2023

QUANTIZATION FOR A SET OF DISCRETE DISTRIBUTIONS ON THE SET OF NATURAL NUMBERS

Juan Gomez

Haily Martinez

Mrinal Kanti Roychowdhury

Alexis Salazar

Daniel J. Vallez

QUANTIZATION FOR A SET OF DISCRETE DISTRIBUTIONS ON THE SET OF NATURAL NUMBERS

[1]JUAN GOMEZ, [2]HAILY MARTINEZ, [3]MRINAL K. ROYCHOWDHURY, [4]ALEXIS SALAZAR, AND [5]DANIEL J. VALLEZ

Abstract. The quantization scheme in probability theory deals with finding a best approximation of a given probability distribution by a probability distribution that is supported on finitely many points. In this paper, first we state and prove a theorem, and then give a conjecture. We verify the conjecture by a few examples. Assuming that the conjecture is true, for a set of discrete distributions on the set of natural numbers we have calculated the optimal sets of n-means and the nth quantization errors for all positive integers n. In addition, the quantization dimension is also calculated.

1. Introduction

The most common form of quantization is rounding-off. Its purpose is to reduce the cardinality of the representation space, in particular, when the input data is real-valued. It has broad applications in communications, information theory, signal processing and data compression (see [GG, GL1, GL2, GN, P, Z1, Z2]). Let $\mathbb{R}^d$ denote the d-dimensional Euclidean space equipped with the Euclidean norm $\|\cdot\|$, and let P be a Borel probability measure on $\mathbb{R}^d$. Then, the nth *quantization error* for P, with respect to the squared Euclidean distance, is defined by

$$V_n := V_n(P) = \inf \Big\{ V(P;\alpha) : \alpha \subset \mathbb{R}^d, 1 \leq \text{ card}(\alpha) \leq n \Big\},$$

where $V(P;\alpha) = \int \min\limits_{a\in\alpha} \|x-a\|^2 dP(x)$ represents the distortion error due to the set α with respect to the probability distribution P, and for a set A, card(A) represents the cardinality of the set A. A set α for which the infimum occurs and contains no more than n points is called an *optimal set of n-means*, and is denoted by $\alpha_n := \alpha_n(P)$. The elements of an optimal set are also called as *optimal quantizers.* It is known that for a Borel probability measure P if its support contains infinitely many elements and $\int \|x\|^2 dP(x)$ is finite, then an optimal set of n-means always exists and has exactly n-elements [AW, GKL, GL1, GL2]. The number

$$D(P) := \lim_{n\to\infty} \frac{2\log n}{-\log V_n(P)}, \tag{1}$$

if it exists, is called the *quantization dimension of P*. Quantization dimension measures the speed at which the specified measure of the error goes to zero as n tends to infinity. For a finite set $\alpha \subset \mathbb{R}^d$ and $a \in \alpha$, by $M(a|\alpha)$ we denote the set of all elements in $\mathbb{R}^d$ which are the nearest to a among all the elements in α, i.e., $M(a|\alpha) = \{x \in \mathbb{R}^d : \|x-a\| = \min_{b\in\alpha} \|x-b\|\}$. $M(a|\alpha)$ is called the *Voronoi region* generated by $a \in \alpha$. On the other hand, the set $\{M(a|\alpha) : a \in \alpha\}$ is called the *Voronoi diagram* or *Voronoi tessellation* of $\mathbb{R}^d$ with respect to the set α. The following proposition provides further information on the Voronoi regions generated by an optimal set of n-means (see [GG, GL2]).

Proposition 1.1. *Let α be an optimal set of n-means, $a \in \alpha$, and $M(a|\alpha)$ be the Voronoi region generated by $a \in \alpha$, i.e.,*

$$M(a|\alpha) = \{x \in \mathbb{R}^d : \|x-a\| = \min_{b\in\alpha} \|x-b\|\}.$$

Then, for every $a \in \alpha$,

(i) $P(M(a|\alpha)) > 0$,
(ii) $P(\partial M(a|\alpha)) = 0$,

2010 *Mathematics Subject Classification.* 60Exx, 94A34.
Key words and phrases. Discrete distribution, optimal sets, quantization error.

(iii) $a = E(X : X \in M(a|\alpha))$, *and*
(iv) P*-almost surely the set* $\{M(a|\alpha) : a \in \alpha\}$ *forms a Voronoi partition of* $\mathbb{R}^d$.

From the above proposition, we can say that if α is an optimal set of n-means for P, then each $a \in \alpha$ is the conditional expectation of the random variable X given that X takes values on the Voronoi region of a. Sometimes, we also refer to such an $a \in \alpha$ as the centroid of its own Voronoi region. In this regard, interested readers can see [DFG, DR, R1].

A vector $(p_1, p_2, p_3, \cdots)$ is called a probability distribution if $0 < p_j < 1$ for all $j \in \mathbb{N}$ and $j \geq 0$ such that $\sum_{j\geq 1} p_j = 1$. Notice that $(p_1, p_2, p_3, \cdots)$ can be a finite, or an infinite vector, where by a finite vector it is meant that the number of coordinates in the vector is a finite number, otherwise it is called an infinite vector.

For a Borel probability measure P on the set $\mathbb{R}$ of real numbers let U be the largest open subset of $\mathbb{R}$ such that $P(U) = 0$, then $\mathbb{R} \setminus U$ is called the *support* of the probability measure P. For example, when an unbiased die is thrown one time, then P is a Borel probability measure on the real line with

$$\text{support}(P) = \{1, 2, 3, 4, 5, 6\},$$

and the associated probability distribution $(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6})$. When an unbiased coin is tossed two times, then P is a Borel probability measure on $\mathbb{R}^2$ with

$$\text{support}(P) = \{(1, 1), (1, 2), (2, 1), (2, 2)\},$$

and the associated probability distribution $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$, where 1 stands for 'Head' and 2 stands for 'Tail'.

Definition 1.2. *Let* $(p_1, p_2, \cdots, p_{k-1})$ *be a permutation of the set* $\{\frac{1}{2}, \frac{1}{2^2}, \cdots, \frac{1}{2^{k-1}}\}$, *where* $k \in \mathbb{N} := \{1, 2, \cdots\}$ *with* $k \geq 2$. *Define a probability measure* P *on the set* $\mathbb{R}$ *of real numbers with the support the set* $\mathbb{N}$ *of natural numbers as follows:*

$$P := \sum_{j=1}^{k-1} p_j \delta_j + \sum_{j=k}^{\infty} \frac{1}{2^j} \delta_j,$$

where for $x \in \mathbb{R}$ *the function* δ_x *represents the dirac measure, i.e., for any subset* $A \subseteq \mathbb{R}$, *we have* $\delta_x(A) = 1$ *if* $x \in A$, *and zero otherwise.*

Let us now state the following theorem and the conjecture.

