MATHEMATICS OF COMPUTATION, VOLUME 30, NUMBER 136
OCTOBER 1976, PAGES 831–837

Asymptotic Normality in Monte Carlo Integration

By Masashi Okamoto

Abstract. To estimate a multiple integral of a function over the unit cube, Haber proposed two Monte Carlo estimators J_1' and J_2' based on $2N$ and $4N$ observations, respectively, of the function. He also considered estimators D_1^2 and D_2^2 of the variances of J_1' and J_2', respectively. This paper shows that all these estimators are asymptotically normally distributed as N tends to infinity.

1. Introduction. Monte Carlo integration is a method to estimate the value of a definite integral of a given real-valued function over a finite region (say, a cube) by observing the value of the function only at a finite number of points in the region which are chosen suitably and stochastically. Kitagawa [4] proposed several estimating methods but he was concerned mainly with the case when the function has a certain prior distribution.

Haber [1], [2] proposed a mesh estimator of the integral and then improved it by means of the idea of "antithetic variates" due to Hammersley and Morton [3]. Specifically, let f be a real-valued function defined over the unit cube G_s in the s-space and set $I = \int_{G_s} f$. Let A_r $(r = 1, \ldots, N)$ be a family of congruent subcubes arising by partitioning G_s so that the interval $[0, 1]$ on the x^i-axis is divided into n_i equal subintervals for each $i = 1, \ldots, s$, where $N = n_1 \cdots n_s$. Let x_r be a random point in A_r chosen independently for each r and let $x_r' = 2c_r - x_r$, where c_r stands for the center of A_r. Then

$$J_1 = \frac{1}{N} \sum_{r=1}^{N} f(x_r) \tag{1.1}$$

and

$$J_2 = \frac{1}{N} \sum_{r=1}^{N} \frac{f(x_r) + f(x_r')}{2} \tag{1.2}$$

are unbiased estimators of I based on N and $2N$ observations of f, respectively.

To estimate the variances of J_1 and J_2 we need replications of observations. Let z_r be another random point in A_r, the z_r being chosen independently of each other and also of the x's. Define similarly $z_r' = 2c_r - z_r$. Then

$$J_1' = \frac{1}{N} \sum_{r=1}^{N} \frac{f(x_r) + f(z_r)}{2} \tag{1.3}$$

and

$$J_2' = \frac{1}{N} \sum_{r=1}^{N} \frac{f(x_r) + f(x_r') + f(z_r) + f(z_r')}{4} \tag{1.4}$$

Received August 26, 1975; revised January 27, 1976.
AMS (MOS) subject classifications (1970). Primary 65D30; Secondary 62E20, 60F05.
Key words and phrases. Monte Carlo integration, Haber's estimators, asymptotic normality, Lyapunov condition.

are again unbiased estimators of I based on $2N$ and $4N$ observations, respectively. By using these random points, the variances of J_1' and J_2' can be estimated unbiasedly by

$$D_1^2 = \frac{1}{4N^2} \sum_{r=1}^{N} \{f(x_r) - f(z_r)\}^2 \tag{1.5}$$

and

$$D_2^2 = \frac{1}{4N^2} \sum_{r=1}^{N} \left\{\frac{f(x_r) + f(x_r')}{2} - \frac{f(z_r) + f(z_r')}{2}\right\}^2, \tag{1.6}$$

respectively.

For $k = 1$ and 2 let C^k denote the set of all real-valued functions defined over G_s and having continuous kth order partial derivatives. In the sequel we say just "$N \longrightarrow \infty$" to indicate that $n_i \longrightarrow \infty$ for every $i = 1, \ldots, s$. Put $n = \min(n_1, \ldots, n_s)$. Haber proved that if $f \in C^1$, then

$$\operatorname{var}(J_1) = \tau_N^2(J_1) + o((Nn^2)^{-1}) \quad \text{as } N \longrightarrow \infty, \tag{1.7}$$

and also that if $f \in C^2$, then

$$\operatorname{var}(J_2) = \tau_N^2(J_2) + o((Nn^4)^{-1}) \quad \text{as } N \longrightarrow \infty, \tag{1.8}$$

where

$$\tau_N^2(J_1) = \frac{1}{12N} \sum_{i=1}^{s} \frac{1}{n_i^2} \int_{G_s} \left(\frac{\partial f}{\partial x^i}\right)^2,$$

$$\tau_N^2(J_2) = \frac{1}{1440N} \left\{2 \sum_{i=1}^{s} \frac{1}{n_i^4} \int_{G_s} \left(\frac{\partial^2 f}{\partial (x^i)^2}\right)^2 + 5 \sum_{i \neq j = 1}^{s} \frac{1}{(n_i n_j)^2} \int_{G_s} \left(\frac{\partial^2 f}{\partial x^i \partial x^j}\right)^2\right\}. \tag{1.9}$$

