OPTIMUM QUANTIZATION FOR MINIMUM DISTORTION

James R. Caprio
Systems Engineering Division
Comptek Research, Inc.
Buffalo, N.Y.

Nancy Westin
Systems Engineering Division
Comptek Research, Inc.
Buffalo, N.Y.

John Esposito
Electrical Engineering Dept.
State University of N.Y. at Buffalo
Buffalo, N.Y.

ABSTRACT

This paper treats the problem of optimal selection of data quantization levels for minimum error.

No assumptions are made regarding the underlying statistics of the process to be quantized. A finite precursor sample of the data is analyzed to infer the underlying distribution. Selection of optimum quantization levels can then be related to the generation of an optimum histogram for the data record. The optimum histogram is obtained by a dynamic programming approach for both least mean square error and minimum Chebychev error criteria.

Transmitted data can then be quantized according to levels specified by the histogram. The process can be repeated periodically either with a new data sample, if the underlying process is nonstationary, or performed on the accumulated record in the stationary case.

INTRODUCTION

An important error source in digital telemetering applications arises in the quantization of analog signals prior to transmission which are to be reconstituted as analog signals upon reception.

The encoding/decoding scheme, mutually agreed to by transmitter and receiver, is as follows. L Quantization intervals $I_1, I_2, ..., I_L$ and corresponding quantization levels $Q_1, Q_2 ... Q_i \in I_i$, are agreed to. The intervals I_i are usually contiguous and cover the total possible range of input signal voltages $x(t)$. At a particular sampling instant t_k, the input

analog signal voltage $x(t_k)$ is quantized to the level Q_i corresponding to the bin interval I_i it has fallen in. The level Q_i (or a symbol representation thereof) is transmitted and the receiver reconstructed signal will be based on the assumption that $x(t)$ had value Q_i at the sampling instant. Of course, an irreparable error is made and even with noiseless channels $x(t)$ can not be reproduced with perfect fidelity at the receiver output. Up to the point of diminishing return where quantization is no greater a contributor to output error than are transmitter, channel and receiver noise, additional levels of quantization will reduce the error; but, at the expense of increased transmitted data rate. Thus, both for fidelity and data rate considerations, the choice of quantization intervals and levels is an important one and should be made under the guidance of a minimum-distortion-with-minimum-levels policy to the extent that such a criterion can be devised. For the rare case in which the stochastic signal $x(t)$ has a known a priori probability density function (pdf) Max, reference 1, has provided an optimum solution. Here we present a method that requires no such a priori knowledge and provides additionally remarkably good estimates of tile unknown underlying pdf.

MINIMAL A POSTERIORI ERROR HISTOGRAM CRITERION

In Max's method the contiguous quantization intervals $I_i = (D_1, D_2)$, $I_2 = [D_2, D_3)$..., $I_L = [D_L, D_L+1)$ $(D_I = -\infty, D_L+1 = \infty)$ and associated quantization levels Q_i are selected to minimize a distortion D which is defined as the a priori expected value of a differentiable . on f of the quantization error;

$$D = E[f(x_{in} - x_{out})] = \sum_{i=1}^{L} \int_{D_i}^{D_{i+1}} f(x-Q_i)p(x)dx$$

By differentiating D with respect the levels Q_i and division points D_i, Max obtains 2L-1 simultaneous equations for the unknown parameters. For the particular case he treats in depth, D is the mean square error $(f(x)=x^2)$ and the system of equations becomes

$$Q_j = 2D_j - Q_{j-1} \qquad j = 2,\ldots,L$$

(1)

Equation (2) shows that Q_j is the centroid of the pdf $p(x)$ between D_j and D_{j+1}.

Besides the obvious deficiency of requiring a known pdf, solution of these equations requires a single parameter search. For example, assume a value for Q_1, solve (2) for D_2, solve (1) for Q_2 and repeat the procedure. If the resulting Q_L is indeed the centroid of $p(x)$ betwee D_L and $D_{L+1} = \infty$ the original Q_1 was the correct choice and the computed Q_i, D_i are optimal. Otherwise a new Q_1 must be tried and the procedure repeated until convergence.

