

ЗАКОН РАСПРЕДЕЛЕНИЯ РЕЗУЛЬТАТОВ ИЗМЕРЕНИЙ СОСТАВА И СВОЙСТВ НЕФТЕПРОДУКТОВ

К.В. Шаталов, А.Д. Черепанова

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

Идентификацию закона распределения результатов измерений состава и свойств нефтепродуктов проводили путем проверки сложной гипотезы с использованием пяти критериев согласия: χ^2 -Пирсона, Колмогорова, Смирнова, ω^2 Крамера-Мизеса-Смирнова; Ω^2 Андерсона-Дарлингга. В качестве возможных функций распределения вероятностей рассматривали 12 симметричных одномодальных законов распределения, а также 66 смесей этих же законов распределения. Целью идентификации являлось нахождение универсального закона распределения (смеси законов распределений) справедливого для всех рассматриваемых величин. Проверка сложной гипотезы о соответствии какому-либо симметричному одномодальному закону распределения показала, что не существует универсального закона распределения справедливого для всех методик измерений состава и свойств нефтепродуктов, наиболее часто не отвергалась гипотеза о соответствии данных обобщенному логистическому распределению, распределению Лапласа и двустороннему экспоненциальному распределению. Проверка сложной гипотезы о соответствии какой-либо смеси симметричных одномодальных законов распределения показала, что эмпирическая функция распределения результатов измерений состава и свойств нефтепродуктов может быть представлена в виде смеси двух нормальных распределений с разными значениями параметров положения и масштаба. При этом для одной и той же выборки значения достигаемого уровня значимости гипотезы о соответствии смеси законов распределений в несколько раз выше среднего значения достигаемого уровня значимости гипотезы о соответствии одному закону распределения. На основе проведенного исследования обоснована вероятностная модель процесса испытаний нефтепродуктов, в рамках которой результат испытаний нефтепродуктов рассматривается как случайная величина с функцией распределения в виде смеси нормальных законов распределения: «основного» с дисперсией, не превышающей установленные требования (при статистически управляемом состоянии процесса испытаний), и «засоряющего» с дисперсией значительно превышающей установленные требования (при статистически неуправляемом состоянии процесса испытаний).

Ключевые слова: нефтепродукты, измерения состава и свойств, закон распределения, идентификация, критерий согласия, вероятностная модель

ВВЕДЕНИЕ

Для метрологического обеспечения измерений состава и свойств нефтепродуктов важное значение имеет доказательность и обоснованность процедур обработки экспериментальных данных, используемых в различных процессах метрологического обеспечения, например, при аттестации методик измерения, определении сертифицированного (аттестованного) значения стандартных образцов.

Корректное применение известных процедур обработки многократных измерений состава и свойств нефтепродуктов, изложенных, например, в ГОСТ 33701-2015 «Определение и применение показателей точности методов испытаний нефтепродуктов», возможно только в случае справедливости допущения о том, что эмпирическая выборка результатов многократных наблюдений одной и той же величины подчиняется нормальному закону распределения случайной величины. Так как величины, характеризующие состав и свойства нефтепродуктов формируются под влиянием множества независимых (слабо зависимых) случайных факторов, основываясь на центральной предельной теореме, можно ожидать, что сумма

большого количества независимых (слабо зависимых) случайных величин с конечными математическими ожиданиями и (примерно одинаковыми) дисперсиями имеет нормальное распределение [1-2].

Практическая проверка данного допущения показала, что фактическое распределение погрешностей результатов измерений состава и свойств нефтепродуктов не соответствует нормальному закону распределения случайной величины. Эмпирический график плотности распределения погрешностей измерения представляет собой одномодальную асимметричную кривую с положительным эксцессом и тяжелыми хвостами [3]. Для формирования обоснованных алгоритмов обработки результатов многократных измерений состава и свойств нефтепродуктов проведены исследования по идентификации формы эмпирического закона распределения величин, характеризующего состав и свойства нефтепродуктов.

Целью исследования являлось нахождение универсального закона распределения (смеси законов распределения) справедливого для большинства величин, характеризующих состав и свойства нефтепродуктов и обоснование вероятностной

модели, описывающей результаты измерений состава и свойств нефтепродуктов.

ОСНОВНАЯ ЧАСТЬ

В качестве исходных данных использовали 21,3 тыс. результатов измерений состава и свойств различных типов и марок нефтепродуктов, полученных в ходе межлабораторных сравнительных испытаний в лабораториях горючего Минобороны России в 1967-2019 гг. и экспериментов по оценке показателей прецизионности квалификационных методов испытаний, проведенных в 2016-2019 гг. Измерения проводились с использованием общепринятых стандартизованных методик, указанных в табл. 1. Для получения однородных исходных данных абсолютные значения были пересчитаны в нормированные:

$$u_i = \frac{x_i - \bar{x}}{\hat{\sigma}},$$

где u_i – нормированное значение i -го члена выборки; x_i – значение i -го члена выборки; \bar{x} – среднее значение выборки; $\hat{\sigma}$ – стандартное отклонение выборки.