Furthermore, since J_k' $(k = 1, 2)$ is the average of two independent realizations of J_k, its variance is half as large as that of J_k. In estimating I by J_k', Haber used D_k as a measure of the error, assuming that J_k' is approximately normally distributed.

The purpose of this paper is to show, first, that J_k $(k = 1, 2)$ is asymptotically normally distributed with mean I and variance $\tau_N^2(J_k)$ as $N \longrightarrow \infty$. This implies the asymptotic normality of J_k'. Next, it is shown that D_k^2 $(k = 1, 2)$ is asymptotically normally distributed with mean $\operatorname{var}(J_k')$ and a proper variance.

2. Asymptotic Normality of J_1 and J_2. Before stating the theorems we introduce some notations. Let

$$I_r = E\{f(x_r)\} = N \int_{A_r} f,$$

$$\delta_r = (\delta_r^1, \ldots, \delta_r^s) = x_r - c_r,$$

$$f_r = f(c_r), \quad f_r^i = \frac{\partial f}{\partial x^i}(c_r) \quad \text{and} \quad f_r^{ij} = \frac{\partial^2 f}{\partial x^i \partial x^j}(c_r),$$

where $x = (x^1, \ldots, x^s)$, for $i, j = 1, \ldots, s$, and $r = 1, \ldots, N$. The L_2-norm will be denoted by $\| \ \|$.

THEOREM 1. *Assume $f \in C^1$. As $N \longrightarrow \infty$, J_1 is asymptotically normally distributed with mean I and variance $\tau_N^2(J_1)$ defined in* (1.9).

Proof. Since $I = \Sigma_{r=1}^N I_r/N$, it holds that

$$J_1 - I = \frac{1}{N} \sum_{r=1}^{N} \{f(x_r) - I_r\}. \tag{2.1}$$

A Taylor series expansion of $f(x_r)$ around the point c_r gives

$$f(x_r) = f_r + \sum_{i=1}^{s} \delta_r^i f_r^i + R_{1r}, \tag{2.2}$$

where the remainder R_{1r} has the following property in view of the uniform continuity of $\partial f/\partial x^i$ in G_s: for any $\epsilon > 0$, $|R_{1r}| \leqslant \epsilon \|\delta_r\|$ for every r, provided N is sufficiently large.

Again, by a Taylor expansion we have

$$I_r = f_r + R_{2r}, \tag{2.3}$$

where $|R_{2r}| \leqslant \epsilon/n$ for every r, provided N is sufficiently large. Substitution of (2.2) and (2.3) into (2.1) yields

$$J_1 - I = S_N + R_N, \tag{2.4}$$

where

$$S_N = \frac{1}{N} \sum_{r=1}^{N} \sum_{i=1}^{s} \delta_r^i f_r^i, \qquad R_N = \frac{1}{N} \sum_{r=1}^{N} (R_{1r} - R_{2r}). \tag{2.5}$$

Since $E(J_1) = I$, to prove the theorem it remains only to verify the following three propositions:

(i) $\operatorname{var}(S_N) \sim \tau_N^2(J_1)$ as $N \longrightarrow \infty$,

where the symbol $\sim$ means that the ratio of the two sides tends to one,

(ii) the sequence S_N satisfies the Lyapunov condition of the central limit theorem (see, e.g., Loève [5, p. 275]), and

(iii) $\operatorname{var}(R_N) = o((Nn^2)^{-1})$ as $N \longrightarrow \infty$.

In fact, in the right-hand side of the identity

$$\frac{J_1 - I}{\tau_N} = \frac{S_N - E(S_N)}{\tau_N} + \frac{R_N - E(R_N)}{\tau_N}$$

the first term converges to the standardized normal distribution because of (i) and (ii), whereas the second term converges in probability to zero because of (iii) and the Chebyshev inequality.