Theorem 1.3. *Let* $P := \sum_{j=1}^{k-1} p_j \delta_j + \sum_{j=k}^{\infty} \frac{1}{2^j} \delta_j$ *be the probability measure as defined by Definition 1.2. Let* $\{a_1, a_2, \cdots, a_{n-3}, a_{n-2}, a_{n-1}, a_n\}$ *be an optimal set of* n*-means with* $n \geq k+2$. *Suppose that* $a_1 = 1, a_2 = 2, \cdots, a_{n-3} = n-3$. *Then, either* $a_{n-2} = n-2, a_{n-1} = Av[n-1, n], a_n = Av[n+1, \infty)$, *or* $a_{n-2} = Av[n-2, n-1], a_{n-1} = Av[n, n+1], a_n = Av[n+2, \infty)$ *with quantization error* $V_n = \frac{2^{3-n}}{3}$, *where for any* $k, \ell \in \mathbb{N}$, $Av[k, \ell]$ *and* $Av[k, \infty)$ *are defined in the next section.*

Example 1.4. Let $(\frac{1}{2^3}, \frac{1}{2^2}, \frac{1}{2})$ be a permutation of the set $\{\frac{1}{2}, \frac{1}{2^2}, \frac{1}{2^3}\}$. Write

$$P := \frac{1}{2^3}\delta_1 + \frac{1}{2^2}\delta_2 + \frac{1}{2}\delta_3 + \sum_{j=4}^{\infty} \frac{1}{2^j}\delta_j.$$

Then, P is a Borel probability measure on $\mathbb{R}$ with support the set $\mathbb{N}$ of natural numbers. Let us assume that $\{a_1, a_2, \cdots, a_n\}$ is an optimal set of n-means for $n = 6$. If $a_1 = 1$, $a_2 = 2$, and $a_3 = 3$, then by Theorem 1.3, we must have the set $\{a_4, a_5, a_6\}$ equals either the set $\{4, Av[5, 6], Av[7, \infty)\}$, or the set $\{Av[4, 5], Av[6, 7], Av[8, \infty)\}$ with quantization error $V_6 = \frac{1}{24}$.

Conjecture 1.5. Let $P := \sum_{j=1}^{k-1} p_j \delta_j + \sum_{j=k}^{\infty} \frac{1}{2^j} \delta_j$ be the probability measure as defined by Definition 1.2. Let $\{a_1, a_2, a_3, \cdots, a_n\}$ be an optimal set of n-means with $n \geq k+2$ such that $a_1 < a_2 < \cdots < a_n$. Then, $a_1 = 1, a_2 = 2, \cdots, a_{n-3} = n-3$.

In this paper, first we give a complete proof of Theorem 1.3. Then, we verify the conjecture by two discrete distributions as mentioned in Remark 3.2.2 and Remark 3.3.2. Under the assumption that the conjecture is true, we calculate the optimal sets of n-means and the nth quantization errors for the two discrete distributions for all $n \in \mathbb{N}$. Once the quantization error is known, the quantization dimension can easily be calculated, see Proposition 3.3.3. In addition, in the last section, we give a proposition Proposition 5.1. By this proposition, we deduce that if $(p_1, p_2, \cdots, p_{k-1})$ is not a permutation of the set $\{\frac{1}{2}, \frac{1}{2^2}, \cdots, \frac{1}{2^{k-1}}\}$, where $k \in \mathbb{N} := \{1, 2, \cdots\}$ with $k \geq 2$, then the conjecture is not true. The general proof of the conjecture is not known yet. Such a problem still remains open.

2. Preliminaries

Let $\mathbb{N} := \{1, 2, 3, \cdots\}$ be the set of natural numbers. Let $(p_1, p_2, p_3, \cdots)$, where $0 < p_j < 1$ for all $j \in \mathbb{N}$ and $\sum_{j=1}^\infty p_j = 1$, be a probability distribution. Let

$$P = \sum_{j=1}^\infty p_j \delta_j,$$

where δ_j is the dirac measure as given in Definition 1.2. Then, P is a discrete probability measure on the set $\mathbb{R}$ of real numbers with the support the set of natural numbers $\mathbb{N}$ associated with the probability distribution $(p_1, p_2, p_3, \cdots)$. In fact, if X is a random variable associated with the probability measure P, and f is the probability mass function, then we have

$$P(X = j) = f(j) = p_j.$$

Define the following notations: For $k, \ell \in \mathbb{N}$, where $k \leq \ell$, write

$$[k, \ell] := \{n : n \in \mathbb{N} \text{ and } k \leq n \leq \ell\}, \text{ and } [k, \infty) := \{n : n \in \mathbb{N} \text{ and } n \geq k\}.$$

Further, write

$$Av[k, \ell] := E\Big(X : X \in [k, \ell]\Big) = \frac{\sum_{n=k}^\ell p_n n}{\sum_{n=k}^\ell p_n}, \ Av[k, \infty) := E\Big(X : X \in [k, \infty)\Big) = \frac{\sum_{n=k}^\infty p_n n}{\sum_{n=k}^\infty p_n},$$

$$Er[k, \ell] := \sum_{n=k}^\ell p_n \Big(n - Av[k, \ell]\Big)^2, \text{ and } Er[k, \infty) := \sum_{n=k}^\infty p_n \Big(n - Av[k, \infty)\Big)^2.$$

Notice that $E(X) := E(X : X \in \text{supp}(P)) = \sum_{n=1}^\infty p_n n$, and so the optimal set of one-mean is the set $\{\sum_{n=1}^\infty p_n n\}$ with quantization error

$$V(P) = \sum_{n=1}^\infty p_n (n - E(X))^2.$$

In the following sections, we give the main results of the paper.

3. Proof of Theorem 1.3 and verifications of Conjecture 1.5

In this section, in the following subsections first we prove Theorem 1.3, and then by two different examples, we verify that Conjecture 1.5 is true. Then, we state and prove Proposition 3.3.3, which gives the quantization dimension of the probability measure P.

3.1. **Proof of Theorem 1.3.** Let $n \geq k + 2$, where $k \geq 2$. The distortion error due to the set $\beta := \{1, 2, \cdots, n-3, n-2, Av[n-1, n], Av[n+1, \infty)\}$ is given by

$$V(P; \beta) = Er[n-1, n] + Er[n+1, \infty) = \frac{2^{3-n}}{3}.$$

Since V_n is the quantization error for n-means, we have $V_n \leq \frac{2^{3-n}}{3}$. Let $\alpha := \{a_1, a_2, a_3, \cdots, a_n\}$ be an optimal set of n-means, where $1 \leq a_1 < a_2 < a_3 < \cdots < a_n < \infty$. Assume that $a_1 = 1, a_2 = 2, \cdots, a_{n-3} = n-3$. Then, the Voronoi region of a_{n-2} must contain the element $n-2$. Suppose that the Voronoi region of a_{n-2} contains the set $\{n-2, n-1, n\}$. Then,

$$V_n \geq Er[n-2, n] = \frac{13}{7} 2^{1-n} > \frac{2^{3-n}}{3} \geq V_n,$$

which is a contradiction. Hence, we can assume that the Voronoi region of a_{n-2} contains only the set $\{n-2\}$ or the set $\{n-2, n-1\}$. Let us consider the following two cases:

Case 1. The Voronoi region of a_{n-2} contains only the set $\{n-2\}$.