После проверки однородности полученных нормированных значений по критериям Андерсона-Дарлинга; Андерсона-Дарлинга-Петита; Лемана-Розенблатта; Смирнова; Жанга и Крускала-Уоллиса они были объединены в 56 выборок, характеристики которых представлены в табл. 1. Процедура получения объединенных выборок более подробно рассмотрена в работе [3]. Объединенные выборки нормированных значений являются представительными выборками из генеральной совокупности результатов, получаемых по той или иной методике измерений состава и свойств нефтепродуктов в целом. Следовательно, идентифицированный по этой выборке закон распределения можно рассматривать как общий закон распределения для всех результатов, получаемых по той или иной методике измерений.

Табл.1. Основные характеристики выборок исходных данных

Наименование величины, методика измерения	Характеристики выборки						
	Число элементов в выборке	Среднее арифметическое	Медиана	Дисперсия	Стандартное отклонение	Коэффициент асимметрии	Коэффициент экссесса
Плотность нефтепродуктов по ГОСТ 3900	1445	-0,0126	0,1019	1,124	1,060	-1,18	9,15
Кинематическая вязкость моторного топлива (при 20 °С) по ГОСТ 33	774	0,0576	0,0195	1,935	1,392	4,87	55,55
Кинематическая вязкость смазочных масел (при 100 °С) по ГОСТ 33	593	-0,0175	-0,0454	1,383	1,177	-1,12	15,34
Фракционный состав моторных топлив по ГОСТ 2177							
- температура начала перегонки	522	-0,0433	0,0467	1,183	1,088	-1,15	5,11
- температура выкипания 10 % фракции	523	-0,0049	0,0679	0,955	0,977	-0,11	0,56
- температура выкипания 50 % фракции	793	0,0000	-0,0130	0,946	0,973	0,12	0,23
- температура выкипания 90 % фракции	740	-0,0208	-0,0522	1,050	1,025	-0,20	2,39
- температура выкипания 95-98 % фракции	854	-0,0010	0,0504	1,020	1,010	-0,27	2,01
Температура вспышки в закрытом тигле по ГОСТ 6356	1030	0,0023	-0,0328	0,970	0,985	0,38	0,43
Температура вспышки в открытом тигле по ГОСТ 4333	488	0,0000	-0,0220	0,951	0,975	-0,13	0,30
Температура помутнения дизельного топлива по ГОСТ 5066	300	0,0000	0,0733	0,953	0,976	-0,17	0,21
Температура застывания по ГОСТ 20287	899	0,0000	-0,0254	0,949	0,974	0,65	0,14
Кислотность моторных топлив по ГОСТ 5985	728	0,0000	-0,0595	0,951	0,975	0,48	0,29
Кислотное число смазочных масел по ГОСТ 5985	301	0,0000	-0,1212	0,950	0,975	0,66	0,63
Щелочное число смазочных масел по ГОСТ 11362	389	0,0000	-0,0369	0,949	0,974	0,09	0,15
Наличие водорастворимых кислот и щелочей (рН водной вытяжки моторных топлив) по ГОСТ 6307	649	0,0018	0,0000	0,944	0,972	-0,10	0,08
Содержание меркаптановой серы в моторных топливах по ГОСТ 17323	134	0,0000	-0,2113	0,955	0,977	0,65	0,52
Зольность по ГОСТ 1461	688	0,0000	-0,0483	0,949	0,974	0,26	0,37
Коксуемость по ГОСТ 19932	392	0,0000	-0,0675	0,946	0,973	0,23	0,40
Химическая стабильность авиа-керосина по СТО 08151164-0271-2017:							
- масса поглощенного кислорода	430	0,0000	-0,0448	0,986	0,993	0,19	-0,07
- максимальная скорость поглощения кислорода	430	0,0000	-0,0843	0,986	0,993	-0,05	0,08
Химическая стабильность автомобильного бензина (доля поглощенного кислорода) по СТО 08151164-007-2009	204	-0,0116	0,0440	1,008	1,004	-0,16	-0,36
Содержание фактических смол в моторных топливах по ГОСТ 8489	882	-0,0015	-0,1430	0,953	0,976	0,52	0,09
Коррозионная активность авиакеросина в динамических условиях по СТО 08151164-0246-2017:							
- коррозионная активность, пластинка из бронзы ВБ-23НЦ	270	0,0000	0,0114	0,985	0,993	0,35	0,55
- коррозионная активность, пластинка из меди	270	0,0000	-0,1094	0,985	0,993	0,51	0,00
- забивка контрольного фильтро-элемента продуктами коррозии, пластинка из бронзы ВБ-23НЦ	270	0,0000	-0,2181	0,985	0,993	0,80	3,21

