The Annals of Mathematical Statistics
1970, Vol. 41, No. 4, 1249–1259

BOUNDS AND ASYMPTOTES FOR THE PERFORMANCE[1] OF MULTIVARIATE QUANTIZERS

BY PETER ELIAS

Massachusetts Institute of Technology

1. Introduction. Let $\mathbf{x} = (x_1, x_2, \cdots, x_N)$ be a vector-valued random variable with probability measure μ defined on the Lebesgue-measurable subsets of N-dimensional Euclidean space E^N. Let $\{R_i\}$, $1 \leqq i \leqq K$, be a set of K Lebesgue-measurable disjoint subsets of E^N, with

$$\sum_{i=1}^{K} \mu(R_i) = 1. \tag{1.1}$$

We define the K-region quantizer Q with quantization regions R_i as a function mapping the portion of E^N covered by the union of the R_i onto the integers 1 to K, given by

$$Q(\mathbf{x}) = i \qquad \text{for} \quad \mathbf{x} \in R_i. \tag{1.2}$$

Such a quantizer may be used as a model for the grouping of N-variate data (statistics), the quantization of signals (communications engineering) and analog-digital conversion (data-processing). It maps each $\mathbf{x}$ into the integer index i, $1 \leqq i \leqq K$, which labels the region R_i in which $\mathbf{x}$ falls, and saves only the value of i for further processing.

Quantization simplifies the handling of data, but introduces an error in the representation of $\mathbf{x}$, since $\mathbf{x}$ must be estimated by some function $\hat{\mathbf{x}}(i)$ of $Q(\mathbf{x}) = i$ alone. The first exploration of a quantization problem seems to be due to Sheppard [7], who analyzed the effect of quantization error on the estimate of the variance of the distribution μ of a scalar random variable $x(N = 1)$, assuming a smooth μ and equal intervals for the R_i.

Panter and Dite [6], also for $N = 1$, use the mean square value of the difference between x and its estimate $\hat{x}(Q(x))$

$$\overline{[x - \hat{x}(Q(x))]^2} = \int_{E^1} [x - \hat{x}(Q(x))]^2 \, d\mu \tag{1.3}$$

as a measure of performance. For fixed K they seek the minimum of this measure by moving the boundaries between the intervals R_i and by choosing the K values of $\hat{x}(i)$. For absolutely continuous μ with sufficiently smooth density $f = d\mu/dx$ they show that the minimum attainable value of their error measure is asymptotic in K to

$$\overline{[x - \hat{x}(Q(x))]^2} \sim (C_2/K^2)\{\int_{E^1} f^{\frac{1}{3}} \, d\lambda\}^3 \tag{1.4}$$

where λ is Lebesgue measure and C_2 is known. They credit this result to Pierre Aigrain.

Received August 21, 1969.

[1] Research sponsored by the Joint Services Electronics Program (Contract DA28-043-AMC-02536(E)).

Zador [8] extends the Panter and Dite results from the mean square performance measure to the mean rth power, $0 < r < \infty$. For $N = 1$ he gets asymptotically in K, for absolutely continuous μ with bounded f,

$$\overline{[x-\hat{x}(Q(x))]^r} \sim (C_r/K^r)\{\textstyle\int_{E^1} f^{1/(1+r)}\,d\lambda\}^{1+r}. \tag{1.5}$$

Here C_r is known. Zador also considers $N > 1$, and gets the general result

$$\overline{[\mathbf{x}-\hat{\mathbf{x}}(Q(\mathbf{x}))]^r} \sim (C_{Nr}/K^{r/N})\{\textstyle\int_{E^N} f^{N/(N+r)}\,d\lambda\}^{(N+r)/N}. \tag{1.6}$$

Again μ is absolutely continuous with bounded f, and λ is Lebesgue measure on E^N. The constants C_{Nr} are not known for $N > 1$.

We consider below the same general problem as Zador, i.e. the minimization of a measure of the error introduced by a quantizer Q having a fixed number K of quantizing regions by varying the shapes of the regions. However we use very different definitions of quantization error, and a different measure of that error. Our results, unlike Zador's, require that the probability measure μ have compact support, vanishing outside some cube in E^N. However we do not need the absolute continuity of μ or the boundedness of its density f, we obtain firm lower bounds which are also estimates asymptotic in K, there are no unknown constants for general N and r, and the results extend to $r = 0$ and $r = \infty$.

We have discussed the case $N = 1$ and given a more complete bibliography, including other quantization problems, in [1].

2. Measures of performance and results. Given a K-region quantizer Q with regions R_i, one measure of performance is derived from a definition of the quantization error or uncertainty in the nth coordinate when $\mathbf{x}$ is in R_i as the *width* $\Delta_i(x_n)$ of R_i in the nth coordinate:

$$\Delta_i(x_n) = \sup_{\mathbf{x}\in R_i}\{x_n\} - \inf_{\mathbf{x}\in R_i}\{x_n\}. \tag{2.1}$$

The performance of Q with respect to the probability measure μ can then be measured by $M_r(Q)$, the rth mean of the quantization errors $\Delta_i(x_n)$, averaging over the N coordinates of each R_i with equal weights and over the different R_i with weights $\mu(R_i)$:

$$M_r(Q) = \{\textstyle\sum_{i=1}^K \mu(R_i) N^{-1} \sum_{n=1}^N \Delta_i^r(x_n)\}^{1/r}, \qquad 0 < r < \infty. \tag{2.2}$$