Then, the Voronoi region of a_{n-1} must contain the element $n-1$. Suppose that the Voronoi region of a_{n-1} contains the set $\{n-1, n, n+1, n+2\}$. Then,

$$V_n \geq \frac{97}{15}\, 2^{-n-1} > V_n,$$

which is a contradiction. Assume that the Voronoi region of a_{n-1} contains only the set $\{n-1, n, n+1\}$. Then, as the Voronoi region of a_n contains the set $\{k : k \geq n+2\}$,

$$V_n \geq Er[n-1, n+1] + Er[n+2, \infty) = \frac{5}{7} 2^{2-n} > V_n,$$

which leads to a contradiction. Next, assume that the Voronoi region of a_{n-1} contains only the set $\{n-1\}$. Then, the Voronoi region of a_n contains the set $\{k : k \geq n\}$ yielding

$$V_n = Er[n, \infty) = 2^{2-n} > V_n,$$

which gives a contradiction. This yields the fact that the Voronoi region of a_{n-1} contains only the set $\{n-1, n\}$, and hence, the Voronoi region of a_n contains only the set $\{k : k \geq n+1\}$. Thus, in this case we have $a_{n-2} = n-2, a_{n-1} = Av[n-1, n], a_n = Av[n+1, \infty)$ with quantization error $V_n = \frac{2^{3-n}}{3}$.

Case 2. The Voronoi region of a_{n-2} contains only the set $\{n-2, n-1\}$.

Then, the Voronoi region of a_{n-1} must contain the element n. Suppose that the Voronoi region of a_{n-1} contains the set $\{n, n+1, n+2, n+3\}$. Then,

$$V_n \geq Er[n-2, n-1] + Er[n, n+3] = \frac{59}{5} 2^{-n-2} > \frac{2^{3-n}}{3} \geq V_n,$$

which leads to a contradiction. Assume that the Voronoi region of a_{n-1} contains only the set $\{n, n+1, n+2\}$. Then, as the Voronoi region of a_n contains the set $\{k : k \geq n+3\}$,

$$V_n \geq Er[n-2, n-1] + Er[n, n+2] + Er[n+3, \infty) = \frac{29}{21} 2^{1-n} > V_n,$$

which leads to a contradiction. Hence, we can assume that the Voronoi region of a_{n-1} contains only the set $\{n\}$, or the set $\{n, n+1\}$. If the Voronoi region of a_{n-1} contains only the set $\{n\}$, then

$$V_n \geq Er[n-2, n-1] + Er[n+1, \infty) = \frac{5}{3} 2^{1-n} > V_n,$$

which leads to a contradiction. Hence, we can assume that the Voronoi region of a_{n-1} contains only the set $\{n, n+1\}$, and the Voronoi region of a_n contains only the set $\{k : k \geq n+2\}$ yielding $a_{n-2} = Av[n-2, n-1], a_{n-1} = Av[n, n+1], a_n = Av[n+2, \infty)$ with quantization error $V_n = \frac{2^{3-n}}{3}$.

Case 1 and Case 2 together give the optimal sets of n-means and the nth quantization errors for all positive integers n. Thus, the proof of the theorem Theorem 1.3 is completed. ☐

3.2. **Verification of Conjecture 1.5 when** $(p_1, p_2, p_3, p_4, \cdots) = (\frac{1}{2^2}, \frac{1}{2}, \frac{1}{2^3}, \frac{1}{2^4}, \frac{1}{2^5}, \cdots)$**.** In this case the probability mass function f for the probability measure P on the set of real numbers $\mathbb{R}$ is given by

$$f(j) = \begin{cases} \frac{1}{2^2} & \text{if } j = 1, \\ \frac{1}{2} & \text{if } j = 2, \\ \frac{1}{2^n} & \text{if } j = n \text{ for } n \in \mathbb{N} \text{ and } n \neq 1, 2, \\ 0 & \text{otherwise.} \end{cases}$$

Notice that here $k = 3$ and $(p_1, \cdots, p_{k-1}) = (p_1, p_2) = (\frac{1}{2^2}, \frac{1}{2})$, where $k \in \mathbb{N}$ as defined by Definition 1.2.

Let us now prove the following proposition.

Proposition 3.2.1. *Let $n \geq 5$, and let α_n be an optimal set of n-means for the probability measure P given by*

$$P = \frac{1}{2^2}\delta_1 + \frac{1}{2}\delta_2 + \sum_{j=3}^{\infty} \frac{1}{2^j}\delta_j.$$

Then, α_n must contain the set $\{1, 2, \cdots, (n-3)\}$.

Proof. The distortion error due to the set $\beta := \{1, 2, \cdots, (n-2), Av[n-1, n], Av[n+1, \infty)\}$ is given by

$$V(P; \beta) = Er[n-1, n] + Er[n+1, \infty) = \frac{2^{3-n}}{3}.$$

Since V_n is the quantization error for n-means, we have $V_n \leq \frac{2^{3-n}}{3}$. Let $\alpha_n := \{a_1, a_2, \cdots, a_n\}$ be an optimal set of n-means such that $1 \leq a_1 < a_2 < \cdots < a_n < \infty$. We show that $a_1 = 1, a_2 = 2, \cdots, a_{n-3} = n-3$. We prove it by induction. Notice that the Voronoi region of a_1 must contain the element 1. Suppose that the Voronoi region of a_1 also contains the element 2. Then,

$$V_n > \sum_{j=1}^{2} f(j)(j - Av[1, 2])^2 = \frac{1}{6} \geq \frac{2^{3-n}}{3} \geq V_n,$$

which is a contradiction. Hence, we can conclude that the Voronoi region of a_1 contains only the element 1 yielding $a_1 = 1$. Thus, we can deduce that there exists a positive integer ℓ, where $1 \leq \ell < n-3$, such that $a_1 = 1, a_2 = 2, \cdots, a_\ell = \ell$. We now show that $a_{\ell+1} = \ell + 1$. Notice that the Voronoi region of $a_{\ell+1}$ must contain $\ell + 1$. Suppose that the Voronoi region of $a_{\ell+1}$ also contains the element $\ell + 2$. Then, we have

$$V_n > \sum_{j=\ell+1}^{\ell+2} \frac{1}{2^j}(j - Av[\ell+1, \ell+2])^2 = Er[\ell+1, \ell+2] = \frac{2^{-\ell-1}}{3} \geq \frac{2^{3-n}}{3} \geq V_n,$$

which is a contradiction. Hence, we can conclude that the Voronoi region of $a_{\ell+1}$ contains only the element $\ell + 1$ yielding $a_{\ell+1} = \ell + 1$. Notice that $2 \leq \ell + 1 \leq n - 3$. Thus, by the Principle of Mathematical Induction, we deduce that $a_1 = 1, a_2 = 2, \cdots, a_{n-3} = n-3$. Thus, the proof of the proposition is complete. □

Remark 3.2.2. Proposition 3.2.1 verifies that the conjecture Conjecture 1.5 is true.

3.3. **Verification of Conjecture 1.5 when** $(p_1, p_2, p_3, p_4, \cdots) = (\frac{1}{2^3}, \frac{1}{2^2}, \frac{1}{2}, \frac{1}{2^4}, \frac{1}{2^5}, \cdots)$**.** In this case the probability mass function f for the probability measure P on the set of real numbers $\mathbb{R}$ is given by

$$f(j) = \begin{cases} \frac{1}{2^3} & \text{if } j = 1, \\ \frac{1}{2} & \text{if } j = 3, \\ \frac{1}{2^n} & \text{if } j = n \text{ for } n \in \mathbb{N} \text{ and } n \neq 1, 3, \\ 0 & \text{otherwise.} \end{cases}$$

Notice that here $k = 4$ and $(p_1, \cdots, p_{k-1}) = (p_1, p_2, p_3) = (\frac{1}{2^3}, \frac{1}{2^2}, \frac{1}{2})$, where $k \in \mathbb{N}$ as defined by Definition 1.2.