Наименование величины, методика измерения	Характеристики выборки						
	Число элементов в выборке	Среднее арифметическое	Медиана	Дисперсия	Стандартное отклонение	Коэффициент асимметрии	Коэффициент эксцесса
- забивка контрольного фильтроэлемента продуктами коррозии, пластинка из меди	270	0,0000	0,1222	0,985	0,993	0,17	0,39
Термоокислительная стабильность авиакеросина в статических условиях (масса осадка) по ГОСТ 11802	447	0,0000	-0,0671	0,949	0,974	0,35	-0,20
Термоокислительная стабильность авиакеросина в динамических условиях на установке ДТС-2 по СТО 08151164-007-2011:							
- индекс термостабильности	316	0,0000	0,1070	0,984	0,992	-0,41	-0,39
- перепад давления на фильтре	316	0,0000	0,0425	0,984	0,992	-0,26	-0,61
- температура начала образования отложений	316	0,0000	0,0733	0,984	0,992	-0,05	-0,09
Термоокислительная стабильность авиакеросина в динамических условиях на установке ДТС-4 по СТО 08151164-0305-2019:							
- индекс термостабильности	347	0,0000	0,0204	0,980	0,990	-0,09	-0,34
- перепад давления на фильтре	347	0,0000	-0,0473	0,980	0,990	0,21	-0,38
- температура начала образования отложений	347	0,0000	0,0314	0,980	0,990	-0,01	-0,52
Противоизносные свойства авиакеросина на установке ПСТ-2 по СТО 08151164-0164-2014:							
- диаметр износа большой	320	0,0000	0,0366	0,994	0,997	0,03	-0,54
- диаметр износа малый	320	0,0000	0,0143	0,994	0,997	-0,12	-1,00
- ширина дорожки износа	320	0,0000	0,0933	0,994	0,997	0,06	-0,39
Противоизносные свойства авиакеросина на установке ПСТ-3 по СТО 08151164-0245-2017:							
- диаметр износа большой	244	-0,0221	0,0117	1,086	1,042	-0,32	2,84
- диаметр износа малый	244	0,0189	-0,0127	1,054	1,027	0,54	1,41
- ширина дорожки износа	244	-0,0230	-0,1160	1,096	1,047	-0,78	2,41
Склонность масел к образованию высокотемпературных отложений на установке ВЦМ-1 по СТО 08151164-0132-2012:							
- индекс высокотемпературных отложений	204	-0,0122	0,0533	1,011	1,005	-0,04	-0,55
- температура начала образования отложений	204	0,0000	-0,0170	0,985	0,993	-0,10	0,15
Термоокислительная стабильность моторных топлив в условиях циркуляции на установке ЦИТО-М (скорость снижения температуры топлива за контрольным фильтро-элементом) по СТО 08151164-004-2009	96	-0,0621	-0,0501	1,161	1,077	-0,57	1,22
Характеристики сгорания авиакеросина на установке УНТ-1 по СТО 08151164-008-2009:							
- нагарное число	204	-0,0237	0,0224	1,095	1,046	-0,63	1,46
- дымное число	204	0,0000	0,0214	0,985	0,993	-0,21	-0,47
- световое число	204	0,0000	-0,0254	0,985	0,993	-0,03	-0,38
Нагарообразующие свойства дизельных топлив на установке ИДТ							
- нагарное число	204	-0,0111	-0,0019	1,005	1,003	0,23	-0,46
- масса нагара	204	0,0000	-0,0029	0,985	0,993	0,19	-0,36
Склонность бензинов к образованию отложений во впускной системе двигателя (масса отложений во впускной системе) по СТО 08151164-006-2009	204	-0,0121	0,0772	1,010	1,005	0,04	-0,37
Антиокислительные свойства и коррозионная активность моторных масел на моторной установке ИКМ по СТО 08151164-037-2010:							
- изменение массы верхних шатунных вкладышей	43	0,2129	-0,0091	1,815	1,347	1,45	3,34
- изменение массы нижних шатунных вкладышей	43	0,1222	0,0472	1,482	1,217	1,57	4,68
- общее изменение массы шатунных вкладышей	43	0,0954	-0,1125	1,255	1,120	1,01	1,74
- изменение вязкости моторного масла	43	-0,0833	0,1196	1,216	1,103	-0,95	1,29
Коррозионная активность моторных масел на установке КАММ-01М (изменение массы шатунных вкладышей) по СТО 08151164-037-2010	40	0,2615	0,1246	2,189	1,480	1,92	4,90
Склонность моторных масел к образованию отложений на установке НК (масса лака) по ГОСТ РВ 9150-001-2019	20	-0,1901	-0,1591	1,132	1,064	0,18	-0,50
Склонность моторных масел к образованию отложений на установке СК (масса лака) по ГОСТ РВ 9150-002-2019	20	-0,0901	-0,1823	1,102	1,050	-0,32	-0,48

