Computer Sciences Department
The University of Wisconsin
1210 West Dayton Street
Madison, Wisconsin 53706

ESTIMATING THE ACCURACY OF QUASI MONTE CARLO INTEGRATION*

by

John H. Halton

Technical Report #139

November 1971

*The work described in this paper was done with the support of the National Science Foundation, under Grants GJ-171 and GJ-29986X, for which the author wishes to express his gratitude. It was presented at a Symposium at the Centre de Recherches Mathématiques, University of Montreal, Canada, 9-14 September 1971.

ESTIMATING THE ACCURACY OF QUASI MONTE CARLO INTEGRATION*

by

John H. Halton
The University of Wisconsin
Madison, Wisconsin, U. S. A.

ABSTRACT

The author discusses the doubtful value of error-bounds and estimates of a statistical nature, based on variance-estimators and the Central Limit Theorem, when used in situations where quasi-random (deterministic) sets of points are used to estimate integrals over multi-dimensional intervals. (The doubt extends, of course, to all quasi-random calculations.) He describes an alternative approach, based on discrepancies of point-sets and consequent bounds on the error, for integrands of bounded variation in the sense of Hardy and Krause. Suitable error-estimates, computable during the calculation of the main estimator, are described.

1. Monte Carlo Integration

The problem we shall consider, is that of evaluating an integral of the form

$$\int_{U^k} d\omega^k f(\underset{\sim}{z}) = \int_0^1 dz_1 \int_0^1 dz_2 \cdots \int_0^1 dz_k\ f(z_1, z_2, \ldots, z_k) = \theta; \qquad (1)$$

where U denotes the unit interval $[0,1]$, U^k is the k-dimensional unit hypercube whose points $\underset{\sim}{z} = [z_i]_{i=1}^k$ satisfy the conditions $0 \leq z_i \leq 1$ $(i = 1, 2, \ldots, k)$, ω is the Lebesgue measure on the real line, and ω^k is the k-dimensional Lebesgue product-measure; so that $d\omega^k = dz_1 dz_2 \cdots dz_k$ and $\omega^k(U^k) = 1$. We assume that f is Riemann-integrable in U^k and that

$$|\theta| < \infty. \qquad (2)$$

We shall consider a situation in which either k is very large or $f(\underset{\sim}{z})$ is extremely complicated to evaluate, or both; so that it is not practical to evaluate θ by any traditional quadrature formula.

The basic Monte Carlo method for estimating θ is to generate a <u>canonical</u> <u>random</u> <u>sequence</u>: a sequence $\Xi = [\underset{\sim}{\xi}_n]_{n=0}^{\infty}$ of points $\underset{\sim}{\xi}_n = [\underset{\sim}{\xi}_{ni}]_{i=1}^k$ independently distributed uniformly in U^k, and to compute the estimator

$$t_N = t_N(\Xi) = N^{-1} \Sigma_{n=0}^{N-1} f(\underset{\sim}{\xi}_n). \qquad (3)$$

Then clearly t_N will have mean value

$$\underset{\sim}{E}[t_N] = \underset{\sim}{E}[f] = \theta, \tag{4}$$

and variance

$$\sigma_N^2 = \text{var}\,[t_N] = \underset{\sim}{E}[|t_N - \underset{\sim}{E}[t_N]|^2] = \underset{\sim}{E}[|t_N|^2] - |\theta|^2$$
$$= N^{-1}(\underset{\sim}{E}[|f|^2] - |\theta|^2) = N^{-1}\,\text{var}[f] = N^{-1}\,\sigma^2. \tag{5}$$

