



---

Soviet-era science, translated into English

# L. A. KHALFIN

1962

SovietRxiv

---

View the original and related papers at <https://sovietrxiv.org/items/ru-196201.05157>

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.

**Abstract**

**Full Text**

L. A. KHALFIN

## A STATISTICAL APPROACH TO METHODS OF APPROXIMATE COMPUTATION. EFFECTIVE QUADRATURE FORMULAS

*(Presented by Academician V. I. Smirnov on 26 I 1962)*

In this work, using as an example the problem of finding effective methods (quadrature formulas) for the approximate computation of definite integrals, the idea of a statistical approach to methods of approximate computation is developed. In contrast to the classical deterministic approach, the statistical approach takes into account from the very beginning that the initial data, for example the values of functions, are known with unavoidable errors\* and therefore constitute realizations (sample values) of random variables (functions). For this reason, problems of approximate computation ( "processing" of the initial data) are considered as problems of mathematical statistics, and the quality of approximate computations is naturally assessed by the corresponding statistical criteria.

1. Let  $f(x; \alpha)$ ,  $\alpha \in A$ , where  $A$  is a nondegenerate interval, be a one-parameter class of functions for which the integral

$$S(\alpha) = \int_0^1 f(x; \alpha) dx \quad (1)$$

has meaning.

Within the framework of the classical approach, in which it is assumed that the values  $f(x_k; \alpha)$  at the points  $x_k \in [0, 1]$  ( $k = 1, 2, \dots, N$ ) are known exactly, as quadrature formulas (quadratures)  $S_{\text{class}}^*$  one uses, let us emphasize, linear\*\* formulas (1):

$$S_{\text{class}}^* = \sum_{k=1}^N a_k f(x_k; \alpha), \quad (2)$$

where  $a_k$  and  $x_k$ , independent of  $\alpha$  ( $S$ ), are chosen by some (optimal) method (Simpson, Chebyshev, Gauss, etc. (1)). The quality of the quadrature  $S_{\text{class}}^*$  is assessed by its closeness to  $S(\alpha)$ , i.e., by the magnitude of the remainder term

$$R_N(a_k, x_k; \alpha) = S_{\text{class}}^* - S(\alpha).$$

2. Now let us take into account, in accordance with the main idea, that  $f(x_k; \alpha)$  are not known exactly, i.e., that we may know only realizations (a sample)

$$\tilde{f}(x_1), \tilde{f}(x_2), \dots, \tilde{f}(x_N)$$

of the random function  $\tilde{f}(x; \alpha)$ :

$$\tilde{f}(x; \alpha) = f(x; \alpha) + n(x), \quad (3)$$

where  $n(x)$  is a random process (errors in the determination of  $f(x; \alpha)$ ). Since the statistical characteristics of  $\tilde{f}(x; \alpha)$ , according to (3), are determined by the statistical characteristics of  $n(x)$  and the deterministic functions  $f(x; \alpha)$ , the problem of determining the integral  $S(\alpha)$  can be formulated and studied as a typical parametric problem of mathematical statistics, namely (2-4), how, from the sample  $\tilde{f}(x_1), \tilde{f}(x_2), \dots, \tilde{f}(x_N)$  of random—

\* The presence of fundamentally irremovable errors (noise) is obvious both in the case when the initial data are determined from experiment and in the general case, since the initial data can be introduced into and “processed” by any real computing device only with a finite number of digits.

\*\* Within the framework of the classical approach, the linearity of  $S_{\text{class}}^*$  with respect to  $f(x_k; \alpha)$  is dictated by the very definition of the integral as the limit of such sums.

of the process  $\tilde{f}(x; \alpha)$  to express oneself with respect to  $\alpha(S)^*$ , i.e. with respect to the parameter of the statistical characteristics of the random process  $\tilde{f}(x; \alpha)$ . All that can be done is to choose some quadrature  $S^*$  (an estimator, in statistical language (2-4)), i.e., generally speaking, an arbitrary function of the sample, independent of  $\alpha(S)$ :

$$S^* = S^*(\tilde{f}(x_1), \tilde{f}(x_2), \dots, \tilde{f}(x_N)). \quad (4)$$

It is natural to choose “good” quadratures. Since  $S^*$  (4) depends on the sample, i.e. on the observed values of random variables, it is meaningless to require deterministic closeness of  $S^*$  to  $S(\alpha)$ ; naturally, one should require “closeness” of  $S^*$  to  $S(\alpha)$  in probabilistic language. Thus, it is desirable, for example, that  $S^*$  be an **unbiased quadrature** for all  $\alpha \in A$ , or, if this is impossible, then for some  $\alpha_i \in A$ , i.e. that

$$MS^* = \overline{S^*(\alpha)} = S(\alpha), \quad \alpha \in A \text{ or } \alpha_i \in A. \quad (5)$$