Данные представленные в табл. 1, позволяют дать предварительную оценку закона распределения результатов измерения состава и свойств нефтепродуктов – в большинстве случаев эмпирическое распределение характеризуется

асимметричной кривой с положительным эксцессом и тяжелыми хвостами. Например, характеристики положения выборки измерения кинематической вязкости смазочных масел при положительной температуре (среднее значение и медиана) отличны от

нуля и смещены в отрицательную область: $-0,0175$ и $-0,0454$ соответственно, плотность вероятности достигает максимума в отрицательной области ($Mo(X) = -0,391$). Степень рассеяния значений почти на 20 % больше, чем в случае стандартного нормального распределения – стандартное отклонение составляет 1,177 против 1,0. Коэффициент асимметрии составляет $-1,12$, что свидетельствует о смещении кривой распределения в отрицательную область. Положительное значение коэффициента эксцесса 15,34 свидетельствует о том, что эмпирическое распределение имеет более высокую и острую вершину, чем стандартное нормальное распределение (коэффициент эксцесса равен 0).

Идентификация закона распределения наблюдаемой случайной величины (структурно-параметрическая идентификация) представляет собой задачу выбора параметрической модели закона распределения вероятностей, наилучшим образом соответствующей результатам экспериментальных наблюдений [4-5]. В настоящем исследовании был использован классический способ идентификации путем проверки сложных гипотез о согласии эмпирического и теоретического распределений в соответствии с рекомендациями Р 50.1.033-2001 и Р 50.1.037-2002.

В качестве теоретических распределений $F(x, \theta)$ рассматривали наиболее часто встречающихся на практике одномодальные законы с возможными значениями на всей числовой оси: распределение Лапласа, нормальное распределение, распределение Коши, логистическое распределение, распределение минимального значения, распределение максимального значения, распределение обобщенного минимального значения, распределение *Sh*-Джонсона, двустороннее экспоненциальное распределение, *H*-распределение, обобщенное логистическое распределение (*L*-распределение), распределение Стьюдента [6-7].

Изучение характеристик исследуемых выборок (табл. 1), графиков эмпирических плотностей распределения величин, характеризующих состава и свойства нефтепродуктов показало, что в большинстве случаев мы наблюдаем асимметричную кривую с положительным эксцессом и тяжелыми хвостами. Известно, что такой вид плотности распределения случайной величины может приобретать, когда эмпирическая выборка представляет собой смесь (сумму) двух симметричных распределений F_1 и F_2 вида:

$$F_c(x) = (1 - \varepsilon)F_1(x, \theta_1) + \varepsilon F_2(x, \theta_2),$$

где $F_1(x)$ – «основное» распределение; $F_2(x)$ – «засоряющее» распределение; $0 \leq \varepsilon \leq 1$ – доля «засоряющего» распределения F_2 .

Поэтому в качестве теоретических распределений рассматривали 66 смесей вышеуказанных одномодальных симметричных законов распределений случайных величин.

Проверяемая сложная гипотеза имеет вид:

$$H_0: F(x) \in \{F_0(x, \hat{\theta}), \hat{\theta} \in \theta_i\}, \quad (1)$$

$$H_0: F_c(x) \in \{(1 - \varepsilon)F_1(x, \hat{\theta}_1) + \varepsilon F_2(x, \hat{\theta}_2), \hat{\theta} \in \theta_i\}, \quad (2)$$

где $F(x, \theta)$ – функция распределения вероятностей известного вида, с оценкой параметра распределения θ , вычисленной по той же выборке, принадлежащая пространству параметров θ_i .

Проверку сложной гипотезы вида (1) проводили с использованием пяти критериев согласия: χ^2 -Пирсона, Колмогорова (с поправкой Большева), Смирнова, ω^2 Крамера-Мизеса-Смирнова; Ω^2 Андерсона-Дарлингга, по алгоритмам изложенным в п. 2.3 Р 50.1.033-2001 и п. 3.2 Р 50.1.037-2002 со следующими дополнениями.

Оценки параметров выбранного теоретического распределения проводили по имеющимся данным методом максимального правдоподобия в результате максимизации по θ функции $L(\theta) = \gamma \prod_{i=1}^n f(x_i, \theta)$.

Распределение статистики критерия согласия $G(S|H_0)$ при справедливости гипотезы H_0 определяли согласно п. 2.3.2 Р 50.1.037-2002 методом компьютерного анализа статистических закономерностей с использованием программы статистического анализа одномерных наблюдений «Интервальная статистика» ISW 4.0, разработанной на кафедре теоретической и прикладной информатики Новосибирского государственного технического университета [8].

Гипотеза о соответствии выборки выбранному закону распределения не отвергалась по тому или иному критерию согласия, если достигаемый уровень значимости $P(S > S^*)$ превышал заданный уровень значимости $\alpha = 0,1$.