Even if σ_N^2 is infinite, it is well-known that t_N converges to θ as $N \to \infty$, both <u>in probability</u> (i.e., for any $\varepsilon > 0$, the probability that $|t_N - \theta| \geq \varepsilon$ tends to zero as $N \to \infty$) and <u>almost surely</u> (i.e. the set of sequences Ξ for which $t_N(\Xi)$ does not tend to θ as $N \to \infty$ has probability zero.) (see, e.g., [1], pp. 230 and 245). However, we shall assume here that, as is often the case,

$$\underset{\sim}{E}[|f|^2] = \int_{U^k} d\omega^k\, |f(\underset{\sim}{z})|^2 = \sigma^2 + |\theta|^2 < \infty; \tag{6}$$

so that

$$\sigma_N^2 = \sigma^2 N^{-1} \to 0 \quad \text{as} \quad N \to \infty. \tag{7}$$

The Central Limit Theorem (e.g. [1], p. 293) then implies that the distribution of t_N is asymptotic to a normal distribution with mean θ and variance σ_N^2 as $N \to \infty$; so that the probability, that $t_N < \theta + w\sigma_N$, converges to

$$\mathfrak{N}(w) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{w} e^{-z^2/2}\, dz, \tag{8}$$

as $N \to \infty$; whence the error made in estimating the integral θ with the sum t_N,

$$\delta_N(f) = |t_N - \theta|, \tag{9}$$

is greater than $w\sigma_N$ with probability asymptotically equal to

$$\int_{-\infty}^{-w} e^{-z^2/2}\, dz + \int_{w}^{\infty} e^{-z^2/2}\, dz = 2[1 - \mathfrak{N}(w)]. \tag{10}$$

In particular, we may note that the error $\delta_N(f)$ is greater than σ_N, $2\sigma_N$, or $3\sigma_N$ with probability asymptotically equal to about 0.317, 0.0455, and 0.00270, respectively.

Thus, the <u>standard deviation</u> σ_N is an important measure of the error in a statistical estimate t_N of θ; and consequently, it is useful to have an estimate of σ_N itself. The quantity

$$\begin{aligned} s_N^2 = s_N(\Xi)^2 &= N^{-2} \; \Sigma_{n=0}^{N-1} |f(\underset{\sim}{\xi}_n) - t_N|^2 \\ &= N^{-4} \; \Sigma_{n=0}^{N-1} |(N-1)f(\underset{\sim}{\xi}_n) - \Sigma_{m \neq n} f(\underset{\sim}{\xi}_m)|^2 \\ &= N^{-2} \; \Sigma_{n=0}^{N-1} |f(\underset{\sim}{\xi}_n)|^2 - N^{-3} \; |\Sigma_{n=0}^{N-1} f(\underset{\sim}{\xi}_n)|^2 \end{aligned} \tag{11}$$

is called the <u>sample mean variance</u> and is clearly a simple Monte Carlo estimator for σ_N^2. Its mean value is seen to be†

$$\begin{aligned} \underset{\sim}{E}[s_N^2] &= \underset{\sim}{E}[N^{-3}(N-1) \; \Sigma_{n=0}^{N-1} \; |f(\underset{\sim}{\xi}_n)|^2 - N^{-3} \; \Sigma_n \Sigma_{m<n} \{f(\underset{\sim}{\xi}_m)f(\underset{\sim}{\xi}_n)^* \\ &\quad + f(\underset{\sim}{\xi}_m)^* f(\underset{\sim}{\xi}_n)\}] = N^{-2}(N-1)\underset{\sim}{E}[|f|^2] - N^{-2}(N-1)\theta\theta^*, \end{aligned}$$

---

† Here, and elsewhere in this paper, the * denotes the complex-conjugate quantity.

whence

$$\underset{\sim}{E}[s_N^2] = \frac{N-1}{N^2} \sigma^2 = \frac{N-1}{N} \sigma_N^2 . \tag{12}$$

For this reason, the unbiased estimator

$$s_N'^2 = N^{-1}(N-1)^{-1} \Sigma_{n=0}^{N-1} |f(\underset{\sim}{\xi}_n) - t_N|^2 = \frac{N}{N-1} s_N^2 \tag{13}$$

is often preferred. By the laws of large numbers mentioned above, it is clear that both Ns_N^2 and $Ns_N'^2$ converge in probability and almost surely to σ^2, as $N \to \infty$.