Now, part (i) is essentially equivalent to Theorem 3 of Haber [1] or (1.7); but a proof is given here for the completeness of the proof. By definition, δ_r^i has a uniform distribution over the interval $[-1/(2n_i), 1/(2n_i)]$ independently of each other; and hence, $\operatorname{var}(\delta_r^i) = 1/(12n_i^2)$, which implies that

$$\operatorname{var}(S_N) = \frac{1}{12N^2} \sum_{r=1}^{N} \sum_{i=1}^{s} \frac{1}{n_i^2}(f_r^i)^2 \sim \tau_N^2(J_1) \quad \text{as } N \longrightarrow \infty.$$

Re (ii). Define $X_{Nr} = \Sigma_{i=1}^s \delta_r^i f_r^i$, then

$$\sum_r E(X_{Nr}^2) = N^2 \operatorname{var}(S_N). \tag{2.6}$$

From the inequality $|\delta_r^i| \leqslant 1/(2n_i)$ it follows that

$$\sum_r E|X_{Nr}|^3 \leqslant \frac{1}{8}\sum_r \Big(\sum_i |f_r^i|/n_i\Big)^3 \sim \frac{N}{8}\int \Big(\sum_i \frac{1}{n_i}\left|\frac{\partial f}{\partial x^i}\right|\Big)^3 = O(Nn^{-3}). \tag{2.7}$$

By (i), (2.6) and (2.7)

$$\Big[\sum_r E|X_{Nr}|^3\Big] \Big/ \Big[\sum_r E(X_{Nr}^2)\Big]^{3/2} = O(N^{-1/2}) = o(1),$$

which is the Lyapunov condition for S_N.

Re (iii). Since R_{2r} are constant,

$$\operatorname{var}(R_N) = \sum_r \operatorname{var}(R_{1r})/N^2 \leqslant \sum_r E(R_{1r}^2)/N^2 \leqslant \epsilon^2/(Nn^2).$$

This implies (iii), since ϵ can be made arbitrarily small. Q.E.D.

Remark. The asymptotic normality may be proved by applying the central limit theorem directly to J_1 in the form (1.1), not indirectly to S_N. This approach, however, requires asymptotic expansions of $E\{f^2(x_r)\}$ and I_r up to the terms of order n^{-2} so that a stronger assumption $f \in C^2$ is needed instead of $f \in C^1$. This remark is valid also with Theorems 2, 3 and 4 for which a much stronger assumption $f \in C^k$ ($k = 4$, 4 and 6, respectively) is required.

THEOREM 2. *Assume $f \in C^2$. As $N \longrightarrow \infty$, J_2 is asymptotically normally distributed with mean I and variance $\tau_N^2(J_2)$ defined in* (1.9).

Proof. Similarly as (2.1)

$$J_2 - I = \frac{1}{2N} \sum_{r=1}^{N} \{f(x_r) + f(x_r') - 2I_r\}. \tag{2.8}$$

Expanding $f(x_r)$ and $f(x_r')$ in Taylor series around the point c_r, we find

$$\begin{aligned} f(x_r) &= f_r + \sum_{i=1}^{s} \delta_r^i f_r^i + \frac{1}{2} \sum_{i,j=1}^{s} \delta_r^i \delta_r^j f_r^{ij} + R_{1r}, \\ f(x_r') &= f_r - \sum_{i=1}^{s} \delta_r^i f_r^i + \frac{1}{2} \sum_{i,j=1}^{s} \delta_r^i \delta_r^j f_r^{ij} + R_{1r}', \end{aligned} \tag{2.9}$$

where $|R_{1r}|$ and $|R_{1r}'|$ are bounded from above by $\epsilon\|\delta_r\|^2$ for every r, provided N is sufficiently large.

Taking one more term in the expansion (2.3), we have

$$I_r = f_r + \sum_{i=1}^{s} \frac{1}{24n_i^2} f_r^{ii} + R_{2r}, \tag{2.10}$$

where $|R_{2r}| \leqslant \epsilon/n^2$ for every r. Substitution of (2.9) and (2.10) into (2.8) yields

$$J_2 - I = S_N + R_N,$$

where

(2.11)
$$S_N = \frac{1}{2N} \sum_{r=1}^{N} \left(\sum_{i,j=1}^{s} \delta_r^i \delta_r^j f_r^{ij} - \sum_{i=1}^{s} \frac{1}{12n_i^2} f_r^{ii} \right),$$
$$R_N = \frac{1}{2N} \sum_{r=1}^{N} (R_{1r} + R'_{1r} - 2R_{2r}).$$

Just as for Theorem 1 we have only to prove the following three propositions:

(i) $\operatorname{var}(S_N) \sim \tau_N^2(J_2)$ as $N \longrightarrow \infty$,

(ii) S_N satisfies the Lyapunov condition, and

(iii) $\operatorname{var}(R_N) = o((Nn^4)^{-1})$ as $N \longrightarrow \infty$.