Let us now prove the following proposition.

Proposition 3.3.1. *Let $n \geq 6$, and let α_n be an optimal set of n-means for the probability measure P given by*

$$P = \frac{1}{2^3}\delta_1 + \frac{1}{2^2}\delta_2 + \frac{1}{2}\delta_3 + \sum_{j=4}^{\infty} \frac{1}{2^j}\delta_j.$$

Then, α_n must contain the set $\{1, 2, \cdots, (n-3)\}$.

Proof. The distortion error due to the set $\beta := \{1, 2, \cdots, (n-2), Av[n-1, n], Av[n+1, \infty)\}$ is given by

$$V(P; \beta) = Er[n-1, n] + Er[n+1, \infty) = \frac{2^{3-n}}{3}.$$

Since V_n is the quantization error for n-means, we have $V_n \leq \frac{2^{3-n}}{3}$. Let $\alpha_n := \{a_1, a_2, \cdots, a_n\}$ be an optimal set of n-means such that $1 \leq a_1 < a_2 < \cdots < a_n < \infty$. We show that $a_1 = 1, a_2 = 2, \cdots, a_{n-3} = n-3$. We prove it by induction. The Voronoi region of a_1 must contain the element 1. Suppose that the Voronoi region of a_1 also contains the element 2. Notice that the remaining elements of the set of natural numbers are contained in the union of the Voronoi regions of $a_2, a_3, \cdots, a_n$ with positive distortion error yielding

$$V_n > \sum_{j=1}^{2} f(j)(j - Av[1,2])^2 = \frac{1}{12} \geq \frac{2^{3-n}}{3} \geq V_n,$$

which is a contradiction. Hence, we can conclude that the Voronoi region of a_1 contains only the element 1, yielding $a_1 = 1$. Thus, we can deduce that there exists a positive integer ℓ, where $1 \leq \ell < n-3$, such that $a_1 = 1, a_2 = 2, \cdots, a_\ell = \ell$. We now show that $a_{\ell+1} = \ell + 1$. Notice that the Voronoi region of $a_{\ell+1}$ must contain $\ell + 1$. Suppose that the Voronoi region of $a_{\ell+1}$ also contains the element $\ell + 2$. Then, proceeding in the similar lines as given in Proposition 3.2.1, we can see that a contradiction arises. Hence, we can conclude that the Voronoi region of $a_{\ell+1}$ contains only the element $\ell + 1$ yielding $a_{\ell+1} = \ell + 1$. Notice that $2 \leq \ell + 2 \leq n - 3$. Thus, by the Principle of Mathematical Induction, we deduce that $a_1 = 1, a_2 = 2, \cdots, a_{n-3} = n - 3$. Thus, the proof of the proposition is complete. □

Remark 3.3.2. Proposition 3.3.1 verifies that the conjecture Conjecture 1.5 is true.

Proposition 3.3.3. *Let $P := \sum_{j=1}^{k-1} p_j \delta_j + \sum_{j=k}^{\infty} \frac{1}{2^j} \delta_j$ be the probability measure as defined by Definition 1.2. Assume that Conjecture 1.5 is true. Then, the quantization dimension $D(P)$ exists and equals zero.*

Proof. By Theorem 1.3 and under the assumption that Conjecture 1.5 is true, the nth quantization error for any positive integer $n \geq k + 2$ for the probability measure P, defined by Definition 1.2, is obtained as $V_n(P) = \frac{2^{3-n}}{3}$. Hence, using the formula (1), we have $D(P) = 0$. □

4. Optimal quantization for the two probability distributions described in Section 3

In this section, in the following two subsections we determine the optimal sets of n-means and the nth quantization errors for all positive integers $n \geq 2$ for the two probability measures P given in Subsection 3.2 and Subsection 3.3 under the assumption that Conjecture 1.5 is true.

4.1. **Optimal quantization for P when $(p_1, p_2, p_3, p_4, \cdots) = (\frac{1}{2^2}, \frac{1}{2}, \frac{1}{2^3}, \frac{1}{2^4}, \frac{1}{2^5}, \cdots)$.** Let us give the results in the following propositions.

Proposition 4.1.1. *The optimal set of two-means is given by $\{Av[1,3], Av[4,\infty)\}$ with quantization error $V_2 = \frac{17}{28}$.*

Proof. We see that $Av[1,3] = \frac{13}{7}$, and $Av[4,\infty) = 5$. Since $3 < \frac{1}{2}(\frac{13}{7} + 5) = \frac{24}{7} < 4$, the distortion error due to the set $\beta := \{\frac{13}{7}, 5\}$ is given by

$$V(P;\beta) = Er[1,3] + Er[4,\infty) = \frac{17}{28}.$$

Since V_2 is the quantization error for two-means, we have $V_2 \leq \frac{17}{28}$. Let $\alpha := \{a_1, a_2\}$ be an optimal set of two-means such that $a_1 < a_2$. Since the points in an optimal set are the conditional expectations in their own Voronoi regions, we have $1 \leq a_1 < a_2 < \infty$. Notice that the Voronoi region of a_1 must contain 1. Suppose that the Voronoi region of a_1 contains the set $\{1, 2, 3, 4\}$. Then,

$$V_2 \geq \sum_{j=1}^{4} f(j)(j - Av[1,4])^2 = Er[1,4] = \frac{5}{8} > V_2,$$

which yields a contradiction. Hence, we can assume that the Voronoi region of a_1 contains only the set $\{1\}$ or $\{1, 2\}$, or the set $\{1, 2, 3\}$. Suppose that the Voronoi region of a_1 contains only the set $\{1\}$, and so the Voronoi region of a_2 contains the set $\{n : n \geq 2\}$. Then, we have

$$V_2 = Er[2,\infty) = \frac{7}{6} > V_2,$$

which leads to a contradiction. Hence, we can assume that the Voronoi region of a_1 contains only the set $\{1, 2\}$, or the set $\{1, 2, 3\}$. Suppose that the Voronoi region of a_1 contains only the set $\{1, 2\}$. Then, the Voronoi region of a_2 contains $\{3, 4, 5, \cdots\}$ yielding

$$V_2 = Er[1, 2] + Er[3, \infty) = \frac{2}{3} > V_2,$$

which gives a contradiction. Hence, we can conclude that the Voronoi region of a_1 contains only the set $\{1, 2, 3\}$, and the Voronoi region of a_2 contains the set $\{j : j \geq 4\}$ yielding

$$a_1 = Av[1, 3] \text{ and } a_2 = Av[4, \infty) \text{ with quantization error } V_2 = Er[1, 3] + Er[4, \infty) = \frac{17}{28}.$$

Thus, the proof of the proposition is complete. □

Proposition 4.1.2. *The set* $\{Av[1, 2], Av[3, 4], Av[5, \infty)\}$ *forms the optimal set of three-means with quantization error* $V_3 = \frac{1}{3}$.