In actual practice, Monte Carlo calculations are most often carried out with sequences of numbers, called pseudo-random and quasi-random (the former, if they have passed a series of statistical tests; the latter if they are merely found to be efficient in performing certain classes of Monte Carlo calculations as if they were random sample-sequences), which are generated by deterministic algorithms and have no true randomness at all. The question arises, of how we are to estimate the accuracy of the answers which we obtain by these means. This question has recently led to much discussion in the literature (see, e.g. [9], and [4], pp. 6, 32, 35-47, 50-51, where many additional references to different, and sometimes conflicting, approaches can be found).

In many cases, the user of pseudo-random and quasi-random numbers assumes that the estimators s_N^2 or $s_N'^2$, just like the estimators t_N, still give the right answers, at least in some qualitative sense. However, the remark above, that s_N^2 is an estimator of σ_N^2, gives us the first heuristic hint, that the assumption that s_N^2 gauges the error $\delta_N(f)$ is some way, is unwarranted in general.

2. Discrepancy

Let us now write Ξ for a deterministic sequence of points in U^k; and let us write $\nu_N(\underset{\sim}{z})$ for the number of points among $\underset{\sim}{\xi}_0$, $\underset{\sim}{\xi}_1, \ldots, \underset{\sim}{\xi}_{N-1}$ which satisfy the conditions

$$0 \leq \xi_{ni} < z_i \leq 1 \quad (i = 1, 2, \ldots, k); \tag{14}$$

and put

$$\Omega(\underset{\sim}{z}) = \omega^k(\underset{\sim}{z}) = \Pi_{i=1}^k z_i \tag{15}$$

for the measure of the region defined by (14). Then the <u>local discrepancy</u> of the sequence Ξ at $\underset{\sim}{z}$ is

$$\Delta_N(\underset{\sim}{z}) = N^{-1} \nu_N(\underset{\sim}{z}) - \Omega(\underset{\sim}{z}), \tag{16}$$

and we say that Ξ is <u>equidistributed</u> in U^k if $\Delta_N(\underset{\sim}{z}) \to 0$ as $N \to \infty$, for every $\underset{\sim}{z}$ in U^k. From this we may derive various measures of the imperfection of equidistribution of the sequence Ξ in U^k:

$$\mathcal{M}_N = -\inf_{\underset{\sim}{z} \in U^k} \Delta_N(\underset{\sim}{z}), \tag{17}$$

$$\mathcal{N}_N = \sup_{\underset{\sim}{z} \in U^k} \Delta_N(\underset{\sim}{z}), \tag{18}$$

$$\mathscr{D}_N = \sup_{\underset{\sim}{z}\in U^k} |\Delta_N(\underset{\sim}{z})| = \max(\mathscr{M}_N, \mathscr{N}_N), \tag{19}$$

and

$$\mathscr{T}_N = [\int_{U^k} d\omega^k \, |\Delta_N(\underset{\sim}{z})|^2]^{1/2} \tag{20}$$

are respectively termed the minimum, maximum, extreme (or L^∞), and (root) mean-square (or L^2) discrepancy [$\mathscr{D}_N$ was formerly called just the 'discrepancy', before the L^2-discrepancy (also termed the turpitude, whence the notation) was introduced by Zaremba and Halton [2,9,10] on the basis of earlier ideas of Roth [8], Hammersley [5], and Halton [3].] We note that, for a true canonical random sequence, $\nu_N(\underset{\sim}{z})$ counts the number of successes in a series of N Bernoulli trials with probability of success $\Omega(\underset{\sim}{z})$; so that the mean value of $\Delta_N(\underset{\sim}{z})$ is zero and the mean value of $|\Delta_N(\underset{\sim}{z})|^2$ is $\Omega(\underset{\sim}{z})[1-\Omega(\underset{\sim}{z})]/N$. Thus

$$\underset{\sim}{E}[\mathscr{T}_N^2] = \int_0^1 dz_1 \int_0^1 dz_2 \cdots \int_0^1 dz_k \left| \frac{z_1 z_2 \cdots z_k - z_1^2 z_2^2 \cdots z_k^2}{N} \right| = \frac{\{(\frac{1}{2})^k - (\frac{1}{3})^k\}}{N}. \,^\dagger \tag{21}$$