This definition extends in the usual way (see e.g. Hardy et al [4] paragraph 2.3): $M_0(Q)$ is the geometric mean of the $\Delta_i(x_n)$, and $M_\infty(Q)$ is the largest of the $\Delta_i(x_n)$ which occurs with positive probability:

$$M_0(Q) = \exp\{\textstyle\sum_{i=1}^K \mu(R_i) N^{-1} \sum_{n=1}^N \log \Delta_i(x_n)\} \tag{2.3}$$

$$M_\infty(Q) = \max_{i:\mu(R_i)>0} \max_n \{\Delta_i(x_n)\}.$$

A second measure of performance is given by

$$M_r^*(Q) = \{\textstyle\sum_{i=1}^K \mu(R_i)\lambda(R_i)^{r/N}\}^{1/r}, \qquad 0 < r < \infty \tag{2.4}$$

Here $\lambda(R_i)$ is the Lebesgue measure (volume) of R_i, its Nth root is a typical linear dimension, and $M_r^*(Q)$ is the rth mean of the K Nth roots. M_r^* is like M_r but less coordinate-bound. The two are related by an inequality:

$$(2.5)\qquad \begin{aligned} M_r(Q) &\geqq \{\textstyle\sum_{i=1}^K \mu(R_i)\prod_{n=1}^N \Delta_i(x_n)^{r/N}\}^{1/r} \\ &\geqq \{\textstyle\sum_{i=1}^K \mu(R_i)\lambda(R_i)^{r/N}\}^{1/r} \\ &= M_r^*(Q). \end{aligned}$$

The first line applies the inequality of arithmetic and geometric means to the definition (2.2) of $M_r(Q)$: equality requires that each R_i having positive probability be equally wide in all dimensions, like a sphere or a cube. The second line notes that the volume of R_i is no larger than that of the smallest interval (rectangle aligned with the coordinate axes) which contains it: equality requires that the R_i of positive probability be cubes. The limiting cases $r = 0$, $r = \infty$ give the geometric mean and the essential supremum of the $\lambda(R_i)^{1/N}$, and the inequality (2.5) persists. So does the condition for equality at $r = \infty$. At $r = 0$, however, the arithmetic-geometric inequality becomes an identity and the condition for $M_0 = M_0^*$ is that the R_i of positive probability be rectangles, not necessarily cubes.

A third similar measure, the Nth root of the rth mean of the volumes, needs no separate discussion because it is trivially related to the second:

$$\{\textstyle\sum_{i=1}^K \mu(R_i)\lambda(R_i)^r\}^{1/Nr} = M_{Nr}^*(Q).$$

In the case $N = 1$, all three of these measures are identical.

M_r and M_r^* can be computed from the $\{R_i\}$ and the K numbers $\{\mu(R_i)\}$—i.e. from the grouped data and the decision boundaries between groups—with no other knowledge of μ. M_r is infinite unless all R_i with positive probability are bounded so that μ has compact support, which we henceforth assume. Let Ω_c be the support of μ—i.e. the smallest closed set in E^N which has probability 1. Then Ω_c is compact and can be covered by a finite closed cube. Let Ω be the smallest such cube. Then

$$(2.6)\qquad \mu(\Omega_c) = \mu(\Omega) = 1; \qquad \lambda(\Omega_c) \leqq \lambda(\Omega) < \infty.$$

The measure μ is otherwise arbitrary. It has a Lebesgue decomposition into a singular part μ_s, and an absolutely continuous part μ_a with density f:

$$(2.7)\qquad \mu = \mu_s + \mu_a, \qquad f = d\mu_a/d\lambda$$

where μ_s vanishes except on a set $\Omega_s \subseteq \Omega_c$ of Lebesgue measure $\lambda(\Omega_s) = 0$.

Since the quantizing regions $\{R_i\}$ together cover all those parts of E^N which have positive probability, the union of their closures must cover Ω_c. We now add the requirement that the union of their closures is Ω. Thus their characterization is

$$(2.8)\qquad R_i \cap R_j = \emptyset, \quad i \neq j; \qquad \textstyle\bigcup_{i=1}^K \bar{R}_i = \Omega,$$

and we can state the principal result.

THEOREM. *Let λ be Lebesgue measure on N-dimensional Euclidean space E^N. Let μ be any probability measure defined on the Lebesgue-measurable subsets of E^N,* *having compact support* Ω_c. *Let* μ_s, μ_a *and* f *be defined by the Lebesgue decomposition* (2.7) *of* μ. *Let* Ω *be the smallest closed cube in* E^N *which includes* Ω_c. *Let* $\{R_i\}$ *be any class of* K *subsets of* Ω *which satisfy* (2.8), *and let* Q *be a quantizer with* $\{R_i\}$ as *its quantizing regions.*

Then we have the following two results.