To circumvent these problems we propose a quantization criterion based on an a posteriori measure of the errors $x_i - Q_j$ where Q_j is the quantization level to which the signal sample x_i is encoded and x_i is one of a precursor set of N+1 independent samples of $x(t)$ that we analyze prior to initiating transmission. Thus, based on analysis of the samples $[x_0, x_1, ..., x_N]$, quantization levels and intervals will be selected and relayed to the receiver for use in signal reconstruction. If N is large enough and $x(t)$ is a stationary process these intervals and levels will remain in force through all subsequent transmissions. Otherwise, at prearranged intervals, new levels and intervals based on more recent sample analysis will be transmitted. With this approach the problem of optimum quantization can be formulated in terms of the optimum histogram generation problem depicted in Figure 1. That is, given data sequence $x_0, x_1, ..., x_N$, generate an L-bin histogram with the property that the total error measure ϵ_L of distances $\epsilon_i = x_i - Q_j$ of all samples x_i from their assigned bin centers Q_j is minimized. This problem is solved by dynamic programming methods in the following section for each of two popular error measures;

(i) Least Mean Square (LMS) error

$$\mathcal{E}_L(N) \ = \ \min \ \sum \epsilon_i^2 \tag{3}$$

(ii) Chebychev error

$$\mathcal{E}_L(N) \ = \ \min \left\{ \max_i \ \left| \epsilon_i \right| \right\} \tag{4}$$

In the latter case the resulting quanitization levels (bin centers) Q_L are truly geometrically centered within the quantization interval, while, in the former, LMS quantization levels will be shown to be centroids of the subset of data falling in the optimally selected interval. This is a notable similarity with Max's result (eqn(2)).

SOLUTION BY DYNAMIC PROGRAMMING

We first consider what appears to be an unrelated approximation problem which was solved by Fryer[2] and Bellman[3]. As shown in Figure 2, we are given a set of N+1 pairs $(x_0, y_0), ..., (x_N, y_N)$ and wish to fit L lines to the data such that the overall error (LMS or Chebychev) is minimized. The data are not necessarily equally spaced and the linear segment intervals $[x_0, x_{B_1}], [x_{B_1+1}, x_{B_2}], ..., [x_{B_{L-1}}, x_N]$ may be of unequal length. Clearly, this problem reduces to one of selection of optimum breakpoint indices $B_1, B_2, ... B_{L-1}$. Once two consecutive breakpoints are specified the optimum linear approximation Ax+B within the specified segment can be obtained easily for the LMS case from well known least squares approximation formulas. In the Chebychev case the optimum first-order-linear segment approximation has the property that the minimum error must occur three times, in alternating fashion, above and below the approximation curve.

Scheid [4] has used this property to develop a point exchange algorithm that will yield the optimum solution. The algorithm applies as well to higher order polynomials with the required number of alternations increasing accordingly.

Whatever the error measure let us designate the resulting minimum error in fitting a linear segment from x_i to x_j by H(i,j). For a specified set of breakpoints $B_1, B_2, \ldots B_{L-1}$ the total error is then

$$E_L(N) = H(0, B_1) + \ldots + H(B_{L-1}, N) \tag{5}$$

in the LMS case and for the Chebychev case

$$E_L(N) = \max_i \ H(B_i, B_{i+1}) \tag{6}$$

where $B_0 = 0$, $B_L = N$ and, for either case,

$$\mathcal{E}_L(N) = \min_{B_i} \ E_L(N). \tag{7}$$

Fryer [2] has shown the futility of attempting the brute force solution of investigating all possible breakpoint arrangement to determine that yielding minimum overall error. The number of such arrangements satisifying

$$0 < B_1 < B_2 \ldots < B_{L-1} < N$$

is easily shown to be

$$\binom{N-1}{L-1} = \frac{(N-1)!}{(L-1)!(N-L)!}$$

Thus, for a typical case of 100 points and 10 lines as in an example shown later, the number of possible arrangements exceeds 1.5×10^{14}. Even in the unlikely event that computation speeds will one day permit each of the 10 required H(i,j) errors to be computed in a microsecond, the total computation time would approach 5 centuries.