We note for future use that the following identity holds in the Lebesgue-Stieltjes senses:

$$N^{-1} \Sigma_{n=0}^{N-1} \, g(\underset{\sim}{\xi}_n) = \int_{U^k} d\Delta_N(\underset{\sim}{z}) \, g(\underset{\sim}{z}) + \int_{U^k} d\omega^k \, g(\underset{\sim}{z}). \tag{22}$$

---

† It is interesting to observe that, rather counter-intuitively, this expression decreases as $k \to \infty$.

We shall also require the following definitions. The <u>k-dimensional total variation</u> of a function g in U^k <u>in the sense of Vitali</u> can formally be written as

$$V^k(g) = \int_{U^k} |d^k_{z_1 z_2 \cdots z_k} g(\underset{\sim}{z})|; \tag{23}$$

and this is rigorously defined by considering k arbitrary partitions π_i of U at points $0 = x_{i0} < x_{i1} < \cdots < x_{in_i} = 1$ $(i = 1,2,\ldots,k)$, defining

$$\delta_{z_i} \phi(\underset{\sim}{z}) = \phi(z_1,\ldots,x_{ij_i},\ldots,z_k) - \phi(z_1,\ldots,x_{i(j_i-1)},\ldots,z_k), \tag{24}$$

and

$$\delta^k_{z_1 z_2 \cdots z_k} \phi(\underset{\sim}{z}) = \delta_{z_1} \delta_{z_2} \cdots \delta_{z_k} \phi(\underset{\sim}{z}), \tag{25}$$

where the x_{ij} occur in place of z_i, and the indices j_i are omitted for brevity; and then putting

$$V^k(g) = \sup_{\pi_1,\pi_2,\ldots,\pi_k} \sum_{j_1=1}^{n_1} \sum_{j_2=1}^{n_2} \cdots \sum_{j_k=1}^{n_k} |\delta^k_{z_1 z_2 \cdots z_k} g(\underset{\sim}{z})|. \tag{26}$$

The function g is <u>of bounded variation in the sense of Vitali</u> if $V^k(g)$ is finite. The function f is said to be <u>of bounded variation in the sense of Hardy and Krause</u> (BVHK) if $V^k(f)$ is finite and also $V^h(g)$ is finite whenever $h < k$ and g is a function of h variables obtained by fixing any $(k - h)$ of the z_i to be 1. Let us now write $f_{\mathfrak{S}_h}$ for such a g, where $\mathfrak{S}_h$ is any choice of h of the z_i (there are $\binom{k}{h}$ such choices), and similarly let $\mathscr{D}_{N,\mathfrak{S}_h}$ and $\mathscr{T}_{N,\mathfrak{S}_h}$ denote the L^∞ and L^2 discrepancies of the sequence $\Xi_{\mathfrak{S}_h} = [\underset{\sim}{\xi}_{n,\mathfrak{S}_h}]_{n=0}^{\infty}$, where $\underset{\sim}{\xi}_{n,\mathfrak{S}_h}$ is the h-dimensional vector consisting of the h components of $\underset{\sim}{\xi}_n$ selected in $\mathfrak{S}_h$. Then it was proved, by Koksma [7], for one dimension $(k = 1)$, and by Hlawka [6] and Zaremba [9,10] in general, that, if f is a BVHK function,

$$\delta_N(f) \leq \Sigma_{h=1}^{k} \Sigma_{\mathfrak{S}_h} \mathscr{D}_{N,\mathfrak{S}_h} V^h(f_{\mathfrak{S}_h}), \tag{27}$$

and Zaremba [9,10] proved also that

$$\delta_N(f) \leq \Sigma_{h=1}^{k} \Sigma_{\mathfrak{S}_h} \mathscr{T}_{N,\mathfrak{S}_h} W^h(f_{\mathfrak{S}_h}), \tag{28}$$

