# Random integrations

Kitagawa, Toshio
Mathematical Institute, Faculty of Science, Kyushu University

# RANDOM INTEGRATIONS[1]

## By Tosio KITAGAWA

(*Mathematical Institute, Faculty of Scieuce, Kyushu University*)

**1. Introduction.** The various current studies of comparisions of the relative efficiencies of mechanical quadratures concern themselves with the functional schemes from the deterministic point of views.[2] Neverthless it may be noticed that such efficiencies should not be judged merely from the view point of individual functions, but rather from a broad family of functions to which these methods are to be applied. This means that true criterion of their efficiencies should be relevantly given from a statistical point of view. On the other hand some theories of samples from processes depending upon continuous parameters are now required in designing various sampling surveys.[3] This will also make it necessary to introduce some type of integrations which should be dealt with from a statistical point of view. The object of this paper is to give preliminary answers to these requirements by introducing certain types of random integrations.

In what follows we shall assume that every individual function $f(t)$ in our consideration is defined and continuous in the closed interval $0 \leq t \leq 1$, and that our approximations to the integral of $f(t)$ over the interval $0 \leq t \leq 1$ are finite sums of the form

$$(1.01) \qquad S_A(f) = \sum_{i=1}^{n} A_i f(t_i),$$

where $0 \leq t_1 < t_2 < t_3 < \ldots t_n \leq 1$ and

$$(1.02) \qquad \sum_{i=1}^{n} A_i = 1, \qquad A_i \geq 0 \quad (i = 1, 2, \ldots, n).$$

We shall propose to introduce three types of random integrations in the following manner:

TYPE A. In this case a definite function $f(t)$ is assumed to be fixed. A randomness is introduced in selecting the division-points $0 \leq t_1 < t_2 < t_3 < \ldots < t_n \leq 1$ and hence sometimes the coefficients $A_i$ $(i = 1, 2, \ldots, n)$.

TYPE B. In this case a family of individual functions is assumed to be given with the suitably defined probability measure in itself. The method of approximations by finite sums is assumed to be fixed. Thus the sum $S_A(x)$ may be recognised as random variable when the function $x$ runs through

(1) Communicated at the Meeting of Kyushu Branch Society of Japanese Math Soc. in February 4th, 1950.

(2) See for example the following literatures: C. RUNGE AND H. KONIG, Numerisches Rechnen, 1924. FR. A. WILLERS, Methoden der praktischen Analysis, 1928, L. M. MILNE, Numerical calculus, 1949.

(3) See KITAGAWA [1].

the function-family.

TYPE C. This case is the combination of two randomnesses in Types A and B. We shall consider both a function-family with its probability measure and division-family with its certain probabilistic scheme.

**2. Random integration of type (A).** The formulations of random integration of type (A) depend upon two kinds of procedures. The first procedure concerns itsself with a method of sampling of $t$ points, while the second with a formulation of finite sums. Here we shall restrict ourselves with the following three procedures of samplings of $t$ points:

$(a_1)$ : a sample of $n$ $t$-points is a purely random one chosen from the interval $0 \leq t \leq 1$ (purely random sampling).

$(a_2)$ : the interval, $0 \leq t \leq 1$, is divided into $n$ strata of the equal length $\delta = 1/n$, and a sample $(t_1, t_2, \ldots, t_n)$ is such that each $t_k (1 \leq k \leq n)$ is randomly drawn from the subinterval $(k-1) \delta \leq t \leq k \delta$ (stratified random sampling).

$(a_3)$ : a point $t$ is drawn randomly from the interval $0 \leq t \leq \delta (=1/n)$, and then a sample $(t_1, t_2, \ldots t_n)$ is determined as $t_k = t_1 + (k-1)\delta$. $(k=2, 3, \ldots, n)$ (systematic sampling).

Each of these sampling procedures will define a certain probability measure of the $n$-dimensional euclidian space $(t_1, t_2, \ldots t_n)$, which we shall express by $(\tau_i)$ respectively $(i=1, 2, 3)$ corresponding to each of $(a_i)$ $(i=1, 2, 3)$.

As to the formulations of finite sums we shall take into consideration the following two types:

$(b_1)$ : let $t_{(1)} \leq t_{(2)} \leq \ldots \leq t_{(n)}$ be the order statistics determined from the sample $(t_1, t_2, \ldots, t_n)$ of each of types $(a_1)$, $(a_2)$ and $(a_3)$, and we shall put

$$(2.01) \qquad S^{(1)}(f) = \sum_{k=1}^{n+1} (t_{(k)} - t_{(k-1)}) f(t_{(k)}),$$

where we put, for the sake of convenience, $t_{(0)} = 0$, $t_{(n+1)} = 1$.

