 DOI 10.25699/SSSB.2021.39.5.006

05.11.15

АЛГОРИТМ ОБРАБОТКИ РЕЗУЛЬТАТОВ МНОГОКРАТНЫХ ИЗМЕРЕНИЙ СОСТАВА И СВОЙСТВ НЕФТЕПРОДУКТОВ

К.В. Шаталов

ФАУ 25 Государственный научно-исследовательский институт химмотологии Министерства обороны Российской Федерации, г. Москва

Аннотация: Разработаны новые робастные алгоритмы обработки результатов многократных измерений состава и свойств нефтепродуктов, учитывающие тот факт, что эмпирическая функция распределения результатов измерений состава и свойств нефтепродуктов представляет собой смесь двух нормальных распределений с разными значениями параметров положения и масштаба. В случае измерений состава и свойств нефтепродуктов в качестве робастных оценок параметра положения и параметра масштаба выборки предложено использовать *М*-оценки с предварительным масштабированием на основе модифицированной функции Хампеля. Для нахождения *М*-оценки предложены два итеративных способа вычисления на основе средневзвешенного метода наименьших квадратов, отличающиеся процедурами расчета начальных оценок параметров положения и масштаба выборки. При числе результатов в выборке более двадцати в качестве начальных значений параметров положения и масштаба целесообразно использовать α-урезанное среднее и α-урезанное стандартное отклонение с долей усечения 0,05. При числе результатов в выборке менее двадцати в качестве начальных значений параметра положения и параметра масштаба обоснованно использование робастных оценок, не требующих удаления части данных. В качестве начальной оценки параметра положения предложено использовать оценку Ходжеса – Лемана; в качестве параметра масштаба – медианы абсолютных разностей. Предложенные робастные алгоритмы могут быть использованы при обработке результатов эксперимента по определению показателей прецизионности, правильности и точности методик измерений состава и свойств нефтепродуктов, итогов межлабораторных сравнительных испытаний нефтепродуктов, расчете аттестованного значения стандартных образцов состава и свойств нефтепродуктов, а также в других случаях многократных наблюдений.

Ключевые слова: нефтепродукты, многократные измерения, алгоритмы обработки, робастные оценки параметров положения и масштаба, М-оценка на основе функции Хампеля.

ВВЕДЕНИЕ

При аттестации методик количественного химического анализа в целом, и методик измерений состава и свойств нефтепродуктов в частности, для установления значений показателей прецизионности, правильности и точности используются алгоритмы обработки многократных наблюдений, изложенные в РМГ 61-2010 «Государственная система обеспечения единства измерений. Показатели точности, правильности, прецизионности методик количественного химического анализа. Методы оценки» и ГОСТ 33701-2015 «Определение и применение показателей точности методов испытаний нефтепродуктов». Указанные алгоритмы основаны на допущении о том, что эмпирическая выборка подчиняется нормальному закону распределения случайной величины. Принято считать, что ошибки измерений, характеризующие «отлаженную» измерительную систему, могут хорошо описываться нормальным законом [1].

Важность вопроса о соответствии выборки многократных экспериментальных измерений тому или иному закону распределения случайной величины обусловлена различиями в процедурах оценки основных параметров выборки – параметров положения и масштаба. Известно, что применение алгоритмов расчета в основу, которых положено предположение о нормальном законе распределения случайной величины, к экспериментальным данным, имеющий другой закон распределения вероятностей приводит к значительным искажениям итоговых выводов [2, 3].

При изучении характера распределения результатов и погрешностей измерения состава и свойств нефтепродуктов было установлено, что фактическое распределение погрешностей результатов измерений состава и свойств нефтепродуктов не соответствует нормальному закону распределения случайной величины. Фактический график плотности распределения погрешностей измерения представляет собой одномодальную асимметричную кривую с положительным эксцессом и тяжелыми хвостами. Основным отличием фактического распределения погрешностей измерений состава и свойств нефтепродуктов от нормального распределения является более плотная концентрация результатов в интервале $[-1\sigma; +1\sigma]$. Другим отличием эмпирического распределения является наличие тяжелых хвостов – результатов, имеющих отклонение более 4σ [4].