(I) *The performance measures* M_r *and* M_r^* *defined in* (2.2), (2.3) *and* (2.4) *are bounded below for any nonnegative real* r:

$$K^{1/N}M_r \geqq K^{1/N}M_r^* \geqq \{\textstyle\int_{\Omega_c} f^{N/(N+r)}\,d\lambda\}^{(N+r)/Nr} \qquad 0 < r < \infty \tag{2.9}$$

$$K^{1/N}M_0 \geqq K^{1/N}M_0^* \geqq 0 \qquad r = 0,\ \mu_s(\Omega) > 0 \tag{2.10}$$

$$K^{1/N}M_0 \geqq K^{1/N}M_0^* \geqq \exp\{-(1/N)\textstyle\int_{\Omega_c} f \log f\,d\lambda\} \qquad r = 0,\ \mu_s(\Omega) = 0 \tag{2.11}$$

$$K^{1/N}M_\infty \geqq K^{1/N}M_\infty^* \geqq \lambda(\Omega_c)^{1/N}. \qquad r = \infty \tag{2.12}$$

(II) *The four lower bounds of* (I) *can be approached for large* K *in all cases. Given a fixed* μ *and* r, *it is possible to construct a sequence* $\{Q_m\}$ *of quantizers,* Q_m *having* K_m *quantizing regions and* K_m *being an increasing sequence in* $m = 1, 2, \cdots$, *such that* $K_m^{1/N}M_r(Q_m)$, *and thus a fortiori* $K_m^{1/N}M_r^*(Q_m)$, *converges to the expression on the right corresponding to the given values of* r *and* $\mu_s(\Omega)$ *as* $m \to \infty$.

COMMENTS. Unfortunately no single procedure will generate the sequences of $\{Q_m\}$ required to prove (II) of the theorem in the four different cases. The proof therefore covers (I) and (II) for each case in turn. Before proceeding, three comments are in order.

First, all expressions in (I) are homogeneous of first degree in the length $\lambda(\Omega)^{1/N}$ of the side of the cube Ω. Where expedient the proof assumes $\lambda(\Omega) = 1$: the general case follows from the homogeneous dilation $x_n' = x_n\lambda(\Omega)^{1/N}$.

Second, decomposing the unit cube into $k^N = K$ half-open cubes of side $1/k$ for integer k gives a set $\{R_i\}$ with $K^{1/N}M_r = K^{1/N}M_r^* = 1$ for any μ on the unit cube. The theorem shows how much smaller than 1 the constant on the right can be made by optimum choice of $\{R_i\}$ so as to make $\lambda(R_i)$ small when $\mu(R_i)$ is large.

Third, the singular part μ_s of μ contributes to the lower bound only in (2.12) at $r = \infty$. There if e.g. $\mu_s = \mu$ is a distribution giving positive probability to points in the unit cube with rational coordinates, $\Omega_c = \Omega$ and $K^{1/N}M_\infty = 1$, though $K^{1/N}M_r \to 0$ with K for every finite r.

3. Proof. $0 < r < \infty$. The proof of (2.9) follows by elementary inequalities from the definition (2.4) of $M_r^*(Q)$ for $0 < r < \infty$. Let $p = N/(N+r)$, $p < 1$.

$$\begin{aligned}
[M_r^*(Q)]^r &= K\textstyle\sum_{i=1}^K K^{-1}\{\mu(R_i)\lambda(R_i)^{r/N}\} \\
&\geqq K\{\textstyle\sum_{i=1}^K K^{-1}(\mu(R_i)\lambda(R_i)^{r/N})^{N/(N+r)}\}^{(N+r)/N} \\
&= K^{-r/N}\{\textstyle\sum_{i=1}^K (\mu(R_i)/\lambda(R_i))^p\lambda(R_i)\}^{1/p} \\
&\geqq K^{-r/N}\{\textstyle\sum_{i=1}^K (1/\lambda(R_i)\int_{R_i} f\,d\lambda)^p\lambda(R_i)\}^{1/p} \\
&\geqq K^{-r/N}\{\textstyle\sum_{i=1}^K \int_{R_i} f^p\,d\lambda\}^{1/p} \\
&= K^{-r/N}\{\textstyle\int_\Omega f^p d\lambda\}^{1/p}.
\end{aligned} \tag{3.1}$$

The first inequality is Minkowsky's: equality requires that all terms in the sum be equal, i.e. that

$$\sigma_i = \mu(R_i)^p \lambda(R_i)^{1-p} \tag{3.2}$$

be equal to a constant $\sigma_i = \sigma$, independent of i. The second inequality introduces the density f of the absolutely continuous part μ_a of μ: equality requires that there be no other part, $\mu_a = \mu$, $\mu_s = 0$ in (2.7). The third follows from the convexity $\cap$ of the pth power, $0 < p < 1$: equality requires f to be piecewise constant, $f = \mu(R_i)/\lambda(R_i)$ almost everywhere in each R_i. (We use Gallager's notation [2], calling a function like x^2 "convex $\cup$", which is convex in the sense of Hardy et al [4], and a function like $1-x^2$ "convex $\cap$", which is concave in the sense of [4].)

With the inequality (2.5) between M_r and M_r^*, (3.1) proves (2.9) in the theorem. It is not possible for Q to meet all of the conditions for equality required by both (2.5) and (3.1) for an arbitrary μ. However, they can all be met or approximated so as to prove (II) of the theorem in this case.