Proof. The distortion error due to set $\beta := \{Av[1, 2], Av[3, 4], Av[5, \infty)\}$ is given by

$$V(P; \beta) = Er[1, 2] + Er[3, 4] + Er[5, \infty) = \frac{1}{3}.$$

Since V_3 is the quantization error for three-means, we have $V_3 \leq \frac{1}{3}$. Let $\alpha := \{a_1, a_2, a_3\}$ be an optimal set of three-means. Since the points in an optimal set are the conditional expectations in their own Voronoi regions, we have $1 \leq a_1 < a_2 < a_3 < \infty$. Suppose that the Voronoi region of a_1 contains the set $\{1, 2, 3\}$. Then,

$$V_3 \geq \sum_{j=1}^{3} f(j)(j - Av[1, 3])^2 = Er[1, 3] = \frac{5}{14} > \frac{1}{3} \geq V_3,$$

which leads to a contradiction. Hence, we can assume that the Voronoi region of a_1 contains only the set $\{1\}$, or the set $\{1, 2\}$. For the sake of contradiction, assume that the Voronoi region of a_1 contains only the set $\{1\}$. Then, the Voronoi region of a_2 must contain the element 2. Suppose that the Voronoi region of a_2 contains the set $\{2, 3, 4, 5\}$. Then,

$$V_3 \geq Er[2, 5] = \frac{181}{368} = 0.491848 > V_3,$$

which yields a contradiction. Assume that the Voronoi region of a_2 contains only the set $\{2, 3, 4\}$, and so the Voronoi region of a_3 contains the set $\{n : n \geq 5\}$. Then, the distortion error is

$$V_3 = Er[2, 4] + Er[5, \infty) = \frac{9}{22} = 0.409091 > V_3,$$

which gives a contradiction. Next, assume that the Voronoi region of a_2 contains only the set $\{2, 3\}$, and so the Voronoi region of a_3 contains the set $\{n : n \geq 4\}$. Then, the distortion error is

$$V_3 = Er[2, 3] + Er[4, \infty) = \frac{7}{20} > V_3,$$

which leads to a contradiction. Finally, assume that the Voronoi region of a_2 contains only the set $\{2\}$, and so the Voronoi region of a_3 contains the set $\{n : n \geq 3\}$. Then, the distortion error is

$$V_3 = Er[3, \infty) = \frac{1}{2} > V_3,$$

which gives a contradiction. Thus, we can conclude that the Voronoi region of a_1 contains only the set $\{1, 2\}$. Then, the Voronoi region of a_2 must contain the element 3. Suppose that the Voronoi region of a_2 contains the set $\{3, 4, 5, 6\}$. Then,

$$V_3 \geq \sum_{j=1}^{2} f(j)(j - Av[1, 2])^2 + \sum_{j=3}^{6} f(j)(j - Av[3, 6])^2 = Er[1, 2] + Er[3, 6] = \frac{59}{160} = 0.36875 > V_3,$$

which yields a contradiction. Assume that the Voronoi region of a_2 contains only the set $\{3, 4, 5\}$, and so the Voronoi region of a_3 contains the set $\{n : n \geq 6\}$. Then, the distortion error is

$$V_3 = Er[1, 2] + Er[3, 5] + Er[6, \infty) = \frac{29}{84} = 0.345238 > V_3,$$

which gives a contradiction. Next, assume that the Voronoi region of a_2 contains only the element 3, and so the Voronoi region of a_3 contains the set $\{n : n \geq 4\}$. Then, the distortion error is

$$V_3 = Er[1, 2] + Er[4, \infty) = \frac{5}{12} = 0.416667 > V_3,$$

which yields a contradiction. Hence, we can conclude that the Voronoi region of a_2 contains only the set $\{3, 4\}$ yielding $a_1 = Av[1, 2]$, $a_2 = Av[3, 4]$, and $a_3 = Av[5, \infty)$ with quantization error $V_3 = \frac{1}{3}$. Thus, the proof of the proposition is complete. □

Proposition 4.1.3. *The sets* $\{1, 2, Av[3, 4], Av[5, \infty)\}$ *forms the optimal sets of four-means with quantization error* $V_4 = \frac{1}{6}$.

Proof. The distortion error due to set $\beta := \{1, 2, Av[3, 4], Av[5, \infty)\}$ is given by

$$V(P; \beta) = Er[3, 4] + Er[5, \infty) = \frac{1}{6}.$$

Since V_4 is the quantization error for four-means, we have $V_4 \leq \frac{1}{6}$. Let $\alpha := \{a_1, a_2, a_3, a_4\}$ be an optimal set of four-means. Since the points in an optimal set are the conditional expectations in their own Voronoi regions, we have $1 \leq a_1 < a_2 < a_3 < a_4 < \infty$. Clearly, the Voronoi region of a_1 contains the point 1. Suppose that the Voronoi region of a_1 contains the set $\{1, 2, 3\}$. Then,

$$V_3 \geq \sum_{j=1}^{3} f(j)(j - Av[1, 3])^2 = Er[1, 3] = \frac{5}{14} > \frac{1}{6} \geq V_4,$$

which leads to a contradiction. Hence, we can assume that the Voronoi region of a_1 contains only the set $\{1\}$, or the set $\{1, 2\}$. Suppose that the Voronoi region of a_1 contains only the set $\{1, 2\}$. Then, the remaining elements of the set of natural numbers are contained in the union of the Voronoi regions of a_2, a_3 and a_4. Notice that the total distortion error contributed by the points a_2, a_3 and a_4 are positive. Hence,

$$V_4 > \text{distortion error contributed by the point } a_1 = Er[1, 2] = \frac{1}{6} = V_4,$$

which leads to a contradiction. Hence, the Voronoi region of a_1 cannot contain $\{1, 2\}$, i.e., the Voronoi region of a_1 contains only set $\{1\}$, i.e., $a_1 = 1$. Then, the Voronoi region of a_2 must contain 2. Suppose that the Voronoi region of a_2 contains the set $\{2, 3, 4\}$. Then,

$$V_4 \geq Er[2, 4] = \frac{25}{88} > V_4,$$

which leads to a contradiction. Hence, we can assume that the Voronoi region of a_2 contains only the set $\{2\}$, or the set $\{2, 3\}$. Suppose that the Voronoi region of a_2 contains only the set $\{2, 3\}$. Assume that the Voronoi region of a_3 contains the set $\{4, 5, 6, 7\}$. Then,

$$V_4 \geq Er[2, 3] + Er[4, 7] = \frac{193}{960} > V_4,$$

which leads to a contradiction. Hence, we can assume that the Voronoi region of a_3 contains only the set $\{4\}$, $\{4, 5\}$, or $\{4, 5, 6\}$. Suppose that the Voronoi region of a_3 contains only the set $\{4, 5, 6\}$. Then, the Voronoi region of a_4 contains the set $\{n : n \geq 7\}$. Then,

$$V_4 = Er[2, 3] + Er[4, 6] + Er[7, \infty) = \frac{53}{280} > V_4,$$

which is a contradiction. Suppose that the Voronoi region of a_3 contains only the set $\{4, 5\}$. Then, the Voronoi region of a_4 contains the set $\{n : n \geq 6\}$. Then,

$$V_4 = Er[2, 3] + Er[4, 5] + Er[6, \infty) = \frac{11}{60} > V_4,$$

which leads to a contradiction. Suppose that the Voronoi region of a_3 contains only the set $\{4\}$. Then, the Voronoi region of a_4 contains the set $\{n : n \geq 5\}$. Then,