Поэтому целью исследований была разработка нового алгоритма обработки результатов многократных измерений состава и свойств нефтепродуктов, учитывающего факт отклонения фактического распределения величин, характеризующих состав и свойства нефтепродуктов от нормального закона распределения.

ОСНОВНАЯ ЧАСТЬ

В основу новых алгоритмов обработки экспериментальных данных была положена новая обобщенная вероятностная модель, описывающая результаты многократных измерений состава и свойств нефтепродуктов, как совокупность двух нормальных законов. При хорошо налаженной работе лаборатории (статистически управляемом состоянии процесса испытаний) получаемые результаты с вероятностью $(1-\varepsilon)$ соответствуют «основному» нормальному закону $F(x)=N(\mu_0;\sigma_0^2)$ с дисперсией σ_0^2, не превышающей установленные требования. При отклонениях в работе лаборатории (статистически неуправляемом состоянии) с вероятностью ε получаемые результаты соответствуют «засоряющему» нормальному закону $F_1(x)=N(\mu_1;\sigma_1^2)$ с дисперсией σ_1^2 значительно превышающей установленные требования [5]. Тогда общее распределение результатов измерений состава и свойств нефтепродуктов имеет вид:

$$F(x)=(1-\varepsilon)N(\mu_0;\sigma_0^2)+\varepsilon N(\mu_1;\sigma_1^2), \mu_0\neq\mu_1, \sigma_0^2>\sigma_1^2 \quad (1)$$

где $N(\mu;\sigma)=\frac{1}{\sigma\sqrt{2\pi}}\cdot exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$.

Отличия идеальной модели, основанной на теоретических предположениях о соответствии результатов измерения нормальному закону, и эмпирической модели (1), обоснованной статистически и экспериментально, рассмотрим на следующих примерах.

Рассмотрим, как изменяется качество выборочного среднего арифметического в условиях статистической модели (1) при увеличении доли «засорения» ε от 0 до 0,3 при $\sigma_1=3$ и $\sigma_1=5$. За критерий качества возьмем асимптотическую абсолютную эффективность среднего арифметического [6]:

$$\text{АЭ}_F(\bar{X})=\{\sigma_F^2(\bar{X})I(f)\}^{-1},$$

$$I(f)=\int\left(\frac{f'(x)}{f(x)}\right)^2dF(x),$$

$$\sigma_F^2(\bar{X})=\int_{-\infty}^{+\infty}x^2dF(x)=1+\varepsilon(\sigma_1^2-1)$$

где $I(f)$ – информация Фишера относительно среднего арифметического; $\sigma_F^2(\bar{X})$ – асимптотическая дисперсия среднего арифметического.

Установлено, что среднее арифметическое как оценка параметра положения μ, обладающая свойством оптимальности при нормальном распределении ($\varepsilon=0$), теряет это свойство даже при небольших отклонениях от нормального распределения. При засорении в 1 % ($\varepsilon=0{,}01$) среднее арифметическое утрачивает от 5 до 17 % эффективности, а при засорении в 5 % ($\varepsilon=0{,}05$) эффективность снижается на 20 - 50 % [6].

В работе [7] проведено сравнение двух оценок масштабного параметра стандартного отклонения $\hat{S}_1$ и среднего абсолютных отклонений $\hat{S}_2$ с использованием асимптотической относительной эффективности

$$\text{АОЭ}_F(\hat{S}_2:\hat{S}_1)=\frac{\sigma_F^2(\hat{S}_1)/S_1^2(F)}{\sigma_F^2(\hat{S}_2)/S_2^2(F)}.$$

Установлено, что даже при незначительных отклонениях от нормального распределения свойство оптимальности стандартного отклонения $\hat{S}_1$ быстро исчезает. Достаточно два «плохих» наблюдения на тысячу, чтобы преимущество $\hat{S}_1$ перед $\hat{S}_2$ пропало, а при засорении в 10 % среднее абсолютных отклонений $\hat{S}_2$ в два раза эффективнее, чем стандартного отклонение $\hat{S}_1$ [7].