$(b_2)$ : this means the arithmetic mean

$$(2.02) \qquad S^{(2)}(f) = \frac{1}{n} \sum_{k=1}^{n} f(t_{(k)}).$$

In combination of these two procedures there may be six $(=2 \times 3)$ combinations of procedures, $(a_i b_j)$ $(i=1, 2, 3; j=1, 2)$ each of which belongs to a random integration of type A. Our chief concerns are the relative efficiencies of these six finite sums as the estimates of the integral of $f(x)$ over the interval $0 \leq t \leq 1$.

These six estimates are unbiassed estimates of the integral, and hence the comparisons of variances of the estimates may be adopted as most elementary criterion. For this purpose we shall first notice the following

THEOREM 1. *Let $f(x)$ be a continuous function defined over the interval $0 \leq x \leq 1$. Then the characteristic functions (c. f.) of the stochastic variables defined by the finite sum $(a_\nu b_j)$ $(\nu=1, 2; j=1, 2)$ are given by the integral*

$$(2.03) \qquad E_{D\nu}\left\{ e^{it \, S_\nu^{(j)}(f)} \right\} = \int\int_{D\nu} \ldots \int \exp\left\{ i\, t\, S^{(j)}(f) \right\} dt_{(1)} dt_{(2)} \ldots dt_{(n)}$$

*where* $S_\nu^{(j)}(f)$ *(* $j=1,2$ *) are defined by* (2.01) *and* (2.02), *and the domain of integration* $D_\nu$ *(* $\nu=1,2,3$ *) are defined by*

(2.04)     $D_1$: $0 \leq t^{(1)} \leq t^{(2)} \leq ... \leq t^{(n)} \leq 1$

(2.05)     $D_2$: $(k-1)\delta \leq t^{(k)} \leq k\delta$ $(k=1,2,...,n)$,

*while those c.f. corresponding to* $(a_3 b_2)$ *is given by*

(2.06)     $E_{D_3} \left\{ e^{itS_3^{(1)}(f)} \right\} = \int_0^\delta \exp \left\{ i\,t\,\delta \sum_{k=1}^n f(x_1 + (k-1)\delta) \right\} dx_1.$

This theorem can be readily seen from the definitions of *c.f.*, and will yield us the following evaluations of mean values and variances of the stochastic variables in our concern.

In what follows the mean values and variances of the stochastic variables defined by $(a_\nu, b_j)$ $(\nu=1,2,3; j=1,2)$ will be denoted by $E_{D\nu}\{S^{(j)}\}$ and $V_{D\nu}\{S^{(j)}\}$ respectively. Then we shall observe readily from Theorem 1 the following

THEOREM 2. *Under the hypothesis to Theorem 1, we shall have*

(2.071) $E_{D_1}\left\{ S_1^{(1)}(f) \right\} = \sum_{k=1}^{n+1} \underset{\substack{\tau_1+\tau_2+...+\tau_n=1 \\ \tau_i \geq 0}}{\int \int \cdots \int} \tau_k f(\tau_1+\tau_2+...+\tau_k) d\tau_1 d\tau_2 ... d\tau_n$

(2.072) $V_{D_1}\left\{ S_1^{(1)}(f) \right\} = \sum_{k=1}^{n+1} \underset{\substack{\tau_1+\tau_2+...+\tau_n=1 \\ \tau_i \geq 0}}{\int \int \cdots \int} \tau_k^2 f(\tau_1+\tau_2+...+\tau_k)^2 d\tau_1 d\tau_2 ... d\tau_n$

$\qquad\qquad + 2 \underset{\substack{\tau_1+\tau_2+...\tau_n=1 \\ (i,k)}}{\sum \sum} \tau_i \tau_k f(\tau_1+\tau_2+...+\tau_i) f(\tau_1+...+\tau_k) \cdot d\tau_1 d\tau_2 ... d\tau_n$

$\qquad\qquad\qquad\qquad - \left( E_{D_1}\left\{ S^{(1)} \right\} \right)^2$

(2.081) $E_{D_1}\left\{ S_1^{(2)}(f) \right\} = \int_0^1 f(x) dx$

(2.082) $V_{D_1}\left\{ S_1^{(2)}(f) \right\} = \frac{1}{n} \left[ \int_0^1 \left\{ f(x) \right\}^2 dx - \left( \int_0^1 f(x) dx \right)^2 \right]$

(2.091) $E_{D_2}\left\{ S_2^{(2)}(f) \right\} = \int_0^1 f(x) dx$

(2.092) $V_{D_2}\left\{ S_2^{(2)}(f) \right\} = \frac{1}{n} \left[ \int_0^1 \left\{ f(x) \right\}^2 dx - \frac{1}{\delta} \sum_{k=1}^n \left\{ \int_{(k-1)\delta}^{k\delta} f(x) dx \right\}^2 \right]$

(2.101) $E_{D_3}\left\{ S_3^{(2)}(f) \right\} = \int_0^1 f(x) dx,$

and

$$(2.102) \quad V_{D_3}\left\{S_3^{(2)}(f)\right\} = \frac{1}{n}\int_0^1 \left[\frac{1}{n}\sum_{k=1}^{n}f\left(\frac{y+k-1}{n}\right) - \int_0^1 f(x)dx\right]^2 dy.$$

It will be noted that the variances depend upon the functional behaviours of the individual functions $f(x)$ such as periodicities. The detailed comparisons, taking elementary functions as their illustrations, will give us some informations about comparisons of relative efficiencies of these finite sums.

**3. Random integration of the type (B).** Here we shall consider the family of all continuous functions defined over the interval $0 \leq t \leq 1$, and we shall make use of Wiener measure[4] recently developed by CAMERON and MARTIN. We shall among others appeal to the following Lemma, which we owe to CAMERON-MARTIN [1] to the following effect:

LEMMA[5] *Let $\beta(t)$ be real-valued and of bounded variations over $0 \leq t \leq 1$, with $\beta(1)=0$, and let us put*

$$(3.01) \qquad A = \sqrt{\int_0^1 [\beta(t)]^2 dt}.$$

*Let $G(u)$ be a measurable function defined in $-\infty < u < \infty$. Let us denote by $x(t)$ an element of Wiener space.*

*Then we shall have the following assertions:*

*(a) the necessary and sufficient condition that*

$$(3.02) \qquad G\left[\int_0^1 x(t)d\beta(t)\right]$$

*should be summable in the sense of Wiener, is that $e^{-u^2}G(Au)$ belong to $L_1(-\infty, \infty)$, that is, integrable in the sense of Lebesgue in $-\infty < u < \infty$.*

*(b) Under the condition to (a), the equality holds:*

$$(3.03) \qquad \int_C G\left[\int_0^1 x(t)d\beta(t)\right] d_\omega x = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-u^2}G(Au)du,$$

*where the left-hand integration $d_\omega x$ means that of Wiener over the function space $(C)$.*

*Specially when $G(u) = \exp\{i u \lambda\}$, (3.03) becomes*

$$(3.04) \qquad \int_C \exp\left\{i\lambda\int_0^1 x(t)d\beta(t)\right\} = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-u^2}e^{i\lambda uA}du.$$

The comparisons of relative efficiencies of various types of finite sums usually adopted in the theory of numerical integrations, such as those of

---

(4) N. WIENER [1] defines an average or integral over the space $C$ of all functions $x(t)$ continuous in $0 \leq t \leq 1$ and vanishing at $t=0$. The theory of this integration was developed in PALEY-WIENER [1], and recently more throughly by CAMERON-MARTIN in a series of numerous important papers.

(5) See CAMERON-MARTIN [1], Lemma 3.

NEWTON-COTES, MACLAURIN, TSCHEBYCHEFF and GAUSS, as estimates of the integral of $x(t)$ over $0 \leq t \leq 1$ will be discussed as random integration of type (B). This means, for example, the comparisons of the variances of the linear functional

$$(3.05) \qquad \int_0^1 x(t) d_i \beta(t)$$

in the Wiener measure, whose characteristic function is given by (3.04). Indeed for any fixed system of $n$ selected points $0 \leq t_1 < t_2 < \ldots < t_n \leq 1$, we may and we shall write

$$(3.06) \qquad \sum_{k=1}^n A_k x(t_k) = \int_0^1 x(t) d\alpha(t)$$

by chossing a step-function $\alpha(t)$, and therefore

$$(3.07) \qquad \int_0^1 x(t) dt - \sum_{k=1}^n A_k x(t_k) = \int_0^1 x(t) d_i \beta(t),$$

where we put

$$(3.08) \qquad \beta(t) = t - \alpha(t).$$

It is however to be noted that the element of the space $C$ is restricted with continuous function in $0 \leq t \leq 1$ vanishing at $t = 0$. Consequenty we shall consider the function space $C'$ which may be expressed as $a_0 + x(t)$, where $x(t)$ belong to the Wiener space $C$ and $a_0$ is a random variable independent of $t$ and the function $x$.

In view of the above mentioned Theorem, we shall have

THEOREM 3. *The variance of the random integral of the type (B) defined in the function space (C) with Wiener measure is given by the integral*

$$(3.09) \qquad \frac{1}{2} \int_0^1 [t - \alpha(t)]^2 dt,$$

*where $\alpha(t)$ is defined by (3.06), that is to say, a step-function associated to the formulae of the finite sum.*

It may seem, however, unsuitable to discuss a theory of numerical integrations merely in a family of all continuous functions and to compare relative efficiencies of their procedures in that family. In fact the current theories usually concern themselves with the functions which are continuously differentiable at least to certain higher orders. This will lead us to the adoptions of family of all the continuous functions which are continuously differentiable to a certain $p$-th order. Then we may and we shall consider a function of the type

$$(3.10) \qquad y(t) = a_0 + a_1 t + a_2 t^2 + \ldots + a_{p-1} t^{p-1} + \int_0^t \frac{(t-s)^{p-1}}{(p-1)!} x(s) ds.$$

In order to introduce a random integration of the type (B) we shall assume here that

*Condition* (1°). $a_0, a_1, \ldots, a_{p-1}$ are mutually independent stochastic variables with their finite variances

*Condition* (2°). $x(s)$ belongs to the function space (C) with Wiener measure.

*Condition* (3°). $x(s)$ runs through (C) independently with $(a_0, a_1, a_2, \ldots, a_{p-1})$.

We shall then have

$$(3.11) \quad \int_0^1 y(t)dt - \sum_{k=1}^n A_k y(t_k)$$

$$= \sum_{r=1}^{p-1} a_r \left\{ \frac{1}{r+1} - \sum_{k=1}^n A_k t_k^r \right\} + \frac{1}{p!} \int_0^1 (1-p)^p x(s)ds - \sum_{k=1}^n A_k \int_0^{t_k} \frac{(t_k-s)^{p-1}}{(p-1)!} x(s)ds$$

$$= \sum_{r=1}^{p-1} a_r \left[ \frac{1}{r+1} - \sum_{k=1}^n A_k t_k^r \right] + \int_0^1 x(s)d\beta_p(s), \quad say.$$

This will lead us to the following

THEOREM 4. *For the function-space consisting of $y(t)$ defined in* (3.10) *with the conditions* (1°), (2°) *and* (3°), *the variance of the random integral of the type* (B) *is given by*

$$(3.12) \quad V\left\{ \sum_{k=1}^n A_k y(t_k) \right\}$$

$$= \sum_{k=1}^{p-1} \left[ \frac{1}{r+1} - \sum_{k=1}^n A_k t_k^r \right]^2 \sigma_{r_i}^2 + \frac{1}{2} \int_0^1 [\beta_p(t)]^2 \, dt,$$

*where $\sigma_r^2$ is defined in condition* (1°) *and the function $\beta_p(t)$ defined in* (3.11).

**4. Random integration of the type** (C). The direct combination of two random integrations (A) and (B) will yield us a random integration of the type (C), which may be considered as stochastic variable in the product space $(\tau) \times (C)$ with the product probability measure. This stochastic variable gives an unbiassed estimator of the integral of the original function, if it holds true that

$$(4.01) \qquad \underset{(\tau) \times (C)}{E} \left\{ \sum_{i=1}^n A_i x(t_i) \right\} = \underset{(C)}{E} \left\{ \int_0^1 x(t)dt \right\}.$$

So far as (4.01) holds, our concern may be turned to the variances of the estimator, which is defined as

$$(4.02) \qquad \underset{(\tau)}{E} \left\{ \frac{1}{2} \int_0^1 [\beta_\tau(t)]^2 dt \right\}.$$

It is to be noted here that the function $\beta_\tau(t)$ which is determined by the formulation of finite sum is in itsself a stochastic process. This follows from the fact that the division points $t_1, t_2, \ldots, t_{n-1}$ and $t_n$ are randomly drawn from the interval $0 \le t \le 1$ as in the case of §2. The role of $\beta_\tau(t)$ corresponds to that of $\beta(t)$ in the case of §3. Here we have taken into

consideration a family of functions, not merely an individual determined function. We mention our result in

THEOREM 5. *The variance of random integral of the type* (*C*) *defined in the product space of* (*τ*)×(*C*) *with the product probability measure of* (*τ*) *and Wiener measure is given by* (4.02) *so far as the integral gives an unbiassed estimator of the integral of the original function over* $0 \leq t \leq 1$, *that is,* (4.01) *holds true.*

Note: In the right-hand side of (3.12), some of the summands become zero, since

$$\sum_{k=1}^{n} A_k t_k^r = \frac{1}{r+1} \quad (r=0, 1, 2, \ldots, h).$$

The numerical calculations of the values of (3.09) and (3.12) will be given at another occasion.

## References

CAMERON R, H. and MARTIN, W. T.: [1] Transformations of Wiener integrals under translations, *Ann. Math.*, 54 (1944), 386.

KITAGAWA, T.: [1] Sampling from the processes depending upon a continuous parameter, *Mem. Fac. Kyushu Univ., Ser. A*, 5 (1950) (in press)

WIENER, N.: [1] Generalised Harmonic Analysis, *Acta Math.*, 55 (1930), 117—258, esp. 214—234.

PALEY, E. A. C. and WIENER, N.: [1] Fourier Transforms in the Complex Domain, *Amer. Math. Soc. Colloquium*, XIX, (1934), esp, Chap. IX and X.