It is also desirable that  $S^*$  be concentrated about  $S(\alpha)$ , i.e. that  $M(S^* - S(\alpha))^2$  be minimal, if such a nontrivial minimum exists. As the Rao-Cramér inequality <sup>(2,3)</sup> indicates, in the class of so-called **regular quadratures**\*\* (see the definition of a regular estimator in <sup>(2-4)</sup>) this minimum exists:

$$M(S^* - S(\alpha))^2 \geq \frac{(\partial MS^*/\partial S)^2}{M(\partial \log L/\partial S)^2} = \frac{(\partial MS^*/\partial \alpha)^2}{M(\partial \log \hat{L}/\partial \alpha)^2}, \quad (6)$$

where  $L = L(\tilde{f}(x_1), \tilde{f}(x_2), \dots, \tilde{f}(x_N); S(\alpha)) = \tilde{L}(\tilde{f}(x_1), \dots, \tilde{f}(x_N); \alpha)$  is the likelihood function of the sample <sup>(2,3)</sup>.

**Definition.** A regular quadrature  $S_{\text{eff}}^*$  is called **effective in the class of regular quadratures  $S^*$  with fixed value  $(\partial MS^*/\partial \alpha)^2$**  if for  $S_{\text{eff}}^*$  equality holds in (6).

The necessary and sufficient conditions for the existence of effective quadratures are the necessary and sufficient conditions for the existence of effective estimators\*\*\* <sup>(3)</sup>. If effective quadratures  $S_{\text{eff}}^*$  do not exist, then (6) gives an upper estimate (in this case an overestimate) of the maximally possible accuracy of computing the integral  $S(\alpha)$ . In the class of unbiased regular quadratures, in this case, according to <sup>(4-6)</sup>, one can obtain an analogue of the Rao-Cramér inequality <sup>(6)</sup>, the equality sign in which is attained for the so-called best unbiased regular quadratures <sup>(5,6)</sup>.

If effective quadratures  $S_{\text{eff}}^*$  exist, then on the basis of R. Fisher's theorem <sup>(3)</sup> they can be found as the unique solutions of the likelihood equations.

It is obvious that a priori it does not follow from anywhere that linear quadratures\*\*\*\*

$$S^* = \sum_{k=1}^N a_k \tilde{f}(x_k), \quad (7)$$

in particular those dictated by the classical approach (see above, §1), are effective. It turns out that only for very simple classes of functions are there effective ones among linear quadratures. Thus we arrive at a fact completely unexpected in the classical formulation: effective **nonlinear** quadratures for the approximate computation of integrals.

\* In this exposition we restrict ourselves to the case when  $\alpha = \alpha(S)$  is a single-valued function of  $S$ .

\*\* The class of regular quadratures, as follows from <sup>(2-4)</sup>, is quite broad.

\*\*\* Generally speaking, ones that are difficult to verify.

\*\*\*\* In what follows we shall usually omit the upper summation limit  $N$ .

3. Let  $n(x)$  be a normal uncorrelated process:

$$p(n(x_k)) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2\sigma^2}n^2(x_k)\right]; \quad Mn(x_k)n(x_l) = 0 \quad (k \neq l); \quad (8)$$

then (6) becomes

$$M(S^* - S(\alpha))^2 \geq \frac{\sigma^2(\partial MS^*/\partial\alpha)^2}{\sum_k (\partial f(x_k; \alpha)/\partial\alpha)^2} \geq \frac{\sigma^2(\partial MS^*/\partial\alpha)^2}{N(\partial f(x_j; \alpha)/\partial\alpha)_{\max}^2}. \quad (9)$$

Inequality (9) obviously gives an estimate of the quality of quadratures  $S^*$ , replacing, in the classical formulation, the estimate by means of  $R_N$ .