Следовательно, принятие вероятностной модели (1) требует разработки новых алгоритмов обработки многократных измерений состава и свойств нефтепродуктов, учитывающих наличие в экспериментальных выборках «засоряющих» результатов.

Вероятностная модель (1) соответствует локальной модели робастного анализа, которая рассматривает задачи оценивания эффективных[1] оценок параметра положения μ_0 и параметра масштаба σ_0, по имеющейся выборке данных, исходя из множества распределений, близких к нормальному [6].

Изучение разных классов робастных оценок [6] показало, что в случае измерений состава и свойств нефтепродуктов в качестве оценок параметра положения μ_0 и параметра масштаба σ_0 выборки целесообразно использовать *M*-оценки, с предварительным масштабированием, которые находятся путем решения уравнения:

$$\sum_{i=1}^{n}\psi\left(\frac{X_i-T_n}{S_n}\right)=\sum_{i=1}^{n}\psi(u_i)=0, \qquad (2)$$

где ψ - заданная функция, принадлежащая классу функций C_M, определяет конкретный вид *M*-оценки, а класс функций C_M содержит дифференцируемые, нечетные, возрастающие функции $\psi(x)$, для которых ψ' имеет компактный носитель, ограниченную вариацию и $\int\psi'(x)dF(x)\neq 0$; S_n, T_n – робастные оценки параметров масштаба и положения, построенные по исходной выборке.

В частности использовали *M*-оценку Хампеля, для которой функция ψ имеет вид:

[1] имеющих минимальную дисперсию, как для эмпирического, так и для теоретического распределения

$$\psi_{Ham}(x)=\begin{cases}|x|, & 0\le |x|<a\\ a, & a\le |x|<b\\ a\left(\dfrac{c-|x|}{c-b}\right), & b\le |x|<c\\ 0, & |x|\ge c\end{cases} \qquad (3)$$

где a, b, c – регулируемые параметры [6]. На основании изучения гистограмм распределения результатов испытаний нефтепродуктов были приняты следующие значения $a = 1{,}0$; $b = 2{,}0$; $c = 3{,}0$.

Известно, что численное решение уравнения (3) возможно с помощью различных итерационных схем [8]. Суть их заключается в следующем. По исходной выборке вычисляют робастные оценки параметра положения и параметра масштаба, которые необходимы для нахождения значений функции ψ на первом этапе итеративных вычислений. Затем с помощью того или иного способа численного решения уравнения (3) находят оценку параметра положения $T_{it=1}$. Вычисления повторяют до тех пор пока значения T_{it} не начнут сходиться. В работе [9] показано преимущество итеративной схемы вычислений на основе средневзвешенного метода наименьших квадратов. В этом случае M-оценка с предварительным нормированием вида (3) вычисляется по формулам:

$$T=\frac{\sum_{i=1}^{n}W_i\cdot x_i}{\sum_{i=1}^{n}W_i},$$

$$W_i=\frac{\psi[(x_i-T_{it})/S_{it}]}{(x_i-T_{it})/S_{it}},$$

где W_i – вес i-го значения исходной выборки; T_{it}, S_{it} – оценки параметра положения и параметра масштаба полученные в ходе it-ой итерации.

Предложены два варианта реализации этого алгоритма А и Б, отличающиеся процедурами вычисления начальных оценок параметров положения и масштаба выборки.

При числе результатов в выборке более двадцати **(Алгоритм А)** в качестве начальных значений параметров положения и масштаба целесообразно использовать α-усеченные оценки: α-урезанное среднее и α-урезанное стандартное отклонение. При идентификации закона распределений результатов измерений состава и свойств нефтепродуктов было установлено, что от 92 % до 96 % получаемых результатов находятся в интервале $[-2\sigma; +2\sigma]$. Поэтому при вычислении α-усеченных оценок принимаем значение параметра усечения *α=0,05*, что соответствует получению выборки, урезанной в целом на 10 %. Оставшиеся 90 % результатов образуют одномодальное симметричное ядро, для которого принимаем, что эмпирическая функция распределения приближенно соответствует нормальному распределению.