Fortunately this problem is well suited for application of Bellman's Principle of Optimality. We suppose that the first breakpoint from the left has been chosen perhaps non-optimally at index k. The resulting error for the right-most segment is then H(k+1,N). The Principle of Optimality suggests that we proceed optimally thereafter, i.e. assign the remaining breakpoints to minimize error over the left samples $x_0, \ldots, x_k$. But this minimum left error is $\epsilon_{L-1}(k)$. Thus the total error incurred is

$$E_L(N) = \mathcal{E}_{L-1}(k) + H(k+1, N)$$

for the LMS case, and

$$E_L(N) = \max \left\{ \varepsilon_{L-1}(k), H(k+1,N) \right\}$$

for the Chebychev case. It follows that the overall minimum error satisfies the recursion

$$\varepsilon_L(N) = \min_k \left\{ \varepsilon_{L-1}(k) + H(k+1,N) \right\} \tag{8}$$

for LMS approximation, and for Chebychev approximation

$$\varepsilon_L(N) = \min_k \max \left\{ \varepsilon_{L-1}(k), H(k+1,N) \right\} \tag{9}$$

where in both cases the initial condition is

$$\varepsilon_1(k) = H(0,k) \qquad k = 0,\ldots,N \tag{10}$$

Thus, starting with the ϵ_1 array from eqn.(10) the recursions (8), (9) proceed through each successive value of the number of lines L until we reach the desired maximum number. At each stage the minimization over index k reveals the optimum breakpoints and the H(i,j) computation provides the corresponding optimum linear approximation parameters (slope A and intercept B). Thus, not only is the best L line fit to the samples $x_0,\ldots,x_N$ obtained but, the best J line fit to $x_0,\ldots,x_k$ for J = 1,2,...,L; k=0,1...,N as well. Moreoever, in cases where the optimum solution is non-unique, the foregoing procedure will reveal all possible optimum solutions (breakpoint arrangements).

This is a very elegant solution of an interesting approximation problem; but, what does it have to do with the optimum quantiziation/histogram-generation problem posed earlier? The answer, as portrayed in Figure 3, is that we have solved the histogram problem if we simply sort the input data $x_0,\ldots,x_N$ in increasing order and use the foregoing technique to fit zero-order line segments in an optimal fashion to the resulting monotone sequence. The quantization levels (bin centers) are the resulting zero-order line intercepts, the bin frequencies are given by the number of x_i samples in the segments and the bin widths are determined by the corresponding segment approximation errors. Note that the resulting bins are not contiguous; the method for joining them contiguously is described below. We further observe that owing to the reduction to zero-order approximation, the computations required for the optimum quantum levels and associated errors are quite simple. For the LMS case the segment level Q_i and error H are simply the sample mean and variance of the data in the segment. For the Chebychev case, since the data are sorted, $Q_i = (x_{Ri} + x_{Li})/2$ and $H = (x_{Ri} - x_{Li})/2$ where x_{Ri}, x_{Li} are the extreme right and left data points in the segment.

SELECTION OF CONTIGUOUS-BIN EDGES AND PDF ESTIMATION

Unless it is known a priori that the input signal can not fall in certain ranges the noncontiguous bins of Figure 1B should not be permitted. The gaps of Figure 1B and also arising in the technique illustrated in Figure 2 merely reflect the fact that all possible values of x can not possibly occur in a finite sample unless the process is discrete. To join the j, j+1 bins contiguously at some division point D_j, we simply apply Max's result (1)

$$D_j = \frac{Q_j + Q_{j+1}}{2} \qquad\qquad j=1,\ldots,L-1 \tag{11}$$

provided that it does not cause overlapping bins. To avoid this D_j is set equal to the value closest to (11) that satisifies $x'B_j \le D_j \le x'_{1+Bj}$ where x_i denotes the sorted sequence and B_j are the breakpoints. The specification $D_0 = -\infty$, $D_L = +\infty$ completes the solution of the quantization problem.

The pdf estimates for interior intervals are

$$\hat{p}(Q_j) = \frac{\text{number of points in segment j}}{(D_j - D_{j-1})(N+1)} \qquad j = 2,\ldots,L-1 \tag{12}$$

This estimate is most accurately attributed to the bin center, but, also provides a reasonable estimate for p(x) at all points x within the segment. For the extreme intervals the assignment $D_0 = -\infty$, $D_L - \infty$ will cause the pdf estimate to vanish. We can avoid this if absolute maximum values of x(t) are known a priori. Other possibilities include extending these bins beyond their most exterior sample by some multiple of the segment deviation.