**Theorem 1.** *In the class of regular linear quadratures (7), for an arbitrary class of one-parameter functions  $f(x; \alpha)$ , the following is valid:*

$$M(S_{\text{eff}}^* - S(\alpha))^2 = \frac{\sigma^2 \left( \sum_k a_k \partial f(x_k; \alpha) / \partial \alpha \right)^2}{\sum_n (\partial f(x_n; \alpha) / \partial \alpha)^2} \leq \sigma^2 \sum_k a_k^2, \quad (10)$$

**Theorem 2.** *In order that, in the class of regular linear quadratures (7), unbiased for  $\alpha = \alpha_0$ :*

$$MS^* = \bar{S}^*(\alpha_0) = S(\alpha_0), \quad (11)$$

the quadrature  $S_{\text{eff}}^*$  be effective for  $\alpha = \alpha_0$ , it is necessary and sufficient that:

$$\left[ \sum_k f(x_k; \alpha_0) \frac{\partial f(x_k; \alpha_0)}{\partial \alpha} \right]^2 = \sum_n f^2(x_n; \alpha_0) \sum_m \left( \frac{\partial f(x_m; \alpha_0)}{\partial \alpha} \right)^2. \quad (12)$$

The idea of the proof of the last theorem is that, among the linear quadratures (7) unbiased with respect to  $\alpha = \alpha_0$  (11), the minimum value of

$$M(S^* - S(\alpha_0))^2 :$$

$$M(S^* - S(\alpha_0))^2 = \sum_k a_k^2 \sigma^2, \quad (13)$$

as can be shown, is attained by  $S^*$  with the following  $a_k$ :

$$a_k = \frac{S(\alpha_0)f(x_k; \alpha_0)}{\sum_n f^2(x_n; \alpha_0)}, \quad (14)$$

i.e.,

$$(M(S^* - S(\alpha_0))^2)_{\min} = \frac{\sigma^2 S^2(\alpha_0)}{\sum_k f^2(x_k; \alpha_0)}. \quad (15)$$

At the same time, from (9)

$$M(S_{\text{eff}}^* - S(\alpha_0))^2 = \frac{\sigma^2 S^2(\alpha_0) \left[ \sum_k f(x_k; \alpha_0) \partial f(x_k; \alpha_0) / \partial \alpha \right]^2}{\left[ \sum_n f^2(x_n; \alpha_0) \right]^2 \cdot \sum_m (\partial f(x_m; \alpha_0) / \partial \alpha)^2}. \quad (16)$$

Comparing (15) and (16), we arrive at (12).

**Example 1.** Let  $f(x; \alpha) = \alpha x^m$ , where  $\alpha \in (-\infty, \infty)$ ,  $m \geq 0$ . Then among the linear quadratures (7) there is an unbiased and effective, for all  $\alpha$ ,

$$S_{\text{eff}}^* = \sum_k \frac{1}{m+1} \frac{x_k^m}{\sum_n x_n^{2m}} \tilde{f}(x_k). \quad (17)$$

It is interesting to compare the linear unbiased effective quadrature  $S_{\text{eff}}^*$  (17) with linear quadratures (unbiased) in the classical formulation <sup>(1)</sup>. If by  $S_{\text{Gauss}}^*$  we denote the quadrature in which  $a_k$  (and  $x_k$ ) are chosen according to the Gauss method <sup>(1)</sup>, then for  $N = 3$ ,  $m = 4$ ,  $DS_{\text{Gauss}}^* = 3.41 DS_{\text{eff}}^*$ , and for  $N = 7$ ,  $m = 5$ ,  $DS_{\text{Gauss}}^* = 6.21 DS_{\text{eff}}^*$ .