Окончательное решение о соответствии выборки выбранному закону распределения принималось по совокупности значений достигаемого уровня значимости $P_i(S > S^*)$ всех пяти критериев согласия, если среднее значение всех $P_i(S > S^*)$ превышало заданный уровень значимости α :

$$\bar{P} = \frac{\sum_{i=1}^5 P_i}{5} \geq \alpha.$$

В этом случае считали, что данная модель теоретического закона не противоречит наблюдаемым данным.

Для вычисления статистики критерия χ^2 -Пирсона проводили группирование данных в выборке. Известно, что значение статистики критерия Пирсона существенно зависит от способа разбиения выборки на интервалы и количества интервалов разбиения. Исходя из литературных данных, нельзя отдать однозначное предпочтение тому или иному способу группирования и вычисления числа интервалов разбиения выборок и использовать их как

универсальный инструмент для проводимых вычислений. В связи с тем, что рассматриваемые выборки содержат нормированные значения, которые теоретически должны иметь одномодальное стандартное нормальное распределение, в процессе исследования подбирали количество интервалов и способ разбиения таким образом, чтобы получаемые гистограммы имели вид одномодальных кривых. Гистограмма в этом случае является наиболее наглядным способом выбрать подходящий метод разбиения выборки на интервалы.

Проверку сложной гипотезы вида (2) проводили с использованием одного критерия согласия – критерия Хи-квадрат Пирсона. Потому что для сложной гипотезы вида (2) критерии Колмогорова, Смирнова, ω^2 Крамера-Мизеса-Смирнова, Ω^2 Андерсона-Дарлинга не имеют предельного распределения, что не позволяет вычислить критическое значение статистики S и достигнутый уровень значимости $P(S>S^*)$ для указанных критериев согласия.

В соответствии с принятым алгоритмом, была проведена идентификация закона распределения в отношении 56 объединенных выборок нормированных значений, характеризующих методики измерений состава и свойств нефтепродуктов. В качестве примера рассмотрим процедуру идентификации закона распределения результатов измерений кинематической вязкости смазочных масел при 100°C по ГОСТ 33 «Нефть и нефтепродукты. Прозрачные и непрозрачные жидкости. Определение кинематической и динамической вязкости». Общая характеристика выборки представлена в табл. 1. Результаты проверки сложной гипотезы вида (1) представлены в табл. 2. Установлено, что при заданном уровне значимости $\alpha=0,1$ не отвергается гипотеза о соответствии значений измерения кинематической вязкости смазочных масел при 100°C двум законам распределения:

- распределению Лапласа с параметром положения $\mu=-0,0577$ и параметром масштаба $\lambda=0,8061$, среднее значение достигаемого уровня значимости по совокупности рассматриваемых критериев составляет $P=0,2051$;

- распределению Su -Джонсона с параметром положения $\mu=-0,1846$ и параметром масштаба $\lambda=0,7835$, параметрами формы $a=-0,1833$ и $b=1,0286$, среднее значение достигаемого уровня значимости по совокупности рассматриваемых критериев составляет $P=0,1272$.

В отношении остальных законов распределения сложная гипотеза (1) уверенно отклоняется, среднее значение достигаемого уровня значимости по совокупности рассматриваемых критериев находится в интервале от 0 до 0,009. В частности, гипотеза о соответствии значений измерения кинематической вязкости смазочных масел при 100°C нормальному

распределению с параметром положения $\mu=0,0099$ и параметром масштаба $\lambda=1,0049$, отвергается при среднем значении достигаемого уровня значимости по совокупности рассматриваемых критериев $P=8,26 \cdot 10^{-6}$.

Результаты проверки сложной гипотезы вида (2) в отношении каждой смеси рассматриваемых законов распределения случайной величины представлены в табл. 3. Установлено, что при заданном уровне значимости $\alpha=0,1$ не отвергается гипотеза о соответствии значений измерения кинематической вязкости смазочных масел при 100°C четырнадцати смесям законов распределения, при этом более половины (8 из 14) идентифицированных смесей распределения представляют собой смеси на основе нормального закона распределения. Максимальную вероятность согласия со сложной гипотезой (2) имеет смесь из 69 % нормального распределения с параметром положения $\mu=0,0920$ и параметром масштаба $\lambda=1,2357$ и 31 % нормального распределения с параметром положения $\mu=-0,1540$ и параметром масштаба $\lambda=0,3839$, значение достигаемого уровня значимости по критерию χ^2 -Пирсона составляет $P=0,4690$.

Сравнивая гипотезы о соответствии объединенной выборки результатов измерения кинематической вязкости смазочных масел при 100°C одной из альтернатив: распределению Лапласа или смеси двух нормальных распределений необходимо отметить следующее. Во-первых, среднее значение достигаемого уровня значимости с гипотезой о соответствии выборки смеси нормальных законов распределения более чем в 2 раза выше, чем вероятность согласия с гипотезой о соответствии выборки симметричному одномодальному закону распределения (закону Лапласа): 0,4690 против 0,2051. Во-вторых, гипотеза о соответствии выборки смеси из двух нормальных распределений с разными значениями параметров положения и масштаба более адекватно описывает имеющуюся выборку – ее асимметричность, большее рассеяние результатов, наличие «тяжелых» хвостов.