RESULTS AND CONCLUSIONS

The following example, illustrate the technique. The x_k samples are integerized, computer-generated random variates given by $x_k = A \sin \theta_k$ where θ_k is uniform on $[0, 2\pi]$. Table 1 lists the sorted sample values and Table 2 provides the dynamic programming 10-segment approximation parameters. The estimation results are graphed in Figure 4. Note that we have assumed that the maximum signal amplitude $|A| = 1000$ is known in order to obtain density estimates for the first and last interval. The resulting 10 bin Chebychev and LMS pdf estimates are observed to approximate the 'true' underlying pdf

$$p(x) = \frac{1}{\pi \sqrt{A^2 - x^2}} \tag{12}$$

quite reasonably. An apparent advantage of the LMS estimates over the Chebychev case is observed at the extreme edges where p(x) is sharply peaked. Since the Chebychev error is symmetric, the estimates $\hat{p}(Q_1)$, $\hat{p}(Q_L)$ are attributed to the bin geometric centers while the corresponding LMS bin centers are data centroids that fall much closer to the data extremes $\pm A$ and thus can more accurately portray the singularity nature of the underlying pdf.

```
-999   -948   -760   -374   -182   130   300   636   798   968
-993   -945   -700   -366   -130   130   317   642   798   970
-993   -933   -669   -309   -113   156   342   662   809   978
-992   -909   -649   -300    -61   156   366   675   878   979
-992   -898   -615   -292     -8   165   398   688   882   986
-991   -882   -573   -284      0   182   414   731   909   992
-981   -870   -551   -275     52   199   438   737   936   993
-968   -848   -515   -241     78   207   461   748   936   993
-958   -843   -422   -224    104   267   580   777   958   997
-948   -766   -382   -216    121   292   608   788   965   999
```

Table 1. Sorted x_k Sample Values

k	Chebychev				LMS			
	B_k	x_{B_k}	Q_k	D_k	B_k	x_{B_k}	Q_k	D_k
0				-1000.0				-1000.0
1	18	-843	-921.0	- 805.8	13	-909	-967.9	- 903.0
2	24	-615	-609.5	- 594.0	20	-760	-838.1	- 724.2
3	28	-422	-497.5	- 398.3	27	-669	-610.3	- 458.7
4	39	-216	-299.0	- 195.0	39	-216	-307.1	- 194.7
5	45	0	- 91.0	19.3	45	0	- 82.3	28.9
6	57	207	129.5	235.0	57	207	140.0	249.8
7	65	414	340.5	431.8	67	461	359.5	500.5
8	69	608	523.0	622.8	74	688	641.6	707.4
9	82	809	722.5	830.5	82	809	773.3	866.6
10	99	999	938.5	1000.0	99	999	959.9	1000.0

**Table 2. Dynamic Programming Breakpoints, Quantization Levels and Histogram Bin Division Points**

Additional research exploring the accuracy of our method is underway. We hope to report shortly on comprehensive evaluations of the quantization method described as compared to Max's ideal results and on the pdf estimation as compared to the popular Kernel method [5], [6] and Nearest Neighbor method [7], [8].

REFERENCES

[1]   J. Max, "Quantizing for Minimum Distortion", IRE Transactions on Information Theory, vol. IT-6, pp. 7-12, March 1960.

[2]   W. Fryer, Cornell Aeronautical Laboratory internal report (unpublished) 1963.

[3]   R. Bellman, "Curve Fitting by Segmented Straight Lines", American Statistical Association Journal, pp. 1079-1084, Sept. 1969.

[4]   F. Scheid, "The Under-Over-Under Theorem", Mathematical Association of America Monthly, pp. 862-871, Nov. 1961.

[5]   M. Rosenblatt, "Remarks on Some Nonparametric Estimates of the Density
        Function", Annals of Mathematical Statistics, vol. 27, pp. 832-837, 1957.

[6]   E. Parzen, "On the Estimation of a Probability Density Function and the Mode",
        Annals of Mathematical Statistics, vol. 33, pp. 1065-1076, 1962.

[7]   D. O. Loftsgaarden and C. P. Quesenberry, "A Nonparametric Estimate of a
        Multivariate Density Function", Annals of Mathematical Statistics, vol. 36, pp.
        1049-1051, 1965.

[8]   L. P. Devroye, "The Uniform Convergence of Nearest Neighbor Regression
        Function Estimators and Their Application in Optimization", IEEE Transactions on
        Information Theory, vol. IT-24, No. 2, pp. 142-151, March 1978.