**Example 2.** Let  $f(x; \alpha) = \exp(-\alpha x)$ ,  $\alpha \geq 0$ . As is not hard to see, among the linear quadratures (7) one cannot choose one that is unbiased for all  $\alpha \in [0, \infty)$ . Among the linear quadratures  $S^*$  (7), unbiased for  $\alpha = \alpha_0$  (11), in this case there can be no efficient ones. Indeed, (14), (15), (16) become

$$a_k = \frac{1}{\alpha_0} (1 - e^{-\alpha_0}) \frac{e^{-\alpha_0 x_k}}{\sum_n e^{-2\alpha_0 x_n}}; \quad (M(S^* - S(\alpha_0))^2)_{\min} = \frac{\sigma^2 (1 - e^{-\alpha_0})^2}{\alpha_0^2 \sum_k e^{-2\alpha_0 x_k}};$$

$$M(S_{\text{eff}}^* - S(\alpha_0))^2 = \frac{\sigma^2 (1 - e^{-\alpha_0})^2 \left[ \sum_k x_k e^{-2\alpha_0 x_k} \right]^2}{\alpha_0^2 \left[ \sum_n e^{-2\alpha_0 x_n} \right]^2 \sum_m x_m^2 e^{-2\alpha_0 x_m}}, \quad (18)$$

and, consequently, efficient quadratures, if they exist, must be nonlinear for this class of functions. This conclusion is valid for a very broad class of functions.

4. As is seen from (9), for the normal uncorrelated process (8), in passing to the limit  $N \rightarrow \infty$ , i.e., as the number of points  $x_k$  becomes denser, the estimate (9) becomes trivial. This is due to the unrealistic assumption that  $n(x_k)$  remain uncorrelated as  $N \rightarrow \infty$ . With the natural accounting for correlation, however, the fundamental inequality (6) remains meaningful. Consider, for example, the following process  $n(x)$ :

$$p(n(x_k)) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2\sigma^2}n^2(x_k)\right],$$

$$Mn(x_k)n(x_l) = \sigma^2 e^{-\gamma|x_k-x_l|} = \sigma^2\beta(|x_k-x_l|) \quad (k \neq l); \quad \Delta x = x_k - x_{k-1}; \quad (19)$$

then (6) becomes

$$M(S^* - S(\alpha))^2 \geq \frac{\sigma^2(\partial MS^*/\partial\alpha)^2}{(\partial f(x_1; \alpha)/\partial\alpha)^2 + \frac{1}{1-\beta^2(\Delta x)} \sum_{k=2}^N [\partial f(x_k; \alpha)/\partial\alpha - \beta(\Delta x) \partial f(x_{k-1}; \alpha)/\partial\alpha]^2}. \quad (20)$$

Passing to the limit  $N \rightarrow \infty$ ,  $\Delta x \rightarrow 0$ , we obtain the following result:

$$M(S^* - S(\alpha))^2 \geq \frac{\sigma^2(\partial MS^*/\partial\alpha)^2}{[\partial f(0; \alpha)/\partial\alpha]^2 + \frac{1}{2\gamma} \int_0^1 [\partial^2 f(x; \alpha)/\partial\alpha \partial x + \gamma \partial f(x; \alpha)/\partial\alpha]^2 dx}, \quad (21)$$

which describes the effect of "saturation" of accuracy in the transition to continuous observation.

5. In the statistical formulation of the problem, in the present work we have regarded the positions of the division points  $x_k \in [0, 1]$  and ( $k = 1, 2, \dots, N$ ) as given. The problem of the optimal placement of the division points in the statistical formulation can be solved and will be presented separately.

In conclusion I express my gratitude to Corresponding Member of the Academy of Sciences of the USSR Yu. V. Linnik, Prof. S. G. Mikhlin, Prof. D. K. Faddeev, V. N. Sudakov, O. V. Shalaevsky, and all persons with whom the ideas of this work were discussed, for interesting discussions and comments.

Leningrad Branch  
of the V. A. Steklov Mathematical Institute  
Academy of Sciences of the USSR

Received  
23 I 1962

## References Cited

1. A. N. Krylov, *Lectures on Approximate Computations*, 1950.
2. U. Grenander, *Stochastic Processes and Statistical Inference*, IL, 1961.
3. H. Cramér, *Mathematical Methods of Statistics*, IL, 1948.
4. B. L. van der Waerden, *Mathematical Statistics*, IL, 1960.
5. A. Bhattacharyya, *Sankhya*, **8**, ch. I (1946); **3**, ch. II, III (1947), **4**, ch. IV (1948).
6. L. N. Bol'shev, *Theory of Probability and Its Applications*, **6**, 319 (1961).

*Note: Figure translations are in progress. See original paper for figures.*

*Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.*