\{x\} \times Y \subset \left(\bigcap_{i=1}^{k(x)} U_{x,y_i} \right) \times Y = \bigcup_{j=1}^{k(x)} \left(\left(\bigcap_{i=1}^{k(x)} U_{x,y_i} \right) \times V_{x,y_j} \right) \subset \bigcup_{j=1}^{k(x)} \big(U_{x,y_j} \times V_{x,y_j} \big) \subset f^{-1}((-\infty, \tilde{r}))$$

for $x \in g^{-1}((-\infty, r))$. Especially, given $x \in g^{-1}((-\infty, r))$, for all $x^* \in \bigcap_{i=1}^{k(x)} U_{x,y_i}$ and $y^* \in Y$ we have $f(x^*, y^*) < \tilde{r}$, and hence $g(x^*) \leqslant \tilde{r} < r$ for each $x^* \in \bigcap_{i=1}^{k(x)} U_{x,y_i}$. From this we can derive

$$g^{-1}((-\infty, r)) = \bigcup_{x \in g^{-1}((-\infty, r))} \left(\bigcap_{i=1}^{k(x)} U_{x,y_i} \right).$$

On the right hand side of the above equality there is a union of open sets in X, which is open, so $g^{-1}((-\infty, r))$ is open as well.

The family $\{(-\infty, r), (r, \infty) : r \in \mathbb{R}\}$ constitutes a subbase of the usual topology of $\mathbb{R}$, which implies that the preimage of every open set of $\mathbb{R}$ by g is open. Thus g is continuous. □

Acknowledgements

We would like to thank the referees for their comments (especially for the second proof of the first part of Theorem 2.10) that helped us to improve the paper.

References

[1] Karol Baron. Weak law of large numbers for iterates of random-valued functions. *Aequationes Math.*, 93(2):415–423, 2019.

[2] Karol Baron and Witold Jarczyk. Random-valued functions and iterative functional equations. *Aequationes Math.*, 67(1-2):140–153, 2004.

[3] Gleb Beliakov, Humberto Bustince Sola, and Tomasa Calvo Sánchez. *A practical guide to averaging functions*, volume 329 of *Studies in Fuzziness and Soft Computing*. Springer, Cham, 2016.

[4] Luísa Borsato, Eduardo Horta, and Rafael Rigão Souza. A characterization of the strong law of large numbers for Bernoulli sequences. arXiv 2008.00318, 2020.

[5] Donald L. Cohn. *Measure theory*. Birkhäuser, Boston, Mass., 1980.

[6] Tuomas Hytönen, Jan van Neerven, Mark Veraar, and Lutz Weis. *Analysis in Banach spaces. Vol. I. Martingales and Littlewood-Paley theory*, volume 63 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]*. Springer, Cham, 2016.

[7] Justyna Jarczyk and Witold Jarczyk. Invariance of means. *Aequationes Math.*, 92(5):801–872, 2018.

[8] Justyna Jarczyk and Witold Jarczyk. Gaussian iterative algorithm and integrated automorphism equation for random means. *Discrete and Continuous Dynamical Systems*, 40(12):6837–6844, 2020.

[9] Olav Kallenberg. *Random measures, theory and applications*, volume 77 of *Probability Theory and Stochastic Modelling*. Springer, Cham, 2017.

[10] Hui Hsiung Kuo. *Gaussian measures in Banach spaces*. Lecture Notes in Mathematics, Vol. 463. Springer-Verlag, Berlin-New York, 1975.

[11] Zsolt Páles and Amr Zakaria. On the local and global comparison of generalized Bajraktarević means. *J. Math. Anal. Appl.*, 455(1):792–815, 2017.

[12] Victor M. Panaretos and Yoav Zemel. *An invitation to statistics in Wasserstein space.* Cham: Springer, 2020.

[13] Cédric Villani. *Topics in optimal transportation*. Providence, RI: American Mathematical Society (AMS), 2003.

[14] Willie Wong. Continuity of the infimum. *Bubbles Bad; Ripples Good, Blog*, 2011.