where

$$W^h(f_{\mathfrak{S}_h}) = \left(\int_{U^h} d\omega^h_{\mathfrak{S}_h} \left|\underset{\sim}{D}^h_{\mathfrak{S}_h} f_{\mathfrak{S}_h}\right|^2\right)^{1/2} \tag{29}$$

with $d\omega^h_{\mathfrak{S}_h}$ denoting the product of the dz_i selected by $\mathfrak{S}_h$, and $\underset{\sim}{D}^h_{\mathfrak{S}_h}$ denoting partial differentiation with respect to these z_i, once each. The bound in (28) is thus finite if all the $\underset{\sim}{D}^h_{\mathfrak{S}_h} f_{\mathfrak{S}_h}$ are bounded in quadratic mean (it suffices for this that $\underset{\sim}{D}^k_{z_1 z_2 \cdots z_k} f$ be bounded in U^k.)

Zaremba proves (27) and (28) in [10] by applying his generalization of the Abel transformation to show that

$$\int_{U^k} f(\underset{\sim}{z})\, d\Delta_N(\underset{\sim}{z}) = \Sigma_{h=1}^{k} (-1)^h \, \Sigma_{\mathfrak{S}_h} \int_{U^h} \Delta_{N,\mathfrak{S}_h}(\underset{\sim}{z}_{\mathfrak{S}_h})\, d_{\mathfrak{S}_h} f_{\mathfrak{S}_h}(\underset{\sim}{z}_{\mathfrak{S}_h}), \tag{30}$$

where $\underset{\sim}{z}_{\mathfrak{S}_h}$ is defined like $\underset{\sim}{\xi}_{n,\mathfrak{S}_h}$ as the vector with components selected by $\mathfrak{S}_h$, and $d_{\mathfrak{S}_h} f_{\mathfrak{S}_h}$ refers to integration with respect to the variation of these variables in f; the theorem holding for functions f which are BVHK. Now (1), (3), (9), and (22) combine to show that $\delta_N(f)$ equals the modulus of the left-hand side of (30); and we get (27) by applying the triangle inequality both to the sum and inside the integrals on the right-hand side of (30), and taking out the supremum $\mathscr{D}_{N,\mathfrak{S}_h}$ of $\Delta_{N,\mathfrak{S}_h}(\underset{\sim}{z}_{\mathfrak{S}_h})$, with (23); while we get (28) by applying the Cauchy-Schwarz inequality to each integral on the right-hand side of (30), with (29). We note that, if the $\underset{\sim}{D}^h_{\mathfrak{S}_h} f_{\mathfrak{S}_h}$ are bounded in mean, we can write (27) with

$$V^h(f_{\mathfrak{S}_h}) = \int_{U^h} d\omega^h_{\mathfrak{S}_h} \, |\underset{\sim}{D}^h_{\mathfrak{S}_h} f_{\mathfrak{S}_h}| . \tag{31}$$

3. Error-bounds for Quasi-Monte Carlo Integration

Let us suppose that we perform a quasi-Monte-Carlo integration by using a quasi-random sequence, whose L^∞ or L^2 discrepancy is known (or at least has a known upper bound), in the estimator t_N. Then (27) or (28) will give us an upper bound for the error made, and its determination will simply require that we estimate the measures of variation V^h or W^h for the functions $f_{\mathfrak{S}_h}$. If we have good reason to believe that the V^h or W^h exist (it is clear that, if some of the W^h exist and the rest of the V^h exist, we can form a mixed formula, intermediate between (27) and (28): each integral on the right of (30) can be bounded in either way, according to convenience), we can form Monte Carlo estimates of the appropriate integrals rather easily, especially if the partial derivatives in (29) and (31) are known or computable without too much trouble.