$$V_4 = Er[2,3] + Er[5,\infty) = \frac{9}{40} > V_4,$$

which leads to a contradiction. Thus, we see that if the Voronoi region of a_2 contains only the set $\{2,3\}$, then a contradiction arises. Hence, we can conclude that the Voronoi region of a_2 contains only the set $\{2\}$, in other words, we have $a_2 = 2$. Then, the Voronoi region of a_3 contains the set $\{3\}$. Suppose that the Voronoi region of a_3 contains the set $\{3,4,5,6\}$, then as before we see a contradiction arises. Hence, the Voronoi region of a_3 contains only the set $\{3\}$, $\{3,4\}$, or the set $\{3,4,5\}$. Notice that if the Voronoi region of a_3 contains only the set $\{3,4,5\}$, then the Voronoi region of a_4 contains the set $\{n : n \geq 6\}$, and if the Voronoi region of a_3 contains only the set $\{3\}$, then the Voronoi region of a_4 contains the set $\{n : n \geq 4\}$. In either of the cases, proceeding as before, we see that a contradiction arises. Hence, we can conclude that the Voronoi region of a_3 contains only the set $\{3,4\}$. Hence, the Voronoi region of a_4 contains $\{n : n \geq 5\}$. Thus, we have

$$a_1 = 1, a_2 = 2, a_3 = [3,4], \text{ and } a_4 = [5,\infty) \text{ with } V_4 = \frac{1}{6}.$$

Thus, the proof of the proposition is complete. □

Proposition 4.1.4. *The sets* $\{1,2,\cdots,n-3, Av[n-2,n-1], Av[n,n+1], Av[n+2,\infty)\}$ *and* $\{1,2,\cdots,n-3,n-2, Av[n-1,n], Av[n+1,\infty)\}$ *form the optimal sets of n-means for all $n \geq 5$ with the quantization error* $V_n = \frac{2^{3-n}}{3}$.

Proof. The proof follows by Theorem 1.3 and Conjecture 1.5 under the assumption that Conjecture 1.5 is true. □

4.2. **Optimal quantization for P when** $(p_1,p_2,p_3,p_4,\cdots) = (\frac{1}{2^3}, \frac{1}{2^2}, \frac{1}{2}, \frac{1}{2^4}, \frac{1}{2^5}, \cdots)$**.** Let us give the results in the following propositions.

Proposition 4.2.1. *The optimal set of two-means is given by* $\{Av[1,3], Av[4,\infty)\}$ *with quantization error* $V_2 = \frac{5}{7}$.

Proof. We see that $Av[1,3] = \frac{17}{7}$, and $Av[4,\infty) = 5$. Since $3 < \frac{1}{2}(\frac{17}{7}+5) = \frac{26}{7} < 4$, the distortion error due to the set $\beta := \{\frac{17}{7}, 5\}$ is given by

$$V(P;\beta) = Er[1,3] + Er[4,\infty) = \frac{5}{7}.$$

Since V_2 is the quantization error for two-means, we have $V_2 \leq \frac{5}{7}$. Let $\alpha := \{a_1, a_2\}$ be an optimal set of two-means such that $a_1 < a_2$. Since the points in an optimal set are the conditional expectations in their own Voronoi regions, we have $1 \leq a_1 < a_2 < \infty$. Notice that the Voronoi region of a_1 must contain 1. Suppose that the Voronoi region of a_1 contains the set $\{1,2,3,4,5\}$. Then,

$$V_2 \geq Er[1,5] = \frac{393}{496} > V_2,$$

which yields a contradiction. Thus, we can conclude that the Voronoi region of a_1 does not contain the point 5. Suppose that the Voronoi region of a_1 contains only the $\{1,2,3,4\}$, and so the Voronoi region of a_2 contains the set $\{n : n \geq 5\}$. Then, we have

$$V_2 = Er[1,4] + Er[5,\infty) = \frac{11}{15} > V_2,$$

which leads to a contradiction. Similarly, we can show that if the Voronoi region of a_1 contains only the set $\{1\}$, or the set $\{1,2\}$, then we get a contradiction. Hence, we can assume that the Voronoi region of a_1 contains only the set $\{1,2,3\}$, and so the Voronoi region of a_2 contains only the set $\{n : n \geq 4\}$. Thus, we have

$$a_1 = Av[1,3],\ a_2 = Av[4,\infty) \text{ with } V_2 = \frac{5}{7}.$$

Thus, the proof of the proposition is complete. □

Proposition 4.2.2. *The sets $\{Av[1,2], Av[3,4], Av[5,\infty)\}$ forms the optimal sets of three-means with quantization error $V_3 = \frac{19}{72}$.*

Proof. The distortion error due to set $\beta := \{Av[1,2], Av[3,4], Av[5,\infty)\}$ is given by

$$V(P;\beta) = Er[1,2] + Er[3,4] + Er[5,\infty) = \frac{19}{72}.$$

Since V_3 is the quantization error for three-means, we have $V_3 \leq \frac{19}{72} = 0.263889$. Let $\alpha := \{a_1, a_2, a_3\}$ be an optimal set of three-means. Since the points in an optimal set are the conditional expectations in their own Voronoi regions, we have $1 \leq a_1 < a_2 < a_3 < \infty$. Suppose that the Voronoi region of a_1 contains the set $\{1,2,3\}$. Then,

$$V_3 \geq Er[1,3] = \frac{13}{28} > \frac{19}{72} \geq V_3,$$

which leads to a contradiction. Hence, we can assume that the Voronoi region of a_1 contains only the set $\{1\}$, or the set $\{1,2\}$. Suppose that the Voronoi region of a_1 contains only the set $\{1\}$. In this case, the Voronoi region of a_2 must contain the element 2. Suppose that the Voronoi region of a_2 contains the set $\{2,3,4\}$. Then,

$$V_3 \geq Er[2,4] = \frac{7}{26} = 0.269231 > V_3,$$

which yields a contradiction. Assume that the Voronoi region of a_2 contains only the set $\{2,3\}$, and so the Voronoi region of a_3 contains the set $\{n : n \geq 4\}$. Then, the distortion error is

$$V_3 = Er[2,3] + Er[4,\infty) = \frac{5}{12} > V_3,$$

which leads to a contradiction. Finally, assume that the Voronoi region of a_2 contains only the set $\{2\}$, and so the Voronoi region of a_3 contains the set $\{n : n \geq 3\}$. Then, the distortion error is

$$V_3 = Er[3,\infty) = \frac{13}{20} > V_3,$$

which leads to a contradiction. Hence, we can assume that the Voronoi region of a_1 contains only the set $\{1,2\}$. Then, the Voronoi region of a_2 must contain the set $\{3\}$. Suppose that the Voronoi region of a_2 contains the set $\{3,4,5,6\}$. Then,

$$V_3 \geq Er[1,2] + Er[3,6] = \frac{151}{416} > V_3,$$

which yields a contradiction. Assume that the Voronoi region of a_2 contains only the set $\{3,4,5\}$, and so the Voronoi region of a_3 contains the set $\{n : n \geq 6\}$. Then, the distortion error is

$$V_3 = Er[1,2] + Er[3,5] + Er[6,\infty) = \frac{35}{114} > V_3,$$

which gives a contradiction. Next, assume that the Voronoi region of a_2 contains only the element 3, and so the Voronoi region of a_3 contains the set $\{n : n \geq 4\}$. Then, the distortion error is

$$V_3 = Er[1,2] + Er[4,\infty) = \frac{1}{3} > V_3,$$

which yields a contradiction. Hence, we can conclude that the Voronoi region of a_2 contains only the set $\{3,4\}$ yielding $a_1 = Av[1,2]$, $a_2 = Av[3,4]$, and $a_3 = Av[5,\infty)$ with quantization error $V_3 = \frac{19}{72}$. ☐