Part (i) is equivalent to the main theorem in Haber [2] or (1.8), while (ii) and (iii) can be verified by using reasoning similar to that in the proof of Theorem 1. Q.E.D.

COROLLARY 1. *For $k = 1$ or 2, assume $f \in C^k$. As $N \longrightarrow \infty$, J'_k follows asymptotically a normal distribution with mean I and variance*

(2.12)
$$\tau_N^2(J'_k) = \frac{1}{2}\, \tau_N^2(J_k).$$

3. Asymptotic Normality of D_1 and D_2. Though the asymptotic normality of D_1 or D_2 may not be so important in practice as that of any estimator of I itself, it can be proved along a similar line of arguments for the latter. First let us consider D_1^2.

THEOREM 3. *Assume $f \in C^1$. As $N \longrightarrow \infty$, D_1^2 is asymptotically normally distributed with mean* $\operatorname{var}(J'_1)$ *and variance*

$$\tau_N^2(D_1^2) = \frac{1}{2880N^3} \left\{ 7 \sum_{i=1}^{s} \frac{1}{n_i^4} \int_{G_s} \left(\frac{\partial f}{\partial x^i} \right)^4 + 10 \sum_{i \neq j=1}^{s} \frac{1}{(n_i n_j)^2} \int_{G_s} \left(\frac{\partial f}{\partial x^i} \frac{\partial f}{\partial x^j} \right)^2 \right\}.$$

Proof. Similarly to (2.2) it is the case that

(3.1)
$$f(z_r) = f_r + \sum_{i=1}^{s} \zeta_r^i f_r^i + R_{2r},$$

where

$$\zeta_r = (\zeta_r^1, \ldots, \zeta_r^s) = z_r - c_r, \qquad |R_{2r}| \leqslant \epsilon \|\zeta_r\| \quad \text{for every } r.$$

Substitution of (2.2) and (3.1) into (1.5) yields

(3.2)
$$D_1^2 = T_N + R_N,$$

where

(3.3)
$$T_N = \frac{1}{4N^2} \sum_{r=1}^{N} \left(\sum_{i=1}^{s} \eta_r^i f_r^i \right)^2,$$
$$R_N = \frac{1}{4N^2} \sum_{r=1}^{N} \left\{ (R_{1r} - R_{2r})^2 + 2(R_{1r} - R_{2r}) \sum_{i=1}^{s} \eta_r^i f_r^i \right\}$$

and $\eta_r^i = \delta_r^i - \zeta_r^i$.

Since $E(D_1^2) = \text{var}(J_1')$, a proof of Theorem 3 can be reduced to the following:

(i) $\text{var}(T_N) \sim \tau_N^2(D_1^2)$ as $N \longrightarrow \infty$,

(ii) T_N satisfies the Lyapunov condition, and

(iii) $\text{var}(R_N) = o((N^3 n^4)^{-1})$ as $N \longrightarrow \infty$.

Re (i). First we have

$$\text{var}(T_N) = \frac{1}{16N^4} \sum_r \text{var}\left(\sum_i \eta_r^i f_r^i\right)^2. \tag{3.4}$$

Since for every i, j and r

$$E(\eta_r^i)^2 = \frac{1}{6n_i^2}, \qquad E(\eta_r^i)^4 = \frac{1}{15n_i^4},$$

$$E(\eta_r^i \eta_r^j)^2 = \frac{1}{36n_i^2 n_j^2} \qquad (i \neq j);$$

and since the expectations of any other monomial of the η's of order 2 or 4 vanish, we find

$$E\left(\sum_i \eta_r^i f_r^i\right)^2 = \frac{1}{6} \sum_i \left(\frac{f_r^i}{n_i}\right)^2,$$

$$E\left(\sum_i \eta_r^i f_r^i\right)^4 = \frac{1}{15} \sum_i \left(\frac{f_r^i}{n_i}\right)^4 + \frac{1}{12} \sum_{i \neq j} \left(\frac{f_r^i f_r^j}{n_i n_j}\right)^2.$$