Quantization problem: Given independent samples $x_i = x(t_i)$, i=0,...,N
generate an L bin histogram which minimizes distances of points from
their respective bin 'centers'.
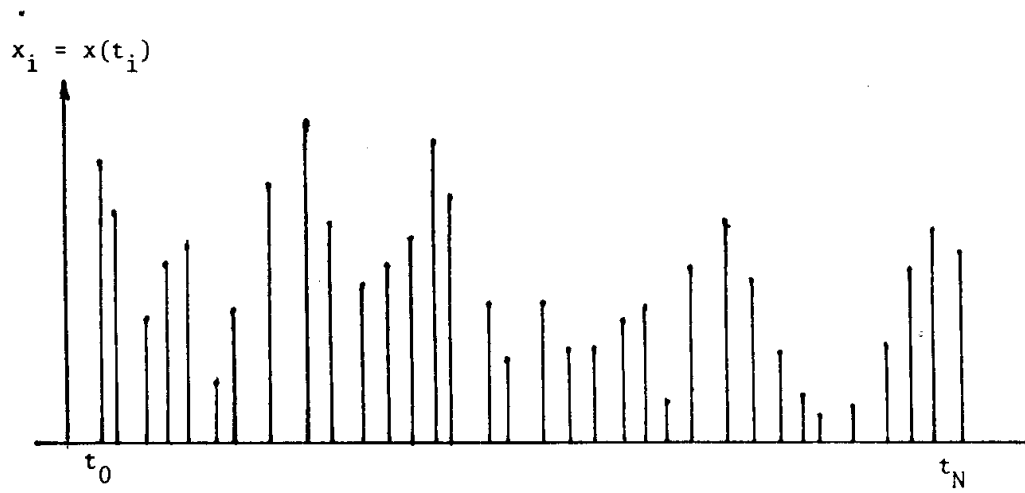
A) Data sequence
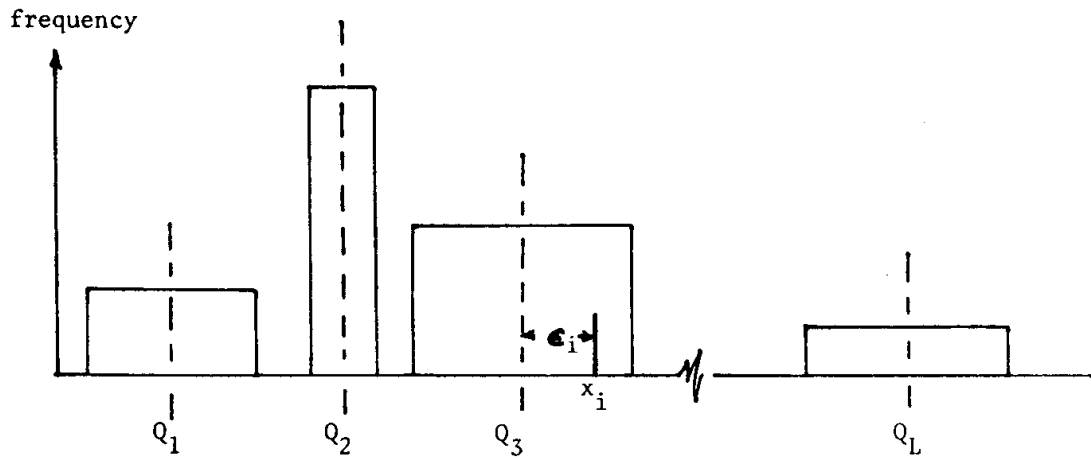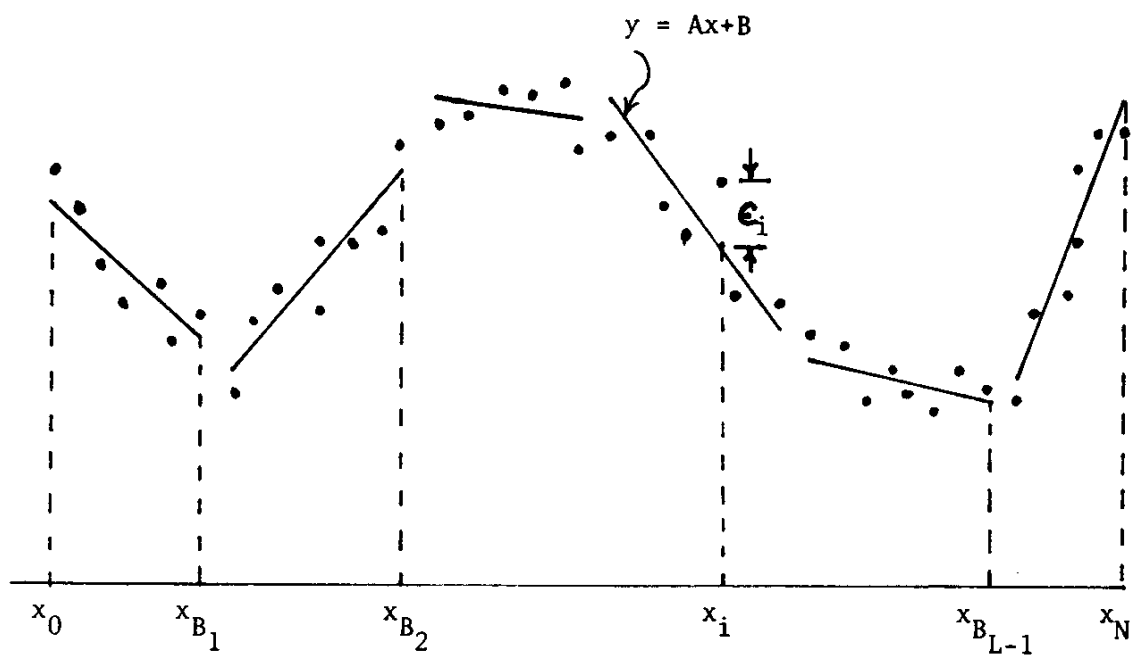

