The minimum of the mean-square deviation of approximations of functions of several independent variables by sums of functions of fewer variables in a multidimensional parallelepiped is investigated. The approximating function yielding the minimum mean-square deviation is obtained and this minimum deviation is calculated.

Several articles [1–3] have been devoted to the study of the existence and uniqueness of best-possible approximations of functions of several variables by sums of functions of fewer variables. Here we consider sums of functions of fewer variables yielding the best-possible approximation in a multidimensional parallelepiped $D$ and calculate the corresponding minimum mean-square deviation.

Let the function
$$\chi (x_1, x_2, \ldots, x_n) = \chi_1(x_1)\chi_2(x_2)\ldots\chi_n(x_n),$$
be given in the $n$-dimensional parallelepiped $D = \{a_i \leq x_i \leq b_i; \; i = 1, 2, \ldots, n\}$, where $\chi_i(x_i)$ is non-decreasing and $\int_{a_i}^{b_i} \chi_i(x_i) \, dx_i = 1$. Consider functions $\varphi(x_1, x_2, \ldots, x_n)$, whose squares are $\chi$-integrable in $D$.

We use the following notation:

- $\sigma_n = \{x_1, x_2, \ldots, x_n\}$ is a set of $n$ independent variables;
- $\sigma^1_q$ (or $\sigma^i_q$, $\sigma_q$, $\sigma$, or $\tau$) is a subset of $\sigma_n$ (the superscript indicates the number of the subset, the subscript indicates the number of variables in the subset);
- $\overline{\sigma^1}$ is the set of variables in $\sigma_n$ but not in $\sigma^1$;
- $\Pi^p_q = \sigma^1_q \cap \sigma^2_q \cap \ldots \cap \sigma^p_q$ is the intersection of the subsets $\sigma^i_q$ (the subscript indicates the number of subsets in the intersection, the superscript indicates the enumeration of these subsets),

\[
\chi(\sigma^i_q) = \chi_i(x_i)\chi_{i+1}(x_{i+1})\ldots\chi_n(x_n);
\quad \sigma^i_q = \{x_{i+1}, x_{i+2}, \ldots, x_n\};
\]

\[
\varphi(\sigma_q) = \varphi(x_1, x_2, \ldots, x_n), \quad f(\sigma_q) = f(x_1, x_2, \ldots, x_n),
\]

is the Lebesgue-Stieltjes integral of $f(\sigma_p)$ over the section $D_{\sigma_q} = \{a_i \leq x_i \leq b_i; \; x_i \in \sigma_q\}$ of $D$ (a q-dimensional hyperplane); this type of integral will be called a "mean" of the function $f(\sigma_p)$ over the set of variables $\sigma_q$ and we write:

\[
\int_{D_{\sigma_q}} f(\sigma_p) \, d\chi(\sigma_q) = \int_{\sigma_q} f(\sigma_p) \, d\chi(\sigma_q) = f(\sigma_p \cap \sigma_q).
\]

We use the following obvious properties of averaging operations:

\[1. \; f(\sigma) = \int_{\sigma} f(\sigma);\]

*a_i and b_i can be infinite.
We call a function \( \varphi(\sigma_n) \) reducible on its set of variables \( \sigma_n \), if there is a finite sum of functions of fewer variables such that \( \varphi(\sigma_n) \Rightarrow \sum f(\sigma') \), \( \sigma' \subset \sigma_n \), almost everywhere in \( D \) with respect to the measure \( \chi(\sigma_n) \); otherwise we call \( \varphi(\sigma_n) \) irreducible.

If \( \sigma_q \subset \sigma_n \), a function \( \psi(\sigma_q) \) is called irreducible if it is irreducible on the set of variables \( \sigma_q \); the function \( \psi(\sigma_n) \equiv \varphi(\sigma_q) \) is always reducible on \( \sigma_n \).

Constant quantities \( \varphi(\sigma_q) \) used in our reasoning will be assumed to be irreducible functions on the empty set of variables.

I. The sum of an irreducible function \( \varphi(\sigma_n) \) and any \( \psi(\sigma) \) \( (\sigma \subset \sigma_n) \) is irreducible.

II. Any irreducible function can be expressed as a sum of irreducible functions of sets of variables, none of which is a subset of another.

We call such a representation normal.

Proof. Let \( \varphi(\sigma_n) \) be reducible. Consider the terms involving \( n-1 \) variables (if such terms exist). Combine terms depending on the same set of variables. If there are reducible terms which are functions of \( n-1 \) variables, express them in reduced form. In this way we obtain an expression for \( \varphi(\sigma_n) \) in which all terms depending on \( n-1 \) variables are irreducible and are functions of sets of variables none of which is included in another set. We separate out these irreducible terms. The remaining terms depending on sets of variables which are subsets of the sets of variables in the irreducible functions already separated out are combined with these irreducible functions. We thus obtain

\[
\varphi(\sigma_n) \Rightarrow \sum_j \psi_j(\tau^j) + \sum f(\sigma^q) ; \tau^j, \sigma^q \subset \sigma_n;
\]

where the \( \psi_j(\tau^j) \) are irreducible, \( \tau^j \neq \tau^{j'} \) for \( j_1 \neq j_2, q_1 \equiv n-2, \) and \( \sigma_{q_1}^{q_1} \neq \tau^j \) for any \( j_1 \) and \( j_2 \).

After repeating this procedure not more than \( n \) times we have

\[
\varphi(\sigma_n) \Rightarrow \sum \psi_i(\tau^i), \quad \tau^i \subset \sigma_n,
\]

where \( \psi_i(\tau^i) \) is irreducible, \( \tau^i \neq \tau^{i'} \) for \( j_1 \neq j_2 \), as was required.

III. The group of sets of variables in a normal representation is unique.

Proof. Let \( \varphi(\sigma_n) \) have two normal representations:

\[
\varphi(\sigma_n) \Rightarrow \sum f(\tau^i), \quad \tau^i \subset \sigma_n;
\]

\[
\varphi(\sigma_n) \Rightarrow \sum \psi_i(\omega^i), \quad \omega^i \subset \sigma_n;
\]

Then

\[
\chi(\sigma_n) \Rightarrow \sum f(\tau^i) - \sum \psi_i(\omega^i) \Rightarrow \sum \Phi_i(\omega^i) \Rightarrow 0,
\]

where

\[
\Phi_i(\omega^i) \Rightarrow \begin{cases}
  f_i(\omega^i) - \sum \psi_k(\omega^i), & \text{if} \quad \sigma^i \equiv \omega^i, \\
  \sum f_k(\omega^i) - \psi_j(\sigma^i), & \text{if} \quad \sigma^i \equiv \tau^k, \\
  f_i(\sigma^i), & \text{if} \quad \sigma^i \equiv \omega^i, \sigma^i \not\supseteq \omega^i, \\
  -\psi_j(\sigma^i), & \text{if} \quad \sigma^i \not\supseteq \tau^i, \sigma^i \not\supseteq \tau^i,
\end{cases}
\]

and \( \sigma^1, \sigma^2, \ldots, \sigma^\nu, \ldots \) are sets none of which is a subset of another.

* A dot over a symbol (as in \( \Rightarrow \) and \( \Rightarrow \)) means "almost everywhere with respect to the measure \( \chi."