Proposition 4.2.3. *The sets $\{Av[1,2], 3, Av[4,5], Av[6,\infty)\}$ forms the optimal sets of four-means with quantization error $V_4 = \frac{1}{6}$.*

Proof. The distortion error due to set $\beta := \{Av[1,2], 3, Av[4,5], Av[6,\infty)\}$ is given by

$$V(P;\beta) = Er[1,2] + Er[4,5] + Er[6,\infty) = \frac{1}{6}.$$

Since V_4 is the quantization error for four-means, we have $V_4 \leq \frac{1}{6} = 0.166667$. Let $\alpha := \{a_1, a_2, a_3, a_4\}$ be an optimal set of four-means. Since the points in an optimal set are the conditional expectations in their own Voronoi regions, we have $1 \leq a_1 < a_2 < a_3 < a_4 < \infty$. Suppose that the Voronoi region of a_1 contains the set $\{1, 2, 3\}$. Then,

$$V_4 \geq Er[1, 3] = \frac{13}{28} > V_4,$$

which leads to a contradiction. Hence, we can assume that the Voronoi region of a_1 contains only the set $\{1\}$, or the set $\{1, 2\}$. Suppose that the Voronoi region of a_1 contains only the set $\{1\}$. In this case, the Voronoi region of a_2 must contain the element 2. Suppose that the Voronoi region of a_2 contains the set $\{2, 3, 4\}$. Then,

$$V_4 \geq Er[2, 4] = \frac{7}{26} = 0.269231 > V_4,$$

which yields a contradiction. Assume that the Voronoi region of a_2 contains only the set $\{2, 3\}$. Then, notice that the Voronoi regions of a_3 and a_4 contain all the elements $\{n : n \geq 4\}$. Thus, the total distortion error contributed by a_3 and a_4 must be positive. This leads to the fact that

$$V_4 > Er[2, 3] = \frac{1}{6} \geq V_4,$$

which gives a contradiction. Assume that the Voronoi region of a_2 contains only the set $\{2\}$. Then, as before, we see that a contradiction arises. Hence, we can assume that the Voronoi region of a_1 contains only the set $\{1, 2\}$. Then, the Voronoi region of a_2 must contain 3. If the Voronoi region of a_2 contains more points using the similar arguments as before, we can show that a contradiction arises. Hence, we can conclude that $a_2 = 3$. Again, using the similar arguments, we can show that the Voronoi region of a_3 contains only the set $\{4, 5\}$, and the Voronoi region of a_4 contains only the set $\{n : n \geq 6\}$. Thus, we have

$$a_1 = Av[1, 2], a_2 = 3, a_3 = Av[4, 5], \text{ and } a_4 = Av[6, \infty) \text{ with quantization error } V_4 = \frac{1}{6}.$$

Thus, the proof of the proposition is complete. □

Proposition 4.2.4. *The sets* $\{1, 2, 3, Av[4, 5], Av[6, \infty)\}$ *forms the optimal sets of five-means with quantization error* $V_5 = \frac{1}{12}$.

Proof. The distortion error due to set $\beta := \{1, 2, 3, Av[4, 5], Av[6, \infty)\}$ is given by

$$V(P; \beta) = Er[4, 5] + Er[6, \infty) = \frac{1}{12}.$$

Since V_5 is the quantization error for five-means, we have $V_5 \leq \frac{1}{12} = 0.0833333$. Let $\alpha := \{a_1, a_2, a_3, a_4, a_5\}$ be an optimal set of five-means such that $a_1 < a_2 < a_3 < a_4 < a_5$. Since the points in an optimal set are the conditional expectations in their own Voronoi regions, we have $1 \leq a_1 < a_2 < a_3 < a_4 < a_5 < \infty$. Clearly, the Voronoi region of a_1 contains the point 1. For the sake of contradiction, assume that the Voronoi region of a_1 contains the set $\{1, 2, 3\}$. Then,

$$V_5 \geq Er[1, 3] = \frac{13}{28} > V_5,$$

which is a contradiction. Next, assume that the Voronoi region of a_1 contains only the set $\{1, 2\}$. Then, notice that the union Voronoi regions of a_2, a_3, a_4, and a_5 contain all the elements $\{n : n \geq 3\}$. Hence, we must have

$$V_5 > Er[1, 2] = \frac{1}{12} \geq V_5,$$

which is a contradiction. Hence, we can conclude that the Voronoi region of a_1 contains only the element 1, i.e., $a_1 = 1$. Clearly, the Voronoi region of a_2 contains the element 2. Suppose the Voronoi region of a_2 contains the set $\{2, 3\}$. Then, we have

$$V_5 \geq Er[2, 3] = \frac{1}{6} > V_5,$$

which give a contradiction. Hence, the Voronoi region of a_2 contains only the element 2, i.e., $a_2 = 2$. Similarly, we can show that $a_3 = 3$. The rest of the proof follows in the similar lines as given in

Proposition 4.1.3. Thus, we see that $a_4 = Av[4,5]$ and $a_5 = Av[6,\infty)$ with quantization error $V_5 = \frac{1}{12}$. Thus, the proof of the proposition is complete. □

Proposition 4.2.5. *The sets* $\{1, 2, \cdots, n-3, Av[n-2, n-1], Av[n, n+1], Av[n+2, \infty)\}$ *and* $\{1, 2, \cdots, n-3, n-2, Av[n-1, n], Av[n+1, \infty)\}$ *form the optimal sets of n-means for all $n \geq 6$ with the quantization error* $V_n = \frac{2^{3-n}}{3}$.

Proof. The proof follows by Theorem 1.3 and Conjecture 1.5 under the assumption that Conjecture 1.5 is true. □

5. Observation and Remarks

In Conjecture 1.5 the probability measure P is defined as $P := \sum_{j=1}^{k-1} p_j\delta_j + \sum_{j=k}^{\infty} \frac{1}{2^j}\delta_j$, where $(p_1, p_2, \cdots, p_{k-1})$ is a permutation of the set $\{\frac{1}{2}, \frac{1}{2^2}, \cdots, \frac{1}{2^{k-1}}\}$, where $k \in \mathbb{N}$ with $k \geq 2$. If $P := \sum_{j=1}^{k-1} p_j\delta_j + \sum_{j=k}^{\infty} \frac{1}{2^j}\delta_j$, and $(p_1, p_2, \cdots, p_{k-1})$ is not a permutation of the set $\{\frac{1}{2}, \frac{1}{2^2}, \cdots, \frac{1}{2^{k-1}}\}$, then Conjecture 1.5 is not true. In this regard, we give the following proposition.