Now consider the sample mean variance s_N^2. By (1), (6), (11), and (22), we see that

$$|Ns_N^2 - \sigma^2| = \left| N^{-1} \Sigma_{n=0}^{N-1} |f(\underset{\sim}{\xi}_n)|^2 - |N^{-1} \Sigma_{n=0}^{N-1} f(\underset{\sim}{\xi}_n)|^2 - \int_{U^k} d\omega^k |f(\underset{\sim}{z})|^2 + \left|\int_{U^k} d\omega^k f(\underset{\sim}{z})\right|^2 \right| = \left| \int_{U^k} d\Delta_N(\underset{\sim}{z}) |f(\underset{\sim}{z})|^2 - \left|\int_{U^k} d\Delta_N(\underset{\sim}{z}) f(\underset{\sim}{z})\right|^2 - \int_{U^k} d\Delta_N(\underset{\sim}{z}) f(\underset{\sim}{z})\theta^* - \int_{U^k} d\Delta_N(\underset{\sim}{z}) f(\underset{\sim}{z})^*\theta \right|. \tag{32}$$

Applying (30) appropriately, we get, as the analog of (27), that

$$|Ns_N^2 - \sigma^2| \le \Sigma_{h=1}^k \Sigma_{\mathfrak{S}_h} \mathscr{D}_{N,\mathfrak{S}_h} [V^h(|f_{\mathfrak{S}_h}|^2) + 2|\theta| V^h(f_{\mathfrak{S}_h})] + [\Sigma_{h=1}^k \Sigma_{\mathfrak{S}_h} \mathscr{D}_{N,\mathfrak{S}_h} V^h(f_{\mathfrak{S}_h})]^2, \tag{33}$$

and the analog of (28) replaces $\mathscr{D}$ by $\mathscr{T}$ and V by W. We observe from this that the condition $\mathscr{D}_{N,\mathfrak{S}_h} \to 0$ as $N \to \infty$ for all h in (33), or the equivalent condition on $\mathscr{T}_{N,\mathfrak{S}_h}$ if f is appropriately differentiable, suffices to confirm rigorously our heuristic conjecture that

s_N^2 is a good estimator for σ_N^2. It follows from this that, when (as often occurs) $\delta_N(f)$ behaves like $N^{-1+\alpha}$ with $\alpha > 0$ arbitrarily small, while σ_N behaves like $N^{-1/2}$, s_N will give no indication of the error in the sample mean t_N.†

What is here proposed, as an alternative, is to abandon the sample mean variance s_N^2 and the estimator $s_N'^2$ (defined in (11) and (13)) as measures of accuracy, in Monte Carlo calculations of integrals in multi-dimensional intervals, using quasi-random sets of points; and to adopt, instead, the bounds given in (27) or (28) (when we have reasonable values or bounds for the discrepancies $\mathcal{D}_{N,\mathfrak{G}_h}$ or $\mathcal{T}_{N,\mathfrak{G}_h}$). For this purpose, at the same time as the estimator t_N is being computed (as defined in (3)), we would evaluate the estimators

$$v_{N,\mathfrak{G}_h} = v_{N,\mathfrak{G}_h}(\Xi) = N^{-1} \Sigma_{n=0}^{N-1} \left| D_{\underset{\sim}{\mathfrak{G}}_h}^h f_{\mathfrak{G}_h} \right| \tag{34}$$

or

$$w_{N,\mathfrak{G}_h} = w_{N,\mathfrak{G}_h}(\Xi) = (N^{-1} \Sigma_{n=0}^{N-1} \left| D_{\underset{\sim}{\mathfrak{G}}_h}^h f_{\mathfrak{G}_h} \right|^2)^{1/2}, \tag{35}$$

of $V^h(f_{\mathfrak{G}_h})$ or $W^h(f_{\mathfrak{G}_h})$, respectively; at least, if it is possible to compute the derivatives $D_{\underset{\sim}{\mathfrak{G}}_h}^h f_{\mathfrak{G}_h}(\underset{\sim}{z}_{\mathfrak{G}_h})$ without too much trouble.

.These considerations are illustrated by the one-dimensional examples given in the next section.