We take $\lambda(\Omega) = 1$. For each integer m divide Ω into 2^{mN} cubes $\{Y_j\}$ of side 2^{-m}, and given $\varepsilon > 0$, subdivide each Y_j into L_j^N smaller cubes, each of side $2^{-m}L_j^{-1}$. These smaller cubes are the quantizing regions R_i of a quantizer Q_m. Then

$$\begin{aligned} K &= \textstyle\sum_{j=1}^{2^{mN}} L_j^N \\ \lambda(Y_j) &= 2^{-mN} \\ \lambda(R_i) &= 2^{-mN} L_j^{-N} \qquad \text{for} \quad R_i \subseteq Y_j. \end{aligned} \tag{3.3}$$

The integer L_j is selected so as to make approximately equal the σ_i in (3.2):

$$L_j = [L_0(\mu(R_i)/\lambda(R_i))^{1/(N+r)} + 1]. \tag{3.4}$$

where the square brackets denote the integer part of their argument. Then

$$(L_j/L_0)^{N+r} > \mu(R_i)/\lambda(R_i) \qquad \text{and} \tag{3.5}$$

$$\begin{aligned} L_j/L_0 &\leqq \{(\mu(R_i)/\lambda(R_i))^{1/(N+r)} + 1/L_0\}^{N+r} \\ &= \left\{\frac{1\cdot(\mu(R_i)/\lambda(R_i))^{1/(N+r)} + (1/L_0)\cdot 1^{1/(N+r)}}{1+(1/L_0)}\right\}^{N+r} (1+(1/L_0))^{N+r} \\ &\leqq \left\{\frac{\mu(R_i)/\lambda(R_i)+1/L_0}{1+(1/L_0)}\right\}(1+(1/L_0))^{N+r}, \end{aligned}$$

since the mean of order $1/(N+r)$ is less than the mean of order 1 (Hardy et al [4], Theorem 16). Given $\varepsilon > 0$, it is clearly possible to choose the integer $L_0 = L_0\,(m, \varepsilon)$ so large as to give

$$(L_j/L_0)^{N+r} < (\mu(R_i)/\lambda(R_i))\,(1+\varepsilon)+\varepsilon. \tag{3.6}$$

Since all R_i are cubes, equality holds in (2.5). Thus

$$
\begin{aligned}
M_r(Q_m)^r &= \textstyle\sum_{j=1}^{2^{mN}} \sum_{i:R_i \subseteq Y_j} \mu(R_i)\lambda(R_i)^{r/N} \\
&= \textstyle\sum_j \mu(Y_j) 2^{-mr} L_j^{-r} \\
&= \textstyle\sum_j (\mu(Y_j)/\lambda(Y_j)) 2^{-m(N+r)} L_j^{-r} \qquad (3.7)\\
&\leqq \textstyle\sum_j (L_j/L_0)^{N+r} 2^{-m(N+r)} L_j^{-r} \\
&= \{\textstyle\sum_j L_j^N\}^{-r/N} \{\textstyle\sum_j (L_j/L_0)^N \lambda(Y_j)\}^{(N+r)/N} \\
&\leqq K^{-r/N} \{\textstyle\sum_j (1+\varepsilon)\mu(Y_j)/\lambda(Y_j)+\varepsilon)^{N/(N+r)} \lambda(Y_j)\}^{(N+r)/N}
\end{aligned}
$$

where the second line uses $\lambda(R_i)$ and the third $\lambda(Y_j)$ from (3.3), the first inequality is (3.5), the second uses K from (3.3) and inequality (3.6).

Taking the pth power of both sides in (3.7) and moving the power of K to the left gives

$$
\begin{aligned}
\{K^{1/N} M_r(Q)\}^{rp} &\leqq \textstyle\sum_j ((1+\varepsilon)\mu(Y_j)/\lambda(Y_j)+\varepsilon)^p \lambda(Y_j) \\
&\leqq \textstyle\sum_j (1+\varepsilon)(\mu(Y_j)/\lambda(Y_j))^p \lambda(Y_j) + \varepsilon^p \sum_j \lambda(Y_j) \qquad (3.8)\\
&\leqq \textstyle\sum_j \lambda(Y_j)(\mu(Y_j)/\lambda(Y_j))^p + \varepsilon \sum_j \mu(Y_j)^p \lambda(Y_j)^{1-p} + \varepsilon^p \\
&\leqq \textstyle\int_\Omega \bar{g}_m{}^p d\lambda + \varepsilon + \varepsilon^p,
\end{aligned}
$$

using $(a+b)^p \leqq a^p + b^p$ $(p \leqq 1)$ in the second line, $(1+\varepsilon)^p \leqq 1+\varepsilon$ and $\lambda(\Omega) = 1$ in the third, and $\mu(\Omega) = 1$ and the Hölder inequality on the second sum in the third line. In the fourth line the step function $\bar{g}_m$ has been introduced to convert the sum over j to an integral over Ω:

$$
\bar{g}_m(\mathbf{x}) = \mu(Y_j)/\lambda(Y_j), \qquad \mathbf{x} \in Y_j. \qquad (3.9)
$$

Now the limit of $\bar{g}_m$ as $m \to \infty$ is the regular derivate of the Lebesgue–Stieltjes measure μ, which is equal to the density f of the absolutely continuous part μ_a of μ a.e. in Ω (see e.g. Munroe [5] paragraphs 41.3, 41.6). Since $p > 0$ $(r < \infty)$, the pth power is continuous and $\bar{g}_m{}^p$ also approaches f^p a.e. in Ω, and thus in measure ([5] paragraph 31.3). So given $\varepsilon > 0$, there is an $m_0(\varepsilon)$ so large that for $m > m_0$, the set $\Omega_\varepsilon = \{x : \bar{g}_m{}^p(\mathbf{x}) > f^p(\mathbf{x}) + \varepsilon\}$ has Lebesgue measure

$$
\lambda(\Omega_\varepsilon) < \varepsilon. \qquad (3.10)
$$

Then

$$
\begin{aligned}
\textstyle\int_\Omega \bar{g}_m{}^p d\lambda &\leqq \textstyle\int_{\Omega-\Omega_\varepsilon} (f^p+\varepsilon)\, d\lambda + \int_{\Omega_\varepsilon} \bar{g}_m{}^p d\lambda \\
&\leqq \textstyle\int_\Omega (f^p+\varepsilon)\, d\lambda + \lambda(\Omega_\varepsilon) \int_{\Omega_\varepsilon} (\bar{g}_m{}^p/\lambda(\Omega_\varepsilon))\, d\lambda \qquad (3.11)\\
&\leqq \textstyle\int_\Omega f^p\, d\lambda + \varepsilon + \lambda(\Omega_\varepsilon)\{\int_{\Omega_\varepsilon} \bar{g}_m\, d\lambda/\lambda(\Omega_\varepsilon)\}^p \\
&\leqq \textstyle\int_\Omega f^p\, d\lambda + \varepsilon + \lambda^{1-p}(\Omega_\varepsilon)
\end{aligned}
$$

using the definition of Ω_ε in the first line, the convexity $\cap$ of the pth power in the third, and the fact that $\bar{g}_m$ is a density, whose integral over $\Omega_\varepsilon \subseteq \Omega$ is $\leqq 1$.

Substituting (3.11) in (3.8) and using (3.10) to bound $\lambda(\Omega_\varepsilon)$ gives, for $m > m_0(\varepsilon)$.

$$\{K^{1/N} M_r(Q_m)\}^{rp} \leqq \int_\Omega f^p \, d\lambda + 2\varepsilon + \varepsilon^p + \varepsilon^{1-p},$$

which together with (2.9) proves the convergence stated in (II) of the theorem for $m \to \infty$ when $0 < r < \infty$, so $0 < p < 1$.

4. Proof. $r = 0$, $\mu_s > 0$. The inequality in (2.10) is trivial, since the geometric mean of nonnegative quantities is nonnegative. To prove (II) of the theorem in this case, let $\mu_s(\Omega) = a > 0$. Construct the cubes Y_j of side 2^{-m} as before. Given $\varepsilon = 2^{-cN} < a$, c an integer, it is possible to choose $m_0(\Omega)$ so large that for $m > m_0$ a finite union Ω_0 of the Y_j can approximate the singular set Ω_s of μ to within ε, in the symmetric difference sense, in both λ and μ: i.e.

$$\lambda(\Omega_s \Delta \Omega_0) < \varepsilon, \qquad \mu(\Omega_s \Delta \Omega_0) < \varepsilon \tag{4.1}$$

(see e.g. Halmos [3] page 58 problem (8)). Since $\lambda(\Omega_s) = 0$ and $\mu(\Omega_s) = \mu_s(\Omega) = a$, this implies that

$$\lambda(\Omega_0) < \varepsilon = 2^{-cN}, \qquad \mu(\Omega_0) \geqq \mu_s(\Omega_0) = b > a - \varepsilon > 0. \tag{4.2}$$

The quantizing regions R_i are the Y_j in $\Omega - \Omega_0$ and subdivisions of those Y_j in Ω_0: we define L_j, not by (3.4), but by

$$\begin{aligned} L_j &= 2^c, j \in J_0; \qquad J_0 = \{j : Y_j \subset \Omega_0\} \\ L_j &= 1, j \in J_1; \qquad J_1 = \{j : Y_j \subset (\Omega - \Omega_0)\}. \end{aligned} \tag{4.3}$$

Then the total number K of quantizing cubes is

$$\begin{aligned} K = \sum_{j=1}^{2^{mN}} L_j^N &= \sum_{j \in J_0} L_j^N + \sum_{j \in J_1} L_j^N \\ &= 2^{cN}(\lambda(\Omega_0)/\lambda(Y_j)) + 1 \cdot (\lambda(\Omega - \Omega_1)/\lambda(Y_j)) \\ &\leqq 2^{cN} 2^{mN} 2^{-cN} + 1 \cdot 2^{mN} \cdot 1 = 2 \cdot 2^{mN}, \end{aligned} \tag{4.4}$$

using (4.2) and $\lambda(\Omega - \Omega_1) \leqq \lambda(\Omega) = 1$ in the last line.

The geometric mean quantization error $M_0(Q)$ is the exponential of the mean logarithm of the quantization error. Using logarithms to the base 2, it can be bounded:

$$\begin{aligned} \log_2 M_0(Q) &= N^{-1} \sum_{i=1}^{K} \mu(R_i) \log_2 \lambda(R_i) \\ &= N^{-1} \mu(\Omega_0)(-Nc - Nm) + N^{-1} \mu(\Omega - \Omega_0)(-Nm) \\ &= -m - \mu(\Omega_0)c \leqq -m - bc \end{aligned} \tag{4.5}$$

from (4.2). Exponentiating and using the bound on K of (4.4),

$$K^{1/N} M_0(Q) \leqq 2^{1/N} \cdot 2^m \cdot 2^{-m} \cdot 2^{-bc} = 2^{(1/N) - bc} \tag{4.6}$$

which can be made arbitrarily small by taking c sufficiently large, proving the top line of (2.10).

5. Proof. $r = 0$, $\mu_s(\Omega) = 0$. The proof of (2.11) parallels that of (2.9), starting with the definition of the (geometric mean) quantizing error $M_0{}^*(Q)$ from (2.3):

$$M_0{}^*(Q) = \exp\{N^{-1}\textstyle\sum_{i=1}^K \mu(R_i)\log\lambda(R_i)\}$$

$$= \exp\left\{N^{-1}\sum_{i=1}^K \mu(R_i)\log\mu(R_i) - N^{-1}\sum_{i=1}^K\left(\frac{\mu(R_i)}{\lambda(R_i)}\log\frac{\mu(R_i)}{\lambda(R_i)}\right)\lambda(R_i)\right\}$$

$$(5.1)\qquad \geqq \exp\{-N^{-1}\log K\}\exp\left\{-N^{-1}\sum_{i=1}^K \frac{1}{\lambda(R_i)}\int_{R_i} f\,d\lambda \log\frac{1}{\lambda(R_i)}\int_{R_i} f\,d\lambda\,\lambda(R_i)\right\}$$

$$\geqq K^{-1/N}\exp\left\{-N^{-1}\sum_{i=1}^K\left(\frac{1}{\lambda(R_i)}\int_{R_i} f\log f\,d\lambda\right)\lambda(R_i)\right\}$$

$$= K^{-1/N}\exp\{-N^{-1}\textstyle\int_\Omega f\log f\,d\lambda\}.$$

The first inequality in (5.1) is the analog to the Minkowski inequality in (3.1) at $r = 0$: it is the entropy inequality for the K-term discrete probability distribution $\{\mu(R_i)\}: -\sum_{i=1}^K \mu(R_i)\log\mu(R_i) \leqq \log K$ and equality is attained only for equiprobable quantizers with

$$(5.2)\qquad \mu(R_i) = 1/K, \qquad 1 \leqq i \leqq K.$$

Also at this point we use the absolute continuity of μ for the first time (i.e. $\mu_s(\Omega) = 0$ in the second line of (2.8)), to replace $\mu(R_i)$ by the integral of the density f over R_i. The last inequality in (5.1) follows from the convexity $\cup$ of the function $x\log x$, which makes the function of the average smaller than the average of the function: equality requires that f be piecewise constant a.e. and equal to the averaged density function $\bar{f}$, a step-function given by

$$(5.3)\qquad \bar{f}(\mathbf{x}) = \mu(R_i)/\lambda(R_i) = (1/\lambda(R_i))\textstyle\int_{R_i} f\,d\lambda, \qquad \mathbf{x} \in R_i.$$

Thus (5.1) proves (2.11), and thus (I) of the theorem. To prove (II) of the theorem in this case requires the inductive construction of a sequence $\{Q_m\}$ of rectangular, equiprobable quantizers for an arbitrary absolutely continuous μ.

The quantizer Q_0 has $K = 1$ quantizing region, Ω itself, with $\mu(\Omega) = 1$. Q_{m+1} is constructed from Q_m by dividing each of the $K = 2^m$ intervals $\{R_i\}$ in Q_m (each a rectangle aligned with the coordinate axes) into two equiprobable intervals by cutting R_i with an $(N-1)$-flat orthogonal to the greatest dimension of R_i—i.e. orthogonal to a coordinate x_n so chosen that $\Delta_i(x_n) \geqq \Delta_i(x_t)$ for $1 \leqq t \leqq N$, in the notation of (2.1). By the absolute continuity of μ, such equiprobable division will always be possible and will give two quantizing intervals in Q_{m+1} each of which has a width in the coordinate x_n which is strictly positive and strictly less than $\Delta_i(X_n)$. Q_{m+1} thus has $K = 2^{m+1}$ equiprobable rectangular quantizing regions, each with $\lambda(R_i) > 0$ for $m < \infty$.

Let $R(m, x)$ denote that quantizing interval R_i in the quantizer Q_m which contains the point $\mathbf{x}$. In terms of the definition (1.2) of $Q(\mathbf{x})$, $R(m, \mathbf{x}) = R_{Q_m(\mathbf{x})}$. The change in notation is introduced to avoid the double subscript. By construction

$R(m+1, \mathbf{x}) \subset R(m, \mathbf{x})$ so there is a limiting interval $R(\infty, \mathbf{x})$ included in all $R(m, \mathbf{x})$. Then

$$\mu(R(m, \mathbf{x})) = 2^{-m} \tag{5.4}$$

$$\lambda(R(m, \mathbf{x})) > \lambda(R(\infty, x)) \geqq 0.$$

We define the set Ω_1:

$$\Omega_1 = \{\mathbf{x} : \lambda(R(\infty, \mathbf{x})) > 0\}. \tag{5.5}$$

For $\mathbf{x} \in \Omega_1$, by the definition (5.3) of $\bar{f}$ and the properties (5.4) and (5.5),

$$\lim_{m\to\infty} \bar{f}(\mathbf{x}) = \lim_{m\to\infty} \frac{\mu(R(m, \mathbf{x}))}{\lambda(R(m, \mathbf{x}))} \leqq \lim_{m\to\infty} \frac{2^{-m}}{\lambda(R(\infty, \mathbf{x}))} = 0 \tag{5.6}$$

and by the absolute continuity of μ and (5.4),

$$\int_{R(\infty,\mathbf{x})} f\, d\lambda = \lim_{m\to\infty} \int_{R(m,\mathbf{x})} f\, d\lambda = \lim_{m\to\infty} \mu(R(m, \mathbf{x})) = 0. \tag{5.7}$$

Since $R(\infty, \mathbf{x}) > 0$ for $\mathbf{x} \in \Omega_1$, the nonnegative function f must vanish almost everywhere in $R(\infty, \mathbf{x})$, and thus almost everywhere in Ω_1, which is a denumerable set of such intervals of positive Lebesgue measure. This and (5.6) give

$$\lim_{m\to\infty} \bar{f}(\mathbf{x}) = f(\mathbf{x}) = 0 \text{ a.e. in } \Omega_1. \tag{5.8}$$

For x in $\Omega - \Omega_1$, by (5.5) $\lambda(R(m, \mathbf{x})) \to 0$ as m increases. Since it cannot vanish for finite m, at least one width, say $\Delta_i(x_n)$ must decrease to zero in an infinite number of steps. But then so must all widths, since $\Delta_i(x_n)$ must be the largest width each time it decreases, by construction. Thus not only the volume but the diameter of $R(m, \mathbf{x}) \to 0$ for $\mathbf{x}$ in $\Omega - \Omega_1$, and if f is bounded its integral μ has a strong derivate.

$$\lim_{m\to\infty} \bar{f}(\mathbf{x}) = f(\mathbf{x}) \text{ a.e. in } \Omega - \Omega_1 \tag{5.9}$$

(see e.g. Munroe [5] 42.4.1).

If f is bounded, so is its local average $\bar{f}$ and so are $f \log f$ and $\bar{f} \log \bar{f}$. By (5.8) and (5.9) $\bar{f} \to f$ a.e. in Ω, and thus $\bar{f} \log \bar{f} \to f \log f$ a.e. there. So

$$\lim_{m\to\infty} \int_\Omega \bar{f} \log \bar{f}\, d\lambda = \int_\Omega f \log f\, d\lambda \tag{5.10}$$

for bounded f, by the Lebesgue dominated convergence theorem. Since the Q_m are rectangular and equiprobable, equality is attained in (2.5) and in the first inequality in (5.1). Using the definition (5.3) of $\bar{f}$ in (5.1) gives

$$K_m^{1/N} M_0(Q_m) = \exp\{-N^{-1} \int_\Omega \bar{f} \log \bar{f}\, d\lambda\} \tag{5.11}$$

and taking the limit $m \to \infty$, using (5.10), proves (II) of the theorem for bounded f.

If f is not bounded, note from the argument in (5.1) that

$$\int_\Omega \bar{f} \log \bar{f}\, d\lambda \leqq \int_\Omega f \log f\, d\lambda \tag{5.12}$$

so that only a converse inequality is needed to prove (5.10). Decomposing f into two nonnegative components $\bar{f} = \bar{f}_1 + \bar{f}_2$, gives

$$
\begin{aligned}
\int_\Omega \bar{f} \log \bar{f} \, d\lambda &= \int_\Omega \bar{f}_1 \log(\bar{f}_1 + \bar{f}_2) \, d\lambda + \int_\Omega \bar{f}_2 \log(\bar{f}_1 + \bar{f}_2) \, d\lambda \\
&\geqq \int_\Omega \bar{f}_1 \log \bar{f}_1 \, d\lambda + \int_\Omega \bar{f}_2 \log \bar{f}_2 \, d\lambda \qquad (5.13) \\
&\geqq \int_\Omega \bar{f}_1 \log \bar{f}_1 \, d\lambda + \left(\int_\Omega \bar{f}_2 \, d\lambda\right)\left(\log \int_\Omega \bar{f}_2 \, d\lambda\right)
\end{aligned}
$$

where the monotonicity of the logarithm is used in the second line and the convexity $\cup$ of $z \log z$ in the third. Next define $f_1 = \min(f, f_0)$, $f_2 = f - f_1$, $f = f_1 + f_2$, and their local averages, the step functions $\bar{f}_1$ and $\bar{f}_2$:

$$
\begin{aligned}
\bar{f}_1(\mathbf{x}) &= (1/\lambda(R_i)) \int_{R_i} f_1 \, d\lambda \geqq 0, \qquad \mathbf{x} \in R_i \\
\bar{f}_2(\mathbf{x}) &= (1/\lambda(R)) \int_{R_i} f_2 \, d\lambda \geqq 0, \qquad \mathbf{x} \in R_i \qquad (5.14) \\
\bar{f} &= \bar{f}_1 + \bar{f}_2,
\end{aligned}
$$

where f_0 is set so that, given a positive $\varepsilon < 1/e$,

$$\varepsilon > \int_\Omega f_2 \, d\lambda = \int_\Omega \bar{f}_2 \, d\lambda. \qquad (5.15)$$

Now taking limits as $m \to \infty$ in (5.13) and (5.14), and substituting from (5.15), gives

$$\int_\Omega f \log f \, d\lambda \geqq \lim_{m \to \infty} \int_\Omega \bar{f} \log \bar{f} \, d\lambda \geqq \lim_{m \to \infty} \int_\Omega \bar{f}_1 \log \bar{f}_1 \, d\lambda + \varepsilon \log \varepsilon. \qquad (5.16)$$

And f_1, and thus $\bar{f}_1$, is bounded by f_0. Thus precisely the argument of (5.6) to (5.9), with equalities replaced by $\leqq$ in (5.6) and (5.7) and f and $\bar{f}$ replaced by $f_1 (\leqq f)$ and $\bar{f}_1 (\leqq \bar{f})$, proves the analog of (5.10): i.e.

$$\lim_{m \to \infty} \int_\Omega \bar{f}_1 \log \bar{f}_1 \, d\lambda = \int_\Omega f_1 \log f_1 \, d\lambda. \qquad (5.17)$$

Substituting (5.17) in (5.16) and letting $\varepsilon \to 0$, the integral on the right in (5.17) approaches the integral of $f \log f$, by definition of the Lebesgue integral of an unbounded integrand, in the sense that either it converges or both diverge to $+\infty$ (both integrands are bounded below). Substituting the result in (5.11) proves (II) of the theorem for $r = 0$ for an arbitrary f, and thus an arbitrary absolutely continuous μ.

6. Proof. $r = \infty$. To prove (2.12), let Q be any K-region quantizer for a distribution μ of compact support Ω_c. Let $\{R_i\}$ be the quantizing regions of Q, and let U be the union of those R_i which have

$$\max_n \Delta_i(x_n) \leqq M_\infty(Q) \qquad (6.1)$$

and thus also have

$$\lambda(R_i) \leqq M_\infty(Q)^N. \qquad (6.2)$$

Now the closure $\bar{U}$ of U includes Ω_c. For if not there would be a neighborhood in Ω_c not covered by U, and thus a positive probability of choosing an $\mathbf{x}$ not in any of the R_i defined by (6.1), and thus a positive probability of a quantizing error greater than $M_\infty(Q)$, which contradicts the definition (2.3) of $M_\infty(Q)$. The total number K of quantizing regions must be at least as large as the number in $\bar{U}$. Thus by (6.2)

$$KM_{\infty}(Q)^N \geqq \lambda(\bar{U}) \geqq \lambda(\Omega_c), \tag{6.3}$$

which proves (2.12) by taking Nth roots.

To prove (II) of the theorem in this case, note that the open set $\Omega - \Omega_c$ can be approximated for each positive integer m by the union U_m of half-open disjoint cubes of side 2^{-m} and volume 2^{-mN}, each of which has its closure contained in $\Omega - \Omega_c$. Thus given $\varepsilon > 0$ it is possible to find m_0 so large that $m > m_0$ implies that the Lebesgue measure $\lambda(\Omega - \Omega_c - U_m) < \varepsilon$. Then the remainder of the 2^{mN} cubes of side 2^{-m} in Ω cover Ω_c and have Lebesgue measure bounded by

$$\lambda(\Omega - U_m) = \lambda(\Omega_c) + \lambda(\Omega - \Omega_c - U_m) \leqq \lambda(\Omega_c) + \varepsilon. \tag{6.4}$$

Thus their number is bounded above by

$$\frac{\lambda(\Omega_c) + \varepsilon}{2^{-mN}} = 2^{mN}\{\lambda(\Omega_c) + \varepsilon\}. \tag{6.5}$$

Now let the quantizing regions of Q_m consist of those cubes of side 2^{-m} which are not in U_m, together with the region U_m itself. Then Q_m has $M_{\infty}(Q_m)$, the essential supremum of its quantizing interval widths, given by 2^{-m}, since the one possibly larger interval U_m is included in the set $\Omega - \Omega_c$ which has μ-measure zero. The total number of quantizing intervals K_m in Q_m is then, from (6.5), just

$$K_m \leqq 2^{mN}\{\lambda(\Omega_c) + \varepsilon\} + 1. \tag{6.6}$$

Thus for any ε and any $m > m_0(\varepsilon)$, substituting $M_{\infty}(Q_m)$ for 2^{-m} and taking Nth roots in (6.6) gives

$$K_m^{1/N} M_{\infty}(Q_m) \leqq \{\lambda(\Omega_c) + \varepsilon + 2^{-mN}\}^{1/N},$$

so that

$$\lim_{m \to \infty} K_m^{1/N} M_{\infty}(Q_m) \leqq \lambda(\Omega_c)^{1/N}, \tag{6.7}$$

which proves (II) of the theorem, completing the proof of all four cases.

REFERENCES

[1] Elias, P. (1970). Bounds on the performance of optimum quantizers. *IEEE Trans. Information Theory.* **IT–16** 172–184.

[2] Gallager, R. G. (1968). *Information Theory and Reliable Communication.* Wiley, New York.

[3] Halmos, P. R. (1950). *Measure Theory.* Van Nostrand, Princeton.

[4] Hardy, G. H., Littlewood, J. E., and Polya, G. (1934). *Inequalities.* Cambridge Univ.

[5] Munroe, M. E. (1953). *An Introduction to Measure and Integration.* Addison-Wesley, Reading.

[6] Panter, P. F. and Dite, W. (1951). Quantization distortion in pulse-count modulation with nonuniform spacing of levels. *Proc. Inst. Radio Eng.* **39** 44–48.

[7] Sheppard, W. F. (1898). On the calculation of the most probable values of frequency constants for data arranged according to equidistant divisions of a scale. *Proc. London Math. Soc.* **29, 2** 353–380.

[8] Zador, P. (1964). Development and evaluation of procedures for quantizing multivariate distributions. Ph.D. Thesis, Stanford Univ.