Таким образом, проведенная проверка сложных гипотез (1) и (2), анализ самой выборки позволяют принять гипотезу о том, что закон распределения результатов измерения кинематической вязкости смазочных масел при 100°C представляет собой смесь двух нормальных распределений – 69 % нормального распределения с параметром положения $\mu=0,0920$ и параметром масштаба $\lambda=1,2357$ и 31 % нормального распределения с параметром положения $\mu=-0,1540$ и параметром масштаба $\lambda=0,3839$.

Аналогичным образом проводилась идентификация закона распределения по остальным 55 объединенным выборкам нормированных значений. Было установлено следующее.

Табл.2. Результаты проверки сложной гипотезы (1)

Наименование распределения	Параметры распределения				Результаты проверки согласия при уровне значимости $\alpha = 0,1$			
	Масштаба	Положения	Формы 1	Формы 2	Критерий	Значение статистики критерия	Вероятность согласия, P	Итог
Распределение Лапласа	0,8061	-0,0577	нет	нет	χ^2 -Пирсона	21,2159	0,1703	Гипотеза о согласии НЕ ОТВЕРГАЕТСЯ при $P = 0,2051$
					Колмогорова	1,0996	0,0107	
					Смирнова	0,4773	0,7981	
					ω^2	0,1671	0,0297	
					Ω^2	1,2981	0,0162	
Нормальное распределение	1,0049	0,0099	нет	нет	χ^2 -Пирсона	55,3001	$3,17 \cdot 10^{-6}$	Гипотеза о согласии НЕ ОТВЕРГАЕТСЯ при $P = 8,26 \cdot 10^{-6}$
					Колмогорова	1,5971	$1,34 \cdot 10^{-5}$	
					Смирнова	10,2032	$2,46 \cdot 10^{-5}$	
					ω^2	0,5542	0	
					Ω^2	$2 \cdot 10^{100}$	0	
Логистическое распределение	0,5709	-0,0100	нет	нет	χ^2 -Пирсона	34,6734	0,0044	Гипотеза о согласии НЕ ОТВЕРГАЕТСЯ при $P = 0,0009$
					Колмогорова	1,2830	$5,27 \cdot 10^{-5}$	
					Смирнова	6,5845	$2,81 \cdot 10^{-5}$	
					ω^2	0,2633	$5,88 \cdot 10^{-6}$	
					Ω^2	1,5124	$5,53 \cdot 10^{-5}$	
Распределение Су-Джонсона	0,7835	-0,1847	-0,1833	1,0286	χ^2 -Пирсона	14,8662	0,5344	Гипотеза о согласии НЕ ОТВЕРГАЕТСЯ при $P = 0,1272$
					Колмогорова	0,6887	0,0759	
					Смирнова	1,8974	0	
					ω^2	0,0644	0,0256	
					Ω^2	0,7701	$4,67 \cdot 10^{-5}$	
Н-распределение	2,5445	1,7441	2,9479	0,6571	χ^2 -Пирсона	551,8476	0	Гипотеза о согласии НЕ ОТВЕРГАЕТСЯ при $P = 0$
					Колмогорова	11,3811	0	
					Смирнова	518,1234	0	
					ω^2	44,9744	0	
					Ω^2	$2 \cdot 10^{100}$	0	

Табл.3. Результаты проверки сложной гипотезы (2)