---

† Indeed, though the estimates afforded by approximating the right-hand sides of (27) or (28) may be more useful, they too will not give direct evidence of the magnitude of $\delta_N(f)$, since they are only <u>upper</u> bounds.

4. Examples

(a) We consider first the very simple example with $k = 1$ and $f(x) = 2z - 1$. Here $\theta = 0$ and $\sigma^2 = \int_0^1 (2z-1)^2\,dz = \frac{1}{3}$. Only in the one-dimensional case can we explicitly find the set of points, for any given N, for which $\mathscr{D}_N$ (and also $\mathscr{T}_N$) is least; namely,

$$\xi_0 = \frac{1}{2N},\quad \xi_1 = \frac{3}{2N},\quad \xi_2 = \frac{5}{2N},\ldots,\xi_{N-1} = \frac{2N-1}{2N}; \tag{36}$$

and for this set, $\mathscr{D}_N = 1/2N$, while $\mathscr{T}_N^2 = 2N \int_0^{1/2N} z^2 dz = 1/12N^2$. Taking these points, we find, in fact, that

$$t_N = N^{-1} \quad \Sigma_{n=0}^{N-1} \, [2(\frac{2n+1}{2N}) - 1] = 0; \tag{37}$$

and

$$s_N^2 = N^{-2} \; \Sigma_{n=0}^{N-1} \, [4(\frac{2n+1}{2N})^2 - 4(\frac{2n+1}{2N}) + 1] = \frac{N^2-1}{3N^3}, \tag{38}$$

so that

$$Ns_N^2 - \sigma^2 = -\,1/3N^2. \tag{39}$$

Turning to the variations, we see that (27) and (28) become simply

$$\left.\begin{aligned} &\delta_N(f) \le \mathscr{D}_N\, V(f) = V(f)/2N, \\ \text{and}\qquad &\\ &\delta_N(f) \le \mathscr{T}_N\, W(f)/2\sqrt{3}\, N. \end{aligned}\right\} \tag{40}$$

Now, $V(f) = 2$ and $W(f) = 2$. Thus the two upper bounds in (40) become $1/N$ and $1/\sqrt{3}\,N$; but actually $\delta_N(f) = 0$. The estimates v_N and w_N of $V(f)$ and $W(f)$, gotten from the points (36), will be exactly correct; because $f'(z)$ is constant.

(b) Now consider $k = 1$ with $f(z) = 3z^2 - 2z$. Then $\theta = \int_0^1 (3z^2 - 2z)\,dz = 0$ and $\sigma^2 = \int_0^1 (3z^2 - 2z)^2\,dz = \int_0^1 (9z^4 - 12z^3 + 4z^2)\,dz = \frac{2}{15}$. Taking the same optimal point-set (36), we get the same discrepancies: $\mathscr{D}_N = 1/2N$ and $\mathscr{T}_N = 1/2\sqrt{3}N$. With these points, we obtain the estimates

$$t_N = \frac{1}{N}\Sigma_{n=0}^{N-1}\,[3(\frac{2n+1}{2N})^2 - 2(\frac{2n+1}{2N})) = -\frac{1}{4N^2}\ , \tag{41}$$

and

$$s_N^2 = \frac{1}{N^2}\ \Sigma_{n=0}^{N-1}\,[9(\frac{2n+1}{2N})^4 - 12(\frac{2n+1}{2N})^3 + 4(\frac{2n+1}{2N})^2] - \frac{1}{N}\,(-\frac{1}{4N^2})^2$$

$$= \frac{2}{15N} - \frac{1}{3N^3} + \frac{1}{5N^5} = \frac{\sigma^2}{N} - \frac{1}{3N^3} + \frac{1}{5N^5}\ . \tag{42}$$