Proposition 5.1. *For the probability measure P given by* $P := \frac{149}{200}\delta_1 + \frac{1}{200}\delta_2 + \sum_{j=3}^{\infty} \frac{1}{2^j}\delta_j$ *the optimal set of five-means is given by*

$$\{1, Av[2,3], 4, Av[5,6], Av[7,\infty)\}, \text{ or } \{1, Av[2,3], Av[4,5], Av[6,7], Av[8,\infty)\}$$

with quantization error $V_5 = \frac{29}{624}$.

Proof. The distortion error due to set $\beta := \{1, Av[2,3], 4, Av[5,6], Av[7,\infty)\}$ is given by

$$V(P;\beta) = Er[2,3] + Er[5,6] + Er[7,\infty) = \frac{29}{624}.$$

Since V_5 is the quantization error for five-means, we have $V_5 \leq \frac{29}{624} = 0.0464744$. Let us assume that $\alpha := \{a_1, a_2, a_3, a_4, a_5\}$ is an optimal set of five-means such that $a_1 < a_2 < a_3 < a_4 < a_5$. Since the points in an optimal set are the conditional expectations in their own Voronoi regions, we have $1 \leq a_1 < a_2 < a_3 < a_4 < a_5 < \infty$. Clearly, the Voronoi region of a_1 contains the point 1. For the sake of contradiction, assume that the Voronoi region of a_1 contains the set $\{1, 2, 3\}$. Then,

$$V_5 \geq Er[1,3] = \frac{7537}{17500} = 0.430686 > V_5,$$

which is a contradiction. Hence, we can assume that the Voronoi region of a_1 contains only the set $\{1\}$ or the set $\{1, 2\}$. Suppose that the Voronoi region of a_1 contains only the set $\{1, 2\}$. Then, the Voronoi region of a_2 must contain the element 3. Suppose that the Voronoi region of a_2 contains the set $\{3, 4\}$. Then,

$$V_5 \geq Er[1,2] + Er[3,4] = \frac{1399}{30000} = 0.0466333 > V_5,$$

which leads to a contradiction. Hence, we can assume that the Voronoi region of a_2 contain only the element 3, i.e., $a_2 = 3$. Then, the union of the Voronoi regions of a_3, a_4, a_5 contains the set $\{4, 5, 6, \cdots\}$ with associated probability $\frac{1}{2^j}$ for each $j \in \{4, 5, 6, \cdots\}$. Hence, using the similar lines as described in the proof of Theorem 1.3, we can show that

$$\{a_3, a_4, a_5\} \text{ equals the set } \{4, Av[5,6], Av[7,\infty)\}, \text{ or } \{Av[4,5], Av[6,7], Av[8,\infty)\} \tag{2}$$

with the quantization error

$$V_5 = Er[1,2] + Er[5,6] + Er[7,\infty) = \frac{1399}{30000} = 0.0466333 > V_5,$$

which leads to a contradiction. Hence, we can assume that the Voronoi region of a_1 contains only the element 1, i.e., $a_1 = 1$. Then, the Voronoi region of a_2 must contain 2. Suppose that the Voronoi region of a_2 contains the set $\{2, 3, 4\}$. Then,

$$V_5 \geq Er[2,4] = \frac{31}{616} > V_5,$$

which leads to a contradiction. Hence, the Voronoi region of a_2 contains only the set $\{2\}$, or $\{2, 3\}$. Suppose that the Voronoi region of a_2 contains only the set $\{2\}$, i.e., $a_2 = 2$. Then, as $a_1 = 1$, $a_2 = 2$, using the similar lines as described in the proof of Theorem 1.3, we can show that $a_3 = 3, a_4 = Av[4, 5], a_5 = Av[6, \infty)$; or $a_3 = Av[3, 4], a_4 = Av[5, 6], a_5 = Av[7, \infty)$ with quantization error $V_5 = \frac{1}{12} > V_5$, which is a contradiction. Hence, we can assume that the Voronoi region of a_2 contains only the set $\{2, 3\}$. Again, using the similar lines as described in the proof of Theorem 1.3, we can show that $\{a_3, a_4, a_5\}$ equals the set $\{4, Av[5, 6], Av[7, \infty)\}$; or $\{Av[4, 5], Av[6, 7], Av[8, \infty)\}$. Thus, we conclude that the optimal set of five-means is either $\{1, Av[2, 3], 4, Av[5, 6], Av[7, \infty)\}$ or $\{1, Av[2, 3], Av[4, 5], Av[6, 7], Av[8, \infty)\}$ with quantization error $V_5 = \frac{29}{624}$. This completes the proof of the proposition. □

Remark 5.2. Proposition 5.1 implies that Conjecture 1.5 is not true for an arbitrary probability distribution $(p_1, p_2, p_3, \cdots)$ associated with the set of positive integers $\mathbb{N}$.

Remark 5.3. Conjecture 1.5 is verified by two examples given in Subsection 3.2 and Subsection 3.3. We still could not give a general proof of the conjecture. It will be worthwhile to investigate the general proof of the conjecture.

References

[AW] E.F. Abaya and G.L. Wise, *Some remarks on the existence of optimal quantizers*, Statistics & Probability Letters, Volume 2, Issue 6, December 1984, Pages 349-351.

[DFG] Q. Du, V. Faber and M. Gunzburger, *Centroidal Voronoi Tessellations: Applications and Algorithms*, SIAM Review, Vol. 41, No. 4 (1999), pp. 637-676.

[DR] C.P. Dettmann and M.K. Roychowdhury, *An algorithm to compute CVTs for finitely generated Cantor distributions*, Southeast Asian Bulletin of Mathematics (2021) 45: 173-188.

[GG] A. Gersho and R.M. Gray, *Vector quantization and signal compression*, Kluwer Academy publishers: Boston, 1992.

[GKL] R.M. Gray, J.C. Kieffer and Y. Linde, *Locally optimal block quantizer design*, Information and Control, 45 (1980), pp. 178-198.

[GL1] A. György and T. Linder, *On the structure of optimal entropy-constrained scalar quantizers*, IEEE transactions on information theory, vol. 48, no. 2, February 2002.

[GL2] S. Graf and H. Luschgy, *Foundations of quantization for probability distributions*, Lecture Notes in Mathematics 1730, Springer, Berlin, 2000.

[GN] R.M. Gray and D.L. Neuhoff, *Quantization*, IEEE Transactions on Information Theory, Vol. 44, No. 6, October 1998, 2325-2383.

[P] D. Pollard, *Quantization and the Method of k-Means*, IEEE Transactions on Information Theory, 28 (1982), 199-205.

[R1] M.K. Roychowdhury, *Quantization and centroidal Voronoi tessellations for probability measures on dyadic Cantor sets*, Journal of Fractal Geometry, 4 (2017), 127-146.

[Z1] P.L. Zador, *Asymptotic Quantization Error of Continuous Signals and the Quantization Dimension*, IEEE Transactions on Information Theory, 28 (1982), 139-149.

[Z2] R. Zam, *Lattice Coding for Signals and Networks: A Structured Coding Approach to Quantization, Modulation, and Multiuser Information Theory*, Cambridge University Press, 2014.

School of Mathematical and Statistical Sciences, University of Texas Rio Grande Valley, 1201 West University Drive, Edinburg, TX 78539-2999, USA.

Email address: {[1]juan.gomez15, [2]haily.martinez01, [3]mrinal.roychowdhury}@utrgv.edu

Email address: {[4]alexis.salazar01, [5]daniel.vallez01}@utrgv.edu