Состав смесей законов распределений	Параметры распределения				Результаты проверки согласия при уровне значимости $\alpha = 0,1$		
	Масштаба	Положения	Формы 1	Формы 2	Значение статистики критерия χ^2 -Пирсона	Вероятность согласия, P	Итог
Смесь распределений							
69 % Нормальное	1,2357	0,0920	нет	нет	14,7571	0,4690	Гипотеза о согласии НЕ ОТВЕРГАЕТСЯ
31 % Нормальное	0,3839	-0,1540	нет	нет			
Смесь распределений							
47 % Лапласа	1,2212	0,0471	нет	нет	18,2631	0,3087	Гипотеза о согласии НЕ ОТВЕРГАЕТСЯ
53 % Нормальное	0,6578	-0,0721	нет	нет			
Смесь распределений							
77 % Логистическое	0,6803	0,0832	нет	нет	14,9730	0,4533	Гипотеза о согласии НЕ ОТВЕРГАЕТСЯ
23 % Нормальное	0,3255	-0,2035	нет	нет			
Смесь распределений							
53 % Коши	0,4475	-0,1917	нет	нет	14,9260	0,4567	Гипотеза о согласии НЕ ОТВЕРГАЕТСЯ
47 % Нормальное	0,9383	0,2069	нет	нет			
Смесь распределений							
74 % Нормальное	1,1756	0,0948	нет	нет	19,2774	0,1125	Гипотеза о согласии НЕ ОТВЕРГАЕТСЯ
26 % Обобщенное минимального значения	0,7166	-1,3138	5,0444	нет			
Смесь распределений							
69 % Нормальное	1,2377	0,0922	нет	нет	15,7288	0,4003	Гипотеза о согласии НЕ ОТВЕРГАЕТСЯ
31 % Су-Джонсона	2,6323	-0,2269	-0,1897	6,8667			
Смесь распределений							
68 % Нормальное	1,2519	0,0911	нет	нет	19,3499	0,1983	Гипотеза о согласии НЕ ОТВЕРГАЕТСЯ
32 % Двустороннее экспоненциальное	0,5627	-0,1427	1,9250	нет			
Смесь распределений							
71 % Нормальное	1,2196	0,0948	нет	нет	20,6401	0,1487	Гипотеза о согласии НЕ ОТВЕРГАЕТСЯ
29 % L-распределение	0,6237	0,1031	5,0606	7,6047			
Смесь распределений							
24 % Коши	0,3915	0,7103	нет	нет	22,4341	0,1297	Гипотеза о согласии НЕ ОТВЕРГАЕТСЯ
76 % Коши	0,4200	-0,2355	нет	нет			
Смесь распределений							
50 % Минимального значения	1,2588	0,6907	нет	нет	19,3277	0,2520	Гипотеза о согласии НЕ ОТВЕРГАЕТСЯ
50 % Максимального значения	0,4363	-0,3159	нет	нет			

Состав смесей законов распределений	Параметры распределения				Результаты проверки согласия при уровне значимости $\alpha = 0,1$		
	Масштаба	Положения	Формы 1	Формы 2	Значение статистики критерия χ^2 -Пирсона	Вероятность согласия, P	Итог
Смесь распределений					17,6327	0,3458	Гипотеза о согласии НЕ ОТВЕРГАЕТСЯ
45 % Минимального значения	1,1839	0,9027	нет	нет			
55 % Обобщенного минимального значения	1,0479	-1,5470	3,9735	нет			
Смесь распределений					18,2724	0,3081	Гипотеза о согласии НЕ ОТВЕРГАЕТСЯ
18 % Максимального значения	2,4346	-0,8430	нет	нет			
82 % Су-Джонсона	3,6117	-0,8844	-1,2092	5,0844			
Смесь распределений					17,8423	0,3332	Гипотеза о согласии НЕ ОТВЕРГАЕТСЯ
47 % Обобщенное минимального значения	1,3333	0,5524	1,2313	нет			
53 % Обобщенное минимального значения	1,0141	-1,4584	3,8002	нет			
Смесь распределений					23,0421	0,1126	Гипотеза о согласии НЕ ОТВЕРГАЕТСЯ
28 % L-распределение	0,4282	0,2615	2,2694	6,1423			
72 % L-распределение	1,6011	0,2182	3,8765	4,1374			

Проверка гипотезы вида (1) о соответствии какому-либо симметричному одномодальному закону распределения показала, что для двух выборок, характеризующих методику измерения противоизносных свойств авиакеросина на установке ПСТ-2 по величинам «диаметр износа большой» и «ширина дорожки износа» сложная гипотеза вида (1) отвергается для всех исследованных законов распределения случайной величины. Для остальных пятидесяти четырех выборок число законов распределения в отношении, которых гипотеза вида (1) не отвергается находится в интервале от 1 до 10. Так, для объединенной выборки результатов измерения кинематической вязкости моторного топлива при 20°C, температуры вспышки нефтепродуктов в закрытом тигле и др. сложная гипотеза вида (1) не отвергается в отношении одного закона распределения случайной величины. Для объединенной выборки результатов измерения склонности моторных масел к образованию отложений на установке «скользящее кольцо» гипотеза вида (1) не отвергается в отношении десяти законов распределения случайной величины.

Таким образом, результаты проверки гипотезы вида (1) не позволяют сделать вывод о соответствии объединенных выборок нормированных значений, характеризующих методики измерений состава и свойств нефтепродуктов какому-либо одному универсальному закону распределения случайной величины. Можно только отметить, что наиболее часто сложная гипотеза вида (1) не отвергалась в отношении обобщенного логистического распределения (L-распределение), распределения Лапласа и двустороннего экспоненциального распределения.