This implies

$$\text{var}\left(\sum_i \eta_r^i f_r^i\right)^2 = \frac{1}{180} \left\{7 \sum_i \left(\frac{f_r^i}{n_i}\right)^4 + 10 \sum_{i \neq j} \left(\frac{f_r^i f_r^j}{n_i n_j}\right)^2\right\}.$$

Substitution of the last formula into (3.4) proves (i).

Re (ii). Define

$$X_{Nr} = \left(\sum_{i=1}^{s} \eta_r^i f_r^i\right)^2.$$

Then a straightforward calculation gives

$$\sum_r E|X_{Nr}|^3 = O(Nn^{-6});$$

and hence,

$$\left[\sum_r E|X_{Nr}|^3\right] \Big/ \left[\sum_r E(X_{Nr}^2)\right]^{3/2} = O(N^{-1/2}) = o(1).$$

Re (iii). From (3.3) it follows that

$$|R_N| \leqslant \frac{\epsilon^2}{4N^2} \sum_r \left\{(\|\delta_r\| + \|\zeta_r\|)^2 + 2(\|\delta_r\| + \|\zeta_r\|) \frac{1}{n} \sum_i |f_r^i|\right\};$$

and hence,

$$\text{var}(R_N) \leqslant E(R_N^2) = o((N^3 n^2)^{-1}). \quad \text{Q.E.D.}$$

Remark. Though $\tau_N^2(J_1')$ defined in (2.12) is indeed the leading term in the asymptotic expansion of $\text{var}(J_1')$ in n^{-1}, the phrase "mean $\text{var}(J_1')$" in the statement of Theorem 3 cannot be replaced by "mean $\tau_N^2(J_1')$", in general. The reason is that the difference $\mathrm{var}(J_1') - \tau_N^2(J_1')$ is $o((Nn^2)^{-1})$ and hence is not negligible in general as compared with $\tau_N(D_1^2)$ which is of order $(N^{3/2}n^2)^{-1}$.

Similarly, we can prove the following:

THEOREM 4. *Assume $f \in C^2$. As $N \longrightarrow \infty$, D_2^2 is asymptotically normally distributed with mean* $\mathrm{var}(J_2')$ *and variance $\tau_N^2(D_2^2)$ which is a linear combination of seven terms of the form*

$$\sum_{i,j,k,l=1}^{s} \frac{1}{N^3(n_i n_j n_k n_l)^2} \int_{G_s} \left(\frac{\partial^2 f}{\partial x^i \partial x^j} \frac{\partial^2 f}{\partial x^k \partial x^l} \right)^2 ,$$

where some of i, j, k and l are constrained to be equal.

The root-square transformation of the random variable D_k^2 $(k = 1, 2)$ induces the following:

COROLLARY 2. *For $k = 1$ or 2, assume $f \in C^k$. As $N \longrightarrow \infty$, D_k is asymptotically normally distributed with mean $\sigma(J_k')$ and variance* $\frac{1}{4}\tau_N^2(D_k^2)/\mathrm{var}(J_k')$, *where $\sigma(J_k')$ stands for the standard deviation of J_k'.*

Department of Applied Mathematics
Faculty of Engineering Science
Osaka University
Toyonaka, Osaka 560, Japan

1. S. HABER, "A modified Monte-Carlo quadrature," *Math. Comp.*, v. 20, 1966, pp. 361–368. MR **35** # 1178.

2. S. HABER, "A modified Monte-Carlo quadrature. II," *Math. Comp.*, v. 21, 1967, pp. 388–397. MR **38** # 2922.

3. J. M. HAMMERSLEY & K. W. MORTON, "A new Monte Carlo technique: Antithetic variates," *Proc. Cambridge Philos. Soc.*, v. 52, 1956, pp. 449–475. MR **18**, 336.

4. T. KITAGAWA, "Random integrations," *Bull. Math. Statist.*, v. 4, 1950, pp. 15–21. MR **14**, 457.

5. M. LOÈVE, *Probability Theory*, 3rd ed., Van Nostrand, Princeton, N.J., 1963. MR **34** # 3596.