Исходную выборку результатов $x_1, \dots, x_n$ упорядочивают по возрастанию в ряд

$$x_{(1)}\le x_{(2)}\le \dots \le x_{(i)}\le \dots \le x_{(n)}, \qquad (4)$$

где n – количество результатов в исходной выборке.

Для упорядоченной выборки (4) рассчитывают α-усеченные оценки:

- α-урезанное среднее:

$$\bar{X}_\alpha=\frac{1}{n-2k}\sum_{i=k+1}^{n-k}x_{(i)}, \qquad (5)$$

где $\alpha = 0{,}05$ – доля удаляемых значений; $k = [\alpha n]$ – число удаляемых значений, округленное до целой части числа;

- α-урезанное стандартное отклонение:

$$\hat{S}_\alpha=\sqrt{\frac{1}{n-2k}\sum_{i=k+1}^{n-k}\left(x_{(i)}-\bar{X}_\alpha\right)^2}, \qquad (6)$$

Для каждой точки x_i исходной выборки рассчитывают нормированное значение u_i:

$$u_i=\left|\frac{x_i-X^*}{S^*}\right| \qquad (7)$$

В первой итерации при вычислении нормированных значений u_i принимают:

$X^*=\bar{X}_\alpha$, $S^*=\hat{S}_\alpha$.

Вычисляют вес W_i нормированного значения u_i:

$$W_i=\begin{cases}0, & \text{если } |u_i|>3\\ \dfrac{3-u_i}{u_i}, & \text{если } 2<|u_i|\le 3\\ \dfrac{1}{u_i}, & \text{если } 1<|u_i|\le 2\\ 1, & \text{если } |u_i|\le 1\end{cases} \qquad (8)$$

Рассчитывают робастную оценку параметра положения X^* как средневзвешенное всех значений $x_1, \dots, x_i, \dots, x_n$:

$$X^*=\frac{\sum_{i=1}^{n}W_i\cdot x_i}{\sum_{i=1}^{n}W_i} \qquad (9)$$

Вычисления по формулам (5)–(9) составляют первую итерацию. При переходе ко второй итерации рассчитывают новое значение α-урезанного стандартного отклонения, заменяя α-урезанное среднее $\bar{X}_\alpha$ на робастную оценку X^*:

$$S^*=\sqrt{\frac{1}{n-2k}\sum_{i=k+1}^{n-k}\left(x_{(i)}-X^*\right)^2} \qquad (10)$$

Во второй итерации повторяют вычисления по формулам (7)–(9) с использованием ранее найденного значения X^*.

Итерационные вычисления по формулам (7)–(10) повторяют до тех пор, пока робастные значения X^* в двух последних итерациях не начнут сходиться.

Сходимость считают достаточной, если выполняется следующее условие:

$$|X_{it-1}^* - X_{it}^*| < \frac{0{,}1S_{it}^*}{\sqrt{n}} \quad (11)$$

где X_{it-1}^* - робастное значение X^* в *it-1* итерации; X_{it}^* - робастное значение X^* в *it*-ой итерации; S_{it}^* - робастное значение S^* в *it*-ой итерации.

При числе результатов в выборке менее двадцати **(Алгоритм Б)** в качестве начальных значений параметра положения и параметра масштаба целесообразно использовать робастные оценки, не требующие удаления части данных.

В качестве начальной оценки параметра положения применяем оценку Ходжеса – Лемана HL, которая определяется в виде медианы средних Уолша, с использованием всех n^2 пар средних Уолша MW_{ij}

$$HL = med\left\{\left(MW_{ij} = \frac{X_i + X_j}{2}\right), 1 \le i < j \le n\right\}, \quad (12)$$

где $i = 1, \dots, n$ – номер результата в исходной выборке; $j = i, i+1, i+2, \dots, n$ – номер результата в выборке; n – количество результатов в исходной выборке.

Для этого рассчитывают средние Уолша MW_{ij}, которые образуют новую выборку:

$$MW_{11}, \dots, MW_{1n}, MW_{21}, \dots, MW_{2n}, \dots, MW_{n1}, \dots, MW_{nn} \quad (13)$$

где

$$MW_{11} = \frac{X_{i=1} + X_{j=1}}{2}, \dots, MW_{1n} = \frac{X_{i=1} + X_{j=n}}{2};$$
$$MW_{21} = \frac{X_{i=2} + X_{j=1}}{2}, \dots, MW_{2n} = \frac{X_{i=2} + X_{j=n}}{2};$$
$$\dots\dots$$
$$MW_{n1} = \frac{X_{i=n} + X_{j=1}}{2}, \dots, MW_{nn} = \frac{X_{i=n} + X_{j=n}}{2}.$$

Выборку (13) упорядочивают по возрастанию в ряд

$$MW_{(1)} \le MW_{(2)} \le MW_{(3)} \le \dots \le MW_{(n^2)}. \quad (14)$$

Оценку Ходжеса – Лемана HL вычисляют как медиану выборки (14):

$$HL = \begin{cases} \dfrac{MW_{\left(\frac{n^2}{2}\right)} + MW_{\left(\frac{n^2}{2}+1\right)}}{2} - \text{ для четных } n \\ MW_{\left(\frac{n^2+1}{2}\right)} \quad - \text{ для нечетных } n. \end{cases} \quad (15)$$

В качестве начального значения параметра масштаба используют оценку медианы абсолютных разностей $\hat{S}_4$

$$\hat{S}_4 = med\{|X_i - X_j|, 1 \le i < j \le n\}, \quad (16)$$

Для этого рассчитывают абсолютные разности AD_{ij}:

$$AD_{11} = |X_{i=1} - X_{j=1}|, \dots, AD_{1n} = |X_{i=1} - X_{j=n}|$$
$$AD_{21} = |X_{i=2} - X_{j=1}|, \dots, AD_{2n} = |X_{i=2} - X_{j=n}|$$
$$\dots\dots$$
$$AD_{n1} = |X_{i=n} - X_{j=1}|, \dots, AD_{nn} = |X_{i=n} - X_{j=n}|$$

В выборке $AD_{11}, \dots, AD_{1n}, AD_{21}, \dots, AD_{2n}, \dots, AD_{n1}, \dots, AD_{nn}$ отсеивают нулевые абсолютные разности $AD_{ij} = 0$ (пусть их число равно k_1). Ненулевые абсолютные разности AD_{ij} упорядочивают по возрастанию в ряд

$$AD_{(1)} \le AD_{(2)} \le AD_{(3)} \le \dots \le AD_{(n^2-k_1)}. \quad (17)$$

Оценку медианы абсолютных разностей $\hat{S}_4$ вычисляют как медиану выборки (17)

$$\hat{S}_4 = \begin{cases} \dfrac{AD_{\left(\frac{n^2-k_1}{2}\right)} + AD_{\left(\frac{n^2-k_1}{2}+1\right)}}{2} - \text{ для четных } n \\ AD_{\left(\frac{n^2-k_1+1}{2}\right)} \quad - \text{ для нечетных } n \end{cases} \quad (18)$$

3 Для каждой точки x_i исходной выборки рассчитывают нормированное значение u_i по формуле (7). В первой итерации при вычислении нормированных значений u_i принимают:

$X^* = HL, S^* = \hat{S}_4$.

Вычисляют вес W_i нормированного значения u_i по формуле (8).

Рассчитывают робастную оценку параметра положения X^* как средневзвешенное всех значений $x_1, \dots, x_i, \dots, x_n$ по формуле (9).

Вычисления по формулам (12)–(18), (7)–(9) составляют первую итерацию.

Во второй итерации повторяют вычисления по формулам (7)–(9) с использованием ранее найденного значения X^*.

Итерационные вычисления по формулам (7)–(9) повторяют до тех пор, пока значения X^* в двух последних итерациях не начнут сходиться по условию (11).

Представленные робастные алгоритмы были апробированы в ФАУ «25 ГосНИИ химмотологии Минобороны России» при обработке результатов экспериментов по установлению показателей прецизионности, правильности и точности методик измерения показателей, характеризующих эксплуатационные свойства нефтепродуктов, а также при определении аттестованных значений стандартных образцов предприятия.