B) Optimum histogram

LMS distance criterion:

$$\mathcal{E}_L(N) = \min \sum \epsilon_i^2$$

Chebychev distance criterion:

$$\mathcal{E}_L(N) = \min \left\{ \max \left| \epsilon_i \right| \right\}$$

Figure 1. Optimum Quantization in Terms of Histogram Generation

Approximation problem: Fit L lines to N+1 points with minimum total
error $\mathcal{E}_L(N)$.

$$\epsilon_i = y_i - (Ax_i + B)$$

LMS error criterion:
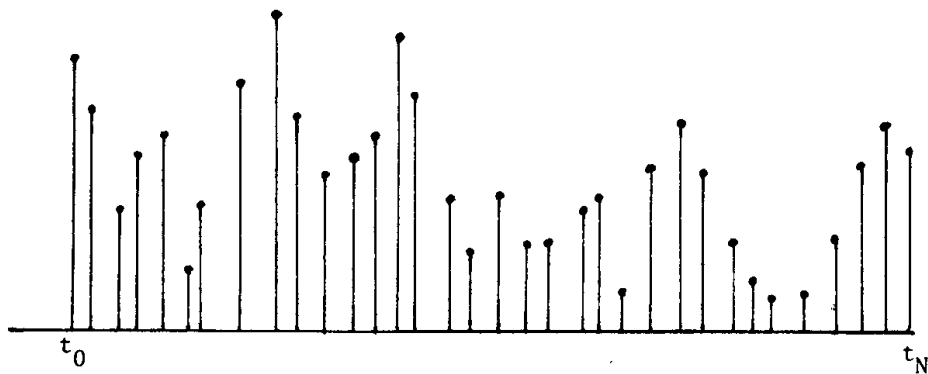
$$\mathcal{E}_L(N) = \min \sum \epsilon_i^2$$
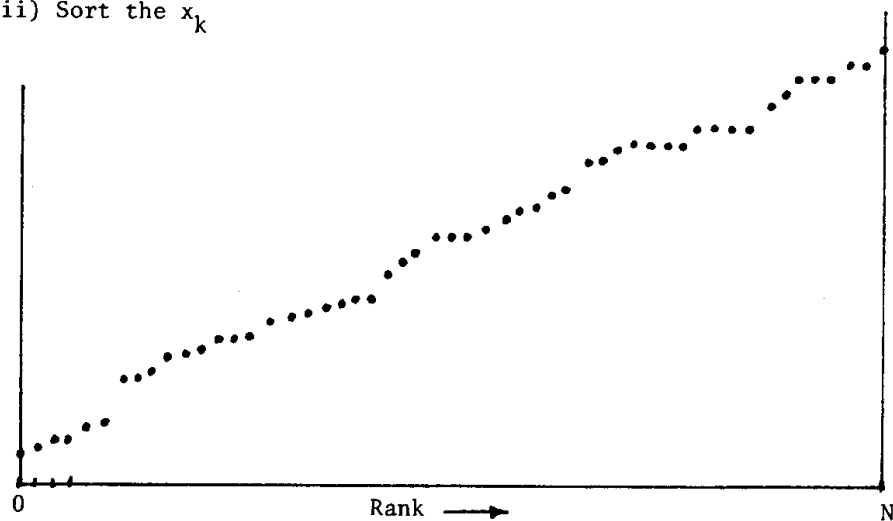
Chebychev error criterion:

$$\mathcal{E}_L(N) = \min \left\{ \max \left| \epsilon_i \right| \right\}$$

Figure 2. Piece-Wise Linear Approximation Problem

(i) Gather independent random data sequence $x_k = x(t_k)$

t_0 t_N

(ii) Sort the x_k

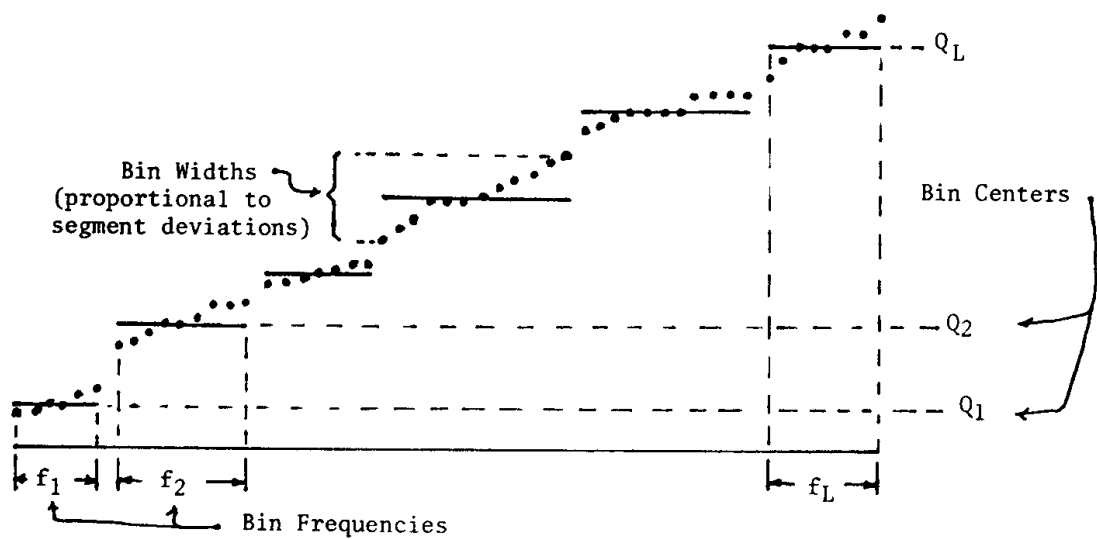0   Rank ⟶   N

(iii) Use dynamic programming to optimally fit <u>zero-order</u> lines

Bin Widths
(proportional to
segment deviations)