Since $f'(z) = 6z - 2$, we see that, for this example,

$$W(f) = \left(\int_0^1 (36z^2 - 24z + 4)\,dz\right)^{1/2} = 2, \tag{43}$$

while, by (31),

$$V(f) = \int_0^1 |6z - 2|\,dz = \int_0^{1/3} (2 - 6z)\,dz + \int_{1/3}^1 (6z - 2)\,dz$$

$$= 1/3 + (1 + 1/3) = 5/3. \tag{44}$$

These values yield the upper bounds

$$\mathscr{D}_N\, V(f) = 5/6N \quad \text{and} \quad \mathscr{T}_N\, W(f) = 1/\sqrt{3}N \tag{45}$$

for $\delta_N(f)$. In actual fact, however, we see from (41) that

$$\delta_N(f) = 1/4N^2. \tag{46}$$

It is of interest, finally, to consider again the estimates v_N and w_N of $V(f)$ and $W(f)$ obtained by the quasi-estimators analogous to t_N, using the same point-set (36). The estimate of $V(f)$ is then

$$v_N = \frac{1}{N}\, \Sigma_{n=0}^{N-1} \left| 6(\frac{2n+1}{2N}) - 2\right| = \frac{1}{N}[\Sigma_{n=0}^{m-1} (2 - \frac{3}{N} - 6\frac{n}{N}) + \Sigma_{n=m}^{N-1} (6\frac{n}{N} + \frac{3}{N} - 2)]$$

$$= \frac{1}{N}[(2-\frac{3}{N})m - \frac{3}{N}m(m-1) + \frac{3}{N}N(N-1) - \frac{3}{N}m(m-1) - (2-\frac{3}{N})(N-m)]$$

$$= (N^2 + 4mN - 6m^2)/N^2,$$

where m is the greatest integer not greater than $\frac{N}{3} + \frac{1}{2}$. It is readily verified that we must have $m = (N+j)/3$, where j is 0 or ± 1. The estimator of $V(f)$ thus reduces to

$$v_N = \frac{5}{3} - \frac{2j^2}{3N^2} = V(f) - \frac{2j^2}{3N^2} \leq V(f). \tag{47}$$

The estimate of $W(f)$ is similarly

$$w_N = \left(\frac{1}{N}\, \Sigma_{n=0}^{N-1} [36(\frac{2n+1}{2N})^2 - 24(\frac{2n+1}{2N}) + 4]\right)^{1/2}$$

$$= \frac{1}{N}[(12N^2 - 3) - 12N^2 + 4N^2]^{1/2} = 2(1 - \frac{3}{4N^2})^{1/2}$$

$$= 2 - \frac{3}{4N^2} - \frac{9}{64N^4} - \frac{27}{512N^6} - \cdots \leq W(f). \tag{48}$$

Thus both estimates are too small, but converge rapidly to the correct values.

References

1. B. V. Gnedenko. The Theory of Probability (Chelsea, New York, 1962).

2. J. H. Halton and S. K. Zaremba. The extreme and L^2 discrepancies of some plane sets. Monatsh. Math. 73 (1969) 316-328.

3. J. H. Halton. On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. Numer. Math. 2 (1960) 84-90, 196.

4. J. H. Halton. A retrospective and prospective survey of the Monte Carlo method. S.I.A.M. Rev. 12 (1970) 1-63.

5. J. M. Hammersley. Monte Carlo methods for solving multivariable problems. Ann. New York Acad. Sci. 86 (1960) 844-874.

6. E. Hlawka. Functions of bounded variation in the theory of uniform distribution. Ann. Mat. Pura Appl. IV:54 (1961) 325-334 (in German.)

7. J. F. Koksma. A general theorem from the theory of uniform distribution modulo 1. Mathematika Zutphen B11 (1942) 7-11 (in Dutch.)

8. K. F. Roth. On irregularities of distribution. Mathematika 1 (1954) 73-79.

9. S. K. Zaremba. The mathematical basis of Monte Carlo and quasi-Monte Carlo methods. S.I.A.M. Rev. 10 (1968) 303-314.

10. S. K. Zaremba. Some applications of multidimensional integration by parts. Ann. Polon. Math. 21 (1968) 85-96.