Практическая реализация разработанных робастных алгоритмов обработки результатов многократных испытаний нефтепродуктов требует достаточно трудоемких вычислений. «Ручной» режим вычислений может быть реализован с использованием формул (2) – (18) в табличном процессоре Excel, входящего в интегральный программный комплекс Microsoft Office. Опыт реализации Алгоритмов А и Б показал, что в «ручном» режиме встречаются ошибки из-за неправильного (некорректного) задания пользователями формул в отдельных ячейках электронных таблиц. Выявление и устранение таких ошибок представляет собой длительную трудоемкую процедуру, так как большинство ошибок имеет скрытый характер и требуется проверка всей цепочки вычислений. Для упрощения использования Алгоритмов А и Б была разработана компьютерная программа «Выборка-2021» (свидетельство о государственной регистрации программы для ЭВМ № 2021660250 от 23.06.2021 г.), позволяющая автоматизировать расчеты. Заинтересованным пользователям программа может быть представлена при запросе по адресу 1499090@mail.ru.

ЗАКЛЮЧЕНИЕ

Разработаны новые робастные алгоритмы обработки результатов многократных измерений состава и свойств нефтепродуктов, учитывающие тот факт, что эмпирическая функция распределения результатов измерений состава и свойств нефтепродуктов представляет собой смесь двух нормальных распределений с разными значениями параметров положения и масштаба.

Предложенные робастные алгоритмы могут быть использованы при обработке результатов эксперимента по определению показателей прецизионности, правильности и точности методик измерений состава и свойств нефтепродуктов, итогов межлабораторных сравнительных испытаний нефтепродуктов, расчете аттестованного значения стандартных образцов состава и свойств нефтепродуктов, а также в других случаях многократных наблюдений.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

21. Налимов, В.В. Применение математической статистики при анализе вещества [Текст]/ В.В. Налимов – М.: Физматиздат, 1960. – 430 с.
22. Лемешко, Б.Ю. Критерии проверки отклонения распределения от нормального закона. Руководство по применению [Текст]/ Б.Ю. Лемешко. – Новосибирск: Издательство НГТУ, 2014. – 192 с.
23. Новицкий, П.В. Оценка погрешностей результатов измерений [Текст]/ П.В. Новицкий, И.А. Зограф. – Л.: Энергоатомиздат. Ленинградское отделение, 1985. – 248 с.
24. Шаталов К.В., Черепанова А.Д. Проверка гипотезы о соответствии погрешностей результатов анализа нефтепродуктов нормальному закону распределения случайной величины [Текст]// Измерительная техника. – 2019. – № 10. – С. 52-60.
25. Шаталов, К.В. Закон распределения результатов измерений состава и свойств нефтепродуктов [Электронный ресурс] / К.В. Шаталов, А.Д. Черепанова // Южно-Сибирский научный вестник – 2021. - №4 – С.16-24. – Режим доступа: hptt://s-sibsb.ru.
26. Шуленин, В. П. Математическая статистика. Ч. 3. Робастная статистика: учебник. [Текст] / В.П. Шуленин – Томск: Издательство НТЛ, 2012. – 520 с.
27. Tukey, J.W. Bias and confidence in not-quite large samples [Текст] / J.W. Tukey // Annals of Mathematical Statistics. – 1958. – V. 29. - №2 – P. 614-623.
28. Лонер, Р.Л. Устойчивые статистические методы оценки данных [Текст] / Р.Л. Лонер, Г.Н. Уилкинсон; пер. с англ. – М.: Машиностроение, 1984. – 229 с.
29. Holland, P.W. Robust regression using interactively reweighted least-squares [Текст] / P.W. Holland, R.E. Welsch // Communications in Statistics – Theory and Methods. - 1977. - Vol. 6 (9). - P. 813−827.

Шаталов Константин Васильевич – к.т.н., доцент, начальник отдела квалификационных испытаний топлив и масел ФАУ «25 ГосНИИ химмотологии Минобороны России», тел. (499)149-90-90, e-mail: 1499090@mail.ru.