Bin Centers

Q_L

Q_2

Q_1

f_1 f_2 f_L
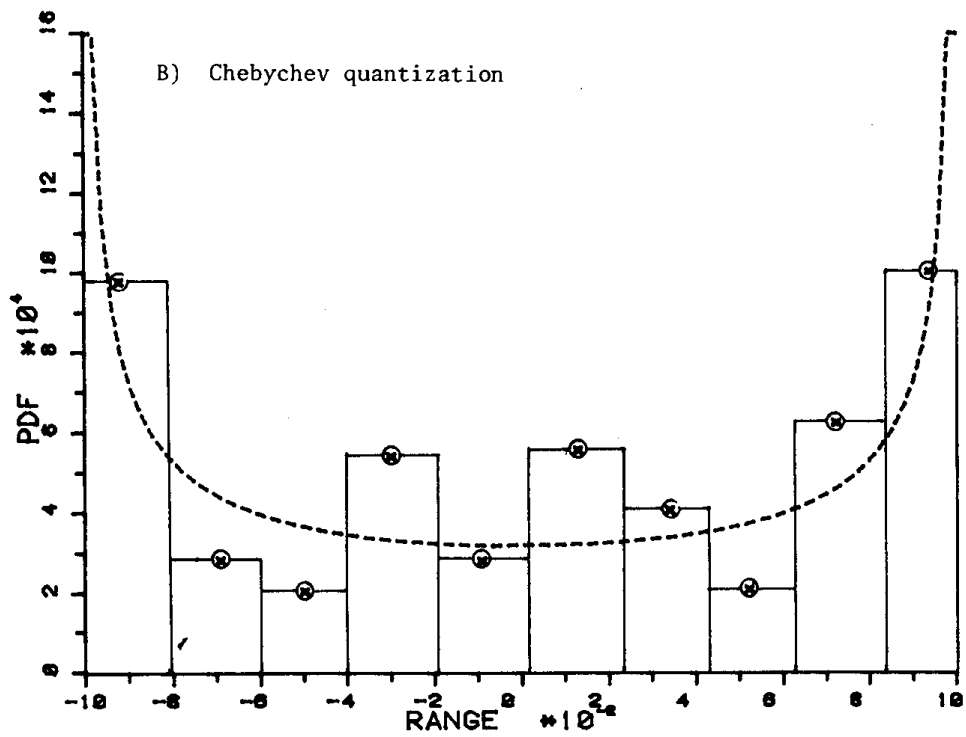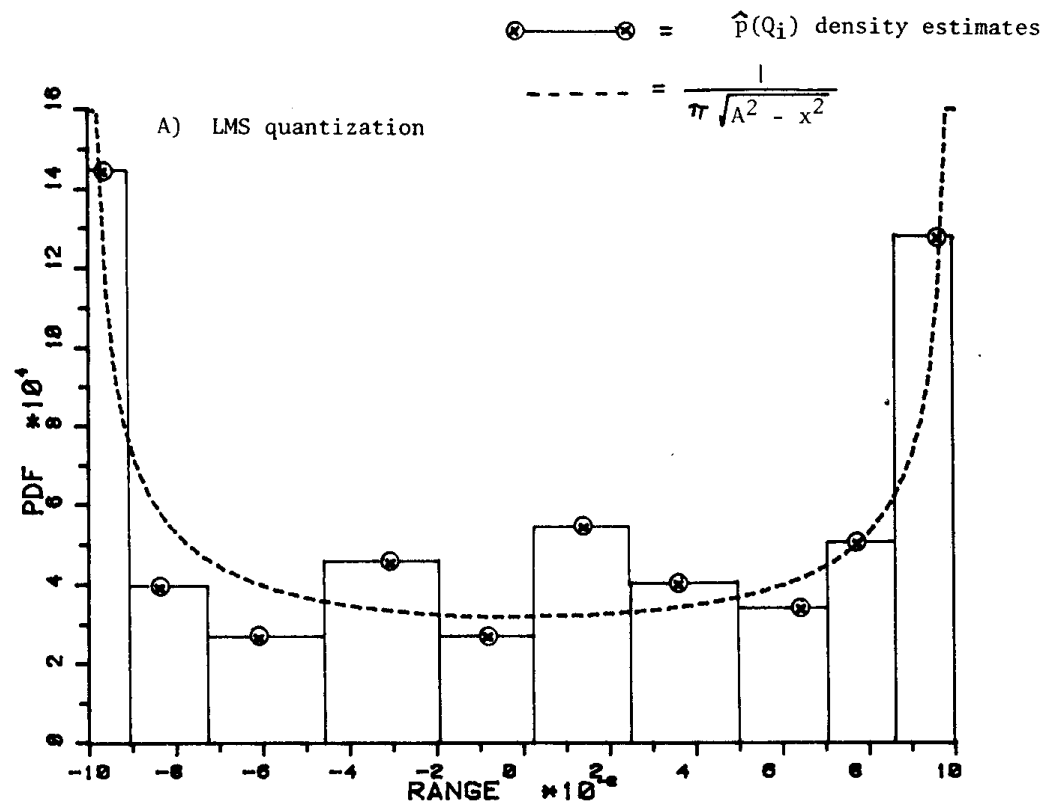
Bin Frequencies

Figure 3. Optimum Histogram Generation Method

Figure 4. pdf Estimates and Quantization Intervals