Результаты проверки сложной гипотезы вида (2) по каждой из 56 объединенных выборок нормированных значений, характеризующих методики измерений состава и свойств нефтепродуктов показали, что для всех выборок не

отвергается гипотеза о соответствии выборки классу смесей распределений на основе нормального закона распределения. При этом для подавляющего большинства выборок (54 из 56 рассматриваемых) не отвергается гипотеза о соответствии смеси из двух нормальных законов распределения. Исключение составляют выборки характеризующие результаты измерения фракционного состава моторных топлив по величине «температура выкипания 90 % фракции» и противоизносных свойств авиакеросина на установке ПСТ-2 по величине «ширина дорожки износа». Более того гипотеза о соответствии выборки смеси из двух нормальных законов распределения, как правило, имеет максимальное значение достигаемого уровня значимости. Только для трех выборок (кислотность моторных топлив, pH водной вытяжки и химическая стабильность автомобильного бензина) достигаемый уровень значимости гипотезы о соответствии выборки смеси из двух нормальных законов распределения имеет второе по величине значение.

Сравнение полученных данных показало, что для одной и той же выборки значения достигаемого уровня значимости по гипотезе (2) в несколько раз выше среднего значения достигаемого уровня значимости (по совокупности критериев) по гипотезе (1). Например, для выборки результатов измерения индекса термостабильности на установке ДТС-2 вероятность согласия с гипотезой о соответствии выборки смеси двух нормальных распределений 51 % нормального распределения с параметром положения $\mu=0,6677$ и параметром масштаба $\lambda=0,6402$ и 49 % нормального распределения с параметром положения $\mu=-0,6679$ и параметром масштаба $\lambda=0,8476$ составляет $P=0,9936$. А для той же самой выборки максимальное среднее значение достигаемого уровня значимости (по совокупности критериев) о соответствии распределению минимального значения с параметром положения $\mu=0,4756$ и параметром масштаба $\lambda=0,8601$ составляет $P=0,2266$.

Проведенные исследования по определению модели закона распределения, позволяют предложить новую обобщенную вероятностную модель, описывающую результаты многократных измерений состава и свойств нефтепродуктов, как совокупности статистически управляемого состояния процесса испытаний, в котором 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, и статистически неуправляемого состояния, когда 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:

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

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

Такая модель позволяет описать фактически наблюдаемую картину. В условиях хорошо налаженной работы лаборатории результаты измерений состава и свойств нефтепродуктов соответствуют «основному» нормальному закону $F(x)$ с небольшой дисперсией σ_0^2 . При нарушениях в работе лаборатории результаты измерений состава и свойств нефтепродуктов соответствуют «засоряющему» нормальному закону $F_1(x)$ с дисперсией σ_1^2 много большей, чем дисперсия σ_0^2 .

ЗАКЛЮЧЕНИЕ

1. Проведенные исследования показали, что гипотеза о согласии эмпирического распределения с теоретическим в виде смеси двух нормальных распределений с разными значениями параметров положения и масштаба не противоречит экспериментальным данным и имеет универсальный характер, справедливый для всех рассмотренных результатов измерений состава и свойств нефтепродуктов. Достижимый уровень значимости по данной гипотезе в несколько раз выше среднего значения достигаемого уровня значимости (по совокупности рассматриваемых критериев) гипотезы о соответствии какому-либо симметричному одномодальному закону распределения.

2. На основе исследований модели закона распределений результатов измерений состава и

свойств нефтепродуктов обоснована обобщенная вероятностная модель процесса испытаний нефтепродуктов, рассматривающая результат испытаний нефтепродуктов как случайную величину с функцией распределения в виде смеси нормальных законов распределения: «основного» с дисперсией, не превышающей установленные требования (при статистически управляемом состоянии процесса испытаний) и «засоряющего» с дисперсией значительно превышающей установленные требования (при статистически неуправляемом состоянии процесса испытаний).

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Бочаров, П.П., Теория вероятностей. Математическая статистика [Текст] / П.П. Бочаров, А.В. Печинкин. – 2-е изд. – М.: Физматлит, 2005. – 296 с.
2. Гнеденко, Б.В. Курс теории вероятностей: учебник [Текст] / Б.В. Гнеденко. – 8-е изд., испр. и доп. – М.: Издательство УРСС, 2005. – 448 с.
3. Шаталов, К.В. Проверка гипотезы о соответствии погрешностей результатов анализа нефтепродуктов нормальному закону распределения случайной величины [Текст] / К.В. Шаталов, А.Д. Черепанова // Измерительная техника. – 2019. – № 10. – С. 52-60.
4. Лемешко, Б.Ю. О задаче идентификации закона распределения случайной составляющей погрешности измерений [Текст] / Б.Ю. Лемешко // Метрология. – 2004. – №7. – С. 8-17.
5. Боровков, А.А. Математическая статистика. Оценка параметров. Проверка гипотез [Текст] / А.А. Боровков. – М.: Наука. Главная редакция физико-математической литературы, 1984. – 472 с.
6. Вадзинский, Р.Н. Справочник по вероятностным распределениям [Текст] / Р.Н. Вадзинский. – СПб.: Наука, 2001. – 295 с.
7. Губарев, В.В. Вