SELECTING REPRESENTATIVE POINTS
IN NORMAL POPULATIONS

BY

S. IYENGAR and H. SOLOMON

TECHNICAL REPORT NO. 330
JANUARY 14, 1983

PREPARED UNDER CONTRACT
NO0014-76-C-0475 (NR-042-267)
FOR THE OFFICE OF NAVAL RESEARCH

Reproduction in Whole or in Part is Permitted
for any purpose of the United States Government

Approved for public release; distribution unlimited.

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
SELECTING REPRESENTATIVE POINTS
IN NORMAL POPULATIONS

BY

S. IYENGAR and H. SOLOMON

TECHNICAL REPORT NO. 330
JANUARY 14, 1983

Prepared Under Contract
N00014-76-C-0475 (NR-042-267)
For the Office of Naval Research

Herbert Solomon, Project Director

Reproduction in Whole or in Part is Permitted
for any Purpose of the United States Government

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
SELECTING REPRESENTATIVE POINTS
IN NORMAL POPULATIONS

by

S. Iyengar and H. Solomon

The representation of a continuous random variable by several
discrete points occurs often in applied probability problems. Quantization
is the term applied to this procedure and optimal quantizers
have been sought by a number of investigators. This requires defining
suitable measures for the error inherent in the procedure and then
constructing quantizing procedures that minimize the expected error.
While the problems that motivate quantization are far ranging, the mathematization leading to solutions is essentially always the same.

Most efforts are devoted to one dimensional random variables.
Obviously two, three, and higher dimensional variables can lead to more
intractability but in this paper we will explore some special cases in two
and three dimensions. The loss function we employ is that of mean square
error. Zador [1963,1982] explored the multivariate normal random variable and its quantization by a random choice of representative points. He
does not restrict himself to mean square error; rather, he defines error
as the $s^{th}$ power of the distance between the random variable and its
quantization and derives results about its asymptotic properties.

The IEEE very recently published a special issue on the topic of
quantization [1982] that collected quite recent work and work by Zador
and other investigators reported but not published as much as 25 years ago.
The papers in the issue arise out of an electrical engineering and information
theory framework and ignore the efforts of workers in other disciplines who in turn are unaware of the work of these authors.

For the one-dimensional normal random variable situation, there are papers by Bofinger [1970] who studied the question of grouping a continuous bivariate normal by selecting intervals on the marginals that would provide the maximum possible correlation between the marginal variables and by Sitgreaves [1961] who arrived at the same bivariate model as Bofinger in connection with a psychometric query on optimal test items for an achievement test. In each case, the univariate normal is quantized in an optimal manner. Maximizing the correlation is equivalent to minimizing mean square error in those cases. Previous workers also are Cox [1957], and Anderberg [1973], each of whom seeks to sectionalize or quantize the univariate normal for subsequent data analysis. Recently Fang and He [1982], motivated by clothing size category representations provide a detailed analysis for the univariate normal and give tables of representative points for \( N = 1, 2, 3, \ldots, 31 \). In an earlier paper in the electrical engineering literature, Max [1960] gives representative points for \( N = 1, 2, 3, \ldots, 26 \). When tabulated values of optimal representative points and interval endpoints are listed there is consensus among the investigators where values can be compared.

A rather early paper on quantization is by Steinhaus [1956]. In that paper, he demonstrates the two necessary (but not sufficient) conditions for optimal quantization, namely, that the optimal representative points are given by \( q_i = \text{E}(X \mid X \leq Q_i) \) when mean squared error is the loss function; and the optimal regions are nearest neighbor regions, namely

\[
Q_i = \{ x : |x-q_i| \leq \min_{1<i<j} |x-q_j| \}.
\]
Let us look into the computation of the optimal quantization of a continuous random variable, $X$. That is, we divide the real line into $N$ disjoint intervals $\{Q_i\}_1^N$, pick a representative point $q_i \in Q_i$, and define $Q(X) = q_i$ whenever $X \in Q_i$. The loss of information is indicated by, say $(X - Q(X))^2$, and we wish to minimize $E(X - Q(X))^2$ by choosing $\{Q_i\}_1^N$ and $\{q_i\}_1^N$ appropriately. We now describe and compare the methods proposed by Lloyd [1957, 1982] and Zador [1963, 1982], and our modification of Zador's method.

Lloyd notes that given the intervals $\{Q_i\}_1^N$, the optimal representative points are given by the centroids $q_i = E(X|X \in Q_i)$. This is, of course, a consequence of the fact that we use mean squared error; if, instead we used $E|X - Q(X)|$ as our criterion, the optimal $q_i$ would be given by the conditional median. He also notes that given $\{q_i\}_1^N$, the optimal intervals are just the nearest neighbor regions, $Q_i = \{x:|x - q_i| \leq \min_{1 \leq i \leq N} |x - q_j| \}$. We have already noted that Steinhaus lists these two conditions.

These two necessary conditions for an optimal quantization suggest an iterative procedure. In particular, we start with points $\{q_i^{(1)}\}$ and define the corresponding optimal intervals $\{Q_i^{(1)}\}$; then we let $q_i^{(2)} = E(X|X \in Q_i^{(1)})$, ..., We repeat this procedure until we have convergence. One important question, then, is when does the procedure converge? Lloyd presents a simple example to show that when the density of $X$ is bimodal, then the iterative procedure may converge to a local minimum and not a global one. Kieffer [1982] has the following positive result: if the density of $X$ is log-concave which is not piecewise affine on $\mathbb{R}$,
then this iterative procedure converges to the unique optimal points at
an exponential rate; that is, if \( q^{(N)} = \{q_i^{(N)}\}_{i=1}^N \), then \( \|q^{(N)} - q^*\| < \alpha^N \)
for all large \( N \) where \( q^* \) is the optimal quantization and \( 0 < \alpha < 1 \).
When \( X \sim N(0,1) \), Lloyd gives a table of optimal points for \( N = 2, 4, 8, 16 \)
and the corresponding mean square errors. Our experience has shown that
the initial points should be chosen symmetrically about zero, else the
procedure converges much more slowly. For a table of the optimal represen-
tation points and errors, see the papers by Lloyd, Max, and Fang and He.

For future reference we write out the mean squared error when
\( X \sim N(0,1) \). Let \( \phi(x) \) and \( \Phi(x) \) be the standard normal density and
distribution functions, respectively, and assume that \( q_1 < q_2 < \cdots < q_N \).
Then

\[
E(X-Q(X))^2 = 1 - 2EXQ(X) + EQ(X)^2
\]

\[
= 1 + 2q_1 \phi\left(-\frac{q_1 + q_2}{2}\right) + 2 \sum_{i=2}^{N-1} q_i \left( \phi\left(\frac{q_i + q_{i+1}}{2}\right) - \phi\left(\frac{q_i - q_{i-1}}{2}\right) \right)
\]

\[
- 2q_N \phi\left(\frac{q_{N-1} + q_N}{2}\right) + 2q_1 \phi\left(\frac{q_1 + q_2}{2}\right) + 2q_N \phi\left(\frac{q_N + q_{N-1}}{2}\right)
\]

\[
+ \sum_{i=2}^{N-1} q_i^2 \left( \phi\left(\frac{q_i + q_{i+1}}{2}\right) - \phi\left(\frac{q_i - q_{i-1}}{2}\right) \right).
\]

Zador proposes a random quantizer: the \( q_i \) are chosen randomly
according to some density \( g \). The mean squared error is then a random
variable and the problem now is to choose \( g \) so that some aspect of the
distribution of the mean squared error is optimized. Zador shows that
\( N^2 \) times the mean squared error has a limiting distribution and he computes the mean of this limit; \( N^2 \text{MSE}_\hat{g} \overset{d}{\to} Z_g \) and \( EZ_g = \mu_g \).

He then shows that by choosing \( g(x) = \frac{1}{3}(x) / \int \frac{1}{3}(t) \, dt \), \( \mu_g \) is minimized. It is clear that the optimality criterion chosen by Zador is quite distinct from Lloyd's criterion. Random quantization in one dimension is not a crucial issue. This is because optimal quantization typically involves only one-dimensional integrals whose computation is efficient. Optimal quantization in higher dimensions, though, rapidly becomes expensive, and it is here that random quantization could be valuable. However, since we know many results for the one-dimensional case, it is of interest to see how random quantization compares in that case. In the one-dimensional case, we now propose several improvements over Zador's scheme.

First, choosing the asymptotically optimal \( g(x) = (1/\sqrt{3}) \phi(x/\sqrt{3}) \) may not do well for finite \( N \). In fact, we shall show that in one case, choosing \( g_0(x) = \frac{1}{\sigma} \phi(\frac{x}{\sigma}) \) with \( \sigma = .545 \) yields substantial improvement. Second, it seems intuitively clear that the optimal points \( \{q_i\} \) ought to be located symmetrically about zero. Zador's scheme does not guarantee this symmetry, but it is fairly easy to modify it to do so.

To illustrate these modifications, we do the analytically tractable cases, \( N=2 \) and \( N=3 \).

When \( N=2 \), assuming \( q_1 < q_2 \), we have that

\[
E(X-Q(X))^2 = 1 + q_1^2 \phi(\frac{q_1 + q_2}{2}) + q_2^2 \phi(-\frac{q_1 + q_2}{2}) - 2(q_2-q_1)\phi(\frac{q_1 + q_2}{2})
\]

\[
= 1 - 2(q_2-q_1)\phi(-\frac{q_1 + q_2}{2}) + q_2^2 - (q_2-q_1)(q_2+q_1)\phi(\frac{q_1 + q_2}{2}).
\]
Under Zador's scheme, we generate \( Y_1, Y_2 \) i.i.d. \( N(0, \sigma^2) \) and set \( q_1 = Y_{(1)}, q_2 = Y_{(2)} \). Then the range \( q_2 - q_1 \) is clearly independent of the mean \( (q_1 + q_2)/2 \). Now let \( X_1, X_2 \) be i.i.d. \( N(0, 1) \). Then

\[
(i) \quad q_2 \overset{d}{=} \sigma X_{(2)} \Rightarrow \text{Eq}^2 = \sigma^2 \text{EX}_{(2)}^2 = 2\sigma^2 \int t^2 \phi(t)\phi(t) dt - \sigma^2.
\]

\[
(ii) \quad E(q_2 - q_1) = \sigma E|X_2 - X_1| = 2\sigma/\sqrt{\pi}.
\]

\[
(iii) \quad E\phi\left(\frac{q_1 + q_2}{2}\right) = E\phi\left(\frac{\sigma}{\sqrt{2}} X\right) = 1/\sqrt{\pi}(2+\sigma^2).
\]

\[
(iv) \quad E(q_1 + q_2)\phi\left(\frac{q_1 + q_2}{2}\right) = 2E \frac{\sigma}{\sqrt{2}} X \phi\left(\frac{\sigma}{\sqrt{2}} X\right) = \frac{\sigma^2}{\sqrt{\pi}(2+\sigma^2)}.
\]

Collecting terms, we see that under Zador's scheme, \( \text{MSE}(\sigma) = E(X - Q(X))^2 = 1 + \sigma^2 - \frac{2\sigma}{\pi} \sqrt{2+\sigma^2} \). The asymptotically optimal \( \sigma^2 \) is 3, for which the mean square error is 1.53. Straightforward numerical work shows, however, that \( \text{MSE}(\sigma) \) is minimized for \( \sigma = .545 \) and \( \text{MSE}(.545) = .77 \), which is almost half of \( \text{MSE}(\sqrt{3}) \).

It is also of interest to know if we get much improvement if we force the random points to be symmetric about the origin. In this case we generate \( Y \sim N(0, \sigma^2) \) and let \( q_1 = -|Y|, q_2 = |Y| \). For this situation, the mean squared error is \( E(1+Y^2 - 4|Y|/\sqrt{2\pi}) = 1+\sigma^2 - \frac{4}{\pi} \sigma \), which is minimized for \( \sigma = \frac{2}{\pi} \) and the minimum value is \( 1 - \frac{4}{\pi^2} = .59 \). A similar computation for \( N=3 \) (in which case \( q_1 = -|Y|, q_2 = 0, \) and \( q_3 = |Y| \)) shows that the mean square error assumes a minimum value of \( .41 \) for \( \sigma = 1.32 \).
We thus have the following short table:

<table>
<thead>
<tr>
<th>N</th>
<th>Optimal</th>
<th>Mean Squared Error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Zador, $\sigma^2=3$</td>
<td>Zador $\sigma_{\text{min}}^2$</td>
</tr>
<tr>
<td>2</td>
<td>.3634</td>
<td>1.53</td>
</tr>
<tr>
<td>3</td>
<td>.1902</td>
<td>.86*</td>
</tr>
</tbody>
</table>

*: Simulation

The results for $N \geq 4$ are analytically intractable. Also, the .86 entry above is a simulation result while all others are exact. The optimal results for $N=2$, $N=3$ are known, Gray and Karnin [1982].

Before we turn to higher dimensional problems, we describe some difficulties in the one dimensional normal case. One might naively expect that the symmetry of the distribution requires that the points be located symmetrically about zero. The conclusion is indeed true for the normal, but the reason is not symmetry alone. To illustrate the difficulties here, we consider the cases $N=2$ and $N=3$ for an arbitrary symmetric density $f$.

When $N=2$, we have, say

$$Q_{ab}(x) = \begin{cases} 
    a & \text{if } x \leq \frac{a+b}{2} \\
    b & \text{if } x > \frac{a+b}{2}
\end{cases}$$

and we seek $a$ and $b$ to minimize $E(X-Q_{ab}(X))^2$. Differentiating this expression with respect to $a$ and $b$ and setting the partial derivatives equal to zero, we get

7
\[ \frac{a+b}{2} \int_{-\infty}^{2} (x-a)f(x)\,dx = 0 \]

and

\[ \int_{a+b}^{\infty} (x-b)f(x)\,dx = 0. \]

If we let \( h(a,b) = \int_{-\infty}^{a+b/2} (x-a)f(x)\,dx \), then the two equations can be rewritten as \( h(a,b) = h(-b,-a) = 0 \). We can now say that if \( (a^*,b^*) \) provides an optimal quantization, then so does \( (-b^*,-a^*) \); this seems to be the only consequence of symmetry. In order to say that \( (a^*,b^*) \) lie symmetrically about zero, we must have that \( h(a,b) = 0 \) has a unique solution. One simple sufficient condition for this is that \( \log f \) be strictly concave (see Fleischer [1964], Trushkin [1982]). We conjecture that a weaker sufficient condition is that \( f \) be unimodal and strictly decreasing from the mode. Notice that one solution is always \( a^* = \mathbb{E}(X|X < 0) \) and \( b^* = \mathbb{E}(X|X > 0) \).

If we require three points, one might invoke a symmetry argument to say that one of the points be at the origin. However, the following intuitively clear example shows that this is not the case. Consider the following class of symmetric bimodal densities:

![Diagram](image-url)
If $\varepsilon$ is very small, then there is virtually no mass between $-L$ and $L$. Thus, if we put one of the points at zero, we are wasting it. If we put two points in the right mode and one in the left, we capture much more information in the random variable. Of course, in this case, we can reflect the asymmetrical quantization without changing the mean square error. We omit the details of such a counterexample. The non-optimality of the symmetric quantizer is a feature of an odd number of points.

The quantization of random vectors in $\mathbb{R}^d$ is in many ways a much more difficult problem than that of ordinary random variables. Even the "simplest" case of $X \sim N(0, I)$ presents many difficulties, as we shall see. First of all, whenever the random variable has a spherically symmetric density, any quantization can be rotated without changing the mean square error. More precisely, we have the following lemma, which is a generalization of the lemma of Gray and Karnin [1982].

**Lemma.** Suppose $X$ has density $f(x) = g(x'x)$ and $Q(X)$ is any quantizer. Then the family of quantizers $\{Q_\Gamma(X)\}_\Gamma$ where $Q_\Gamma(X) = \Gamma'Q(\Gamma X)$ and $\Gamma'\Gamma = I$ have the same mean square error.

**Proof.** Clearly $X \overset{d}{=} \Gamma X$. Thus,

$$E|X-\Gamma'Q(\Gamma X)|^2 = E|\Gamma'(\Gamma X-Q(\Gamma X))|^2 = E|\Gamma X-Q(\Gamma X)|^2 = E|X-Q(X)|^2.$$

Thus, we should not consider the quantizations $Q(X)$ and $\Gamma'Q(\Gamma X)$ as distinct.

As in the one-dimensional case, any quantization has two components, the subsets $\{Q_i\}_i$ of $\mathbb{R}^d$ and the representative points $\{q_i\}_i$ of each subset. Because we use mean square error, we again have the
following necessary conditions for the optimality of a quantizer:

(i) the representative point must be the centroid of the respective subset: \( q_i = E(X|X \in Q_i) \).

(ii) \( Q_i \) is determined by the nearest neighbor rule:

\[
Q_i = \{ x : \| x - q_i \| \leq \min_j \| x - q_j \| \}.
\]

It is clear that if \( Q(X) \) satisfies (i) and (ii), then so does \( \Gamma'Q(\Gamma X) \). Thus, \( \{ \Gamma'Q(\Gamma X) \}_{\Gamma} \) are all fixed points of Lloyd's algorithm.

This is not the real source of the difficulty, however. In general, there will be several distinct local minima. For example in the normal case, if \( d = 2 \) and \( N = 4 \), Gray and Karnin give the following configurations which are fixed points of Lloyd's algorithm.

\[
\begin{array}{c|c|c}
& \bullet & \\
\hline
\bullet & \bullet & \bullet \\
\hline
\bullet & \bullet & \bullet \\
\end{array}
\]

They conjecture that these three are the only fixed points of the algorithm.

Standardizing the error - that is, considering \( \frac{1}{d} E\|X - Q(X)\|^2 \), Gray and Karnin show that the average errors are .3634, .5588, and .4102 for I, II, and III respectively. We call configuration I the product quantizer since it is the Cartesian product of the optimal one-dimensional
quantizer with itself. Quantizer II performs rather poorly. Quantizer III has the intuitive appeal that one point is located at the origin, which is the mode of the distribution. Gray and Karnin comment that it "was a surprise to us that the distortion resulting from [code III] was so much larger than that of [code I]."

However, the intuition that says that one representative point should be at the origin because that is the location of the mode of the distribution can be very misleading. Consider, for instance, the problem of quantizing $X \sim N_2(0, I)$ with three points. Two configurations immediately come to mind:

\[
\begin{align*}
&b(-\frac{1}{2}, \frac{\sqrt{3}}{2}) \\
&b(-\frac{1}{2}, -\frac{\sqrt{3}}{2}) \\
&b(1, 0) \\
&b(-\frac{1}{2}, -\frac{\sqrt{3}}{2}) \\
&(-\mu, 0) \\
&(\mu, 0) \\
&(0, 0) \\
\end{align*}
\]

and

In the first case, the mean square error as a function of $\mu$ is

\[
\begin{align*}
\text{MSE}_1(\mu) &= E\|X-Q(X)\|^2 = E\|X\|^2 + E\|Q(X)\|^2 - 2EX'Q(X) \\
&= 2 + 2\mu^2\phi(-\frac{\mu}{2}) - 4\mu\phi(\frac{\mu}{2}) \\
&= 2 + 2\mu^2\phi(-\frac{\mu}{2}) - 4\mu\phi(\frac{\mu}{2}).
\end{align*}
\]
To find the $\mu$ that minimizes this, we set $\text{MSE}_1(\mu) = 0$ to get
$R(\frac{\mu}{2}) = \frac{1}{\mu}$, where $R(x) = \Phi(-x)/\phi(x)$ is Mills' ratio (for a more thorough discussion of Mills' ratio, see Iyengar [1982]). Straightforward numerical work shows that the minimizing $\mu$ is 1.224 and that the average noise is $\frac{1}{2} \text{MSE}_1(1.224) = .5951$.

The noise for the second configuration is

$$\text{MSE}(b) = 2 + b^2 - 2\text{EX}'Q(x)$$

$$= 2 + b^2 - 2(3) \int_0^\infty \int_{-x\sqrt{3}}^{x\sqrt{3}} xy \phi(y)\phi(x)dydx$$

$$= 2 + b^2 - 3\sqrt{3} \phi(0)b.$$ 

In this case our optimizing $b$ is easily seen to be $\frac{3\sqrt{3}}{2} \phi(0)$, so that the average mean squared error is $1-27/16\pi = .4629$. It is now clear that the second configuration is considerably better than the first one, even though the first one has a point at the origin. (Note that the two configurations do satisfy the two necessary conditions for optimality; we omit a formal proof.)

Quantization of a standard normal in three dimensions provides some new interesting twists. If we use eight points, one obvious choice is the product quantizer; the interesting result here is that the product quantizer can be improved upon. Indeed, Gray and Karnin give three different configurations that beat the product code. In one, there is a point at the origin, two points lie symmetrically on a line orthogonal to a plane formed by a pentagon whose vertices are the other five points. Another quantizer, suggested by N.J.A. Sloane to Gray and Karnin [1982] was obtained by rotating the top square of the product quantizer by $45^\circ$
to obtain a twisted cube. Gray and Karnin report simulation results to show that these quantizers are superior. We did a straightforward (but very expensive) numerical integration to get the following results listed below. Notice that the best of these three does not have a representative point at the origin.

<table>
<thead>
<tr>
<th>Quantizer</th>
<th>Simulation (Gray,Karnin)</th>
<th>Numerical Integration</th>
<th>True</th>
</tr>
</thead>
<tbody>
<tr>
<td>Product</td>
<td>-</td>
<td>.3635</td>
<td>$1 - \frac{2}{\pi} = .3634$</td>
</tr>
<tr>
<td>Pentagon-origin-poles</td>
<td>.3590</td>
<td>.3585</td>
<td>-</td>
</tr>
<tr>
<td>Twisted cube</td>
<td>.3573</td>
<td>.3581</td>
<td>-</td>
</tr>
</tbody>
</table>
References


Lloyd, S. (1957). Least Squares Quantization in PCM, draft manuscript, Mathematical Research Department, Bell Laboratories, New Jersey.


<table>
<thead>
<tr>
<th>4. TITLE (and Subtitle)</th>
<th>SELECTING REPRESENTATIVE POINTS IN NORMAL POPULATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>7. AUTHOR(s)</td>
<td>S. IYENGAR and H. SOLOMON</td>
</tr>
</tbody>
</table>
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS | DEPARTMENT OF STATISTICS  
STANFORD UNIVERSITY  
STANFORD, CALIF. 94305 |
| 11. CONTROLLING OFFICE NAME AND ADDRESS | OFFICE OF NAVAL RESEARCH  
STATISTICS & PROBABILITY PROGRAM CODE 411SP  
ARLINGTON, VA 22217 |
| 15. SECURITY CLASS. (of this report) | UNCLASSIFIED |
| 16. DISTRIBUTION STATEMENT (of this Report) | APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED. |
| 19. KEY WORDS (Continue on reverse side if necessary and identify by block number) | quantization, representation of normal distribution  
Zador's random points representation. |
| 20. ABSTRACT (Continue on reverse side if necessary and identify by block number) | Quantization of the univariate normal arises in a number of applications. One seeks an optimal set of representative points and a number of investigators have written on this problem and prepared tables. In this paper we explore some special cases in two and three dimensions employing mean square error as a loss function. Results are given for these special multivariate situations. |