AN ALGORITHM OF TREATMENT OF THE RESULTS OF MULTIPLE MEASUREMENTS OF COMPOSITION AND PROPERTIES OF PETROLEUM PRODUCTS

K.V. Shatalov
Federal Autonomous Enterprise "The 25-th State Research Institute of Chemmotology, Ministry of Defence of Russian Federation», Moscow

Summary: New robust algorithms of treatment of the results of multiple measurements of composition and properties of petroleum products were developed in respect that empirical distribution function of the results of measurements of composition and properties of petroleum products are the mixture of two normal distributions with different values of position and scale parameters. In case of measurements of composition and properties of petroleum products it has been proposed to use *M*-estimator with pre-scaling based on modified Hampel function as robust estimators of position and scale parameters. To calculation *M*-estimator two iterative methods based on weighted average method of least squares were suggested which differs by procedures of initial estimators of position and scale parameters of sample. In case of more than twenty results in sample, it is expedient to apply α-truncated mean and α-truncated standard deviation with 0,05 truncation share as initial values of position and scale parameters. In case of less than twenty results in sample, it is reasonable to apply robust estimators as initial values of position and scale parameters, which don't require removal of some part of the data. It was proposed to use Hodges-Lehmann estimator as an initial value of position parameter and median of absolute differences as a scale parameter. The proposed robust algorithms can be used in treatment of experiment results on determination of indexes of precision, trueness and accuracy of the methods of measurement of composition and properties of petroleum products; results of interlaboratory comparison tests of petroleum products; calculation of certified value of standard samples of composition and properties of petroleum products and in other cases of multiple observations.

Key words: petroleum products, multiple measurements, algorithms of treatment, robust estimators of position and scale parameters, *M*-estimator based on Hampel function.

REFERENCES

21. Nalimov, V.V. The Application of Mathematical Statistics to Chemical Analysis [Text]/ V.V. Nalimov – Moscow: Physmathizdat, 1960. – 430 p. (in Russian)

22. Lemeshko, B.Yu. Inspection Criteria of Distribution Deviation from Normal Probability Law. Usage Guidelines [Text]/ B.Yu. Lemeshko. – Novosibirsk: NGTU Publishing House, 2014. – 192 p. (in Russian)

23. Novitsky, P.V. Estimation of Errors in Measurement Results [Text]/ P.V. Novitsky, I.A. Zograf. – Leningrad: Energoatomizdat, Leningrad department, 1985. – 248 p. (in Russian)

24. Shatalov, K. V.Verification of the Hypothesis of Compliance of the Errors of the Results of Analyses of Oil Products to the Normal Distribution Law of a Random Variable [Text] / K. V. Shatalov, A. D. Cherepanova // Measurement Techniques. – 2020. - Vol. 62. – P. 905–914.

25. Shatalov, K. V. A distribution law of the results of measurement of composition and properties of petroleum products [Electronic resourse] / K. V. Shatalov, A. D. Cherepanova // South-Siberian scientific bulletin – 2021. - №4 – С.16-24. – The access mode: hpttː//s-sibsb.ru.

26. Shulenin, V.P. Mathematical statistics. Part 3. Robust statistics: coursebook. [Text] / V.P. Shulenin – Tomsk: NTL Publishing House, 2012. – 520 p. (in Russian)

27. Tukey, J.W. Bias and confidence in not-quite large samples [Text] / J.W. Tukey // Annals of Mathematical Statistics. – 1958. – V. 29. - №2 – P. 614-623.

28. Launer, R.L. Robustness in statistics [Text] / R.L. Launer, G.N. Wilkinson. – Academic press, 1979. – 296 p.

29. Holland, P.W. Robust regression using interactively reweighted least-squares [Text] / P.W. Holland, R.E. Welsch // Communications in Statistics – Theory and Methods. 1977. Vol. 6 (9). P. 813−827.

Shatalov Konstantin Vasilievich – Candidate of Engineering Sciences, Associate Professor, Head of the Department of Qualification Assessment of Fuel and Oils Federal Autonomous Enterprise «The 25-th State Research Institute of Himmotology, Ministry of Defence of Russian Federation», tel. (499)149-90-90, e-mail: 1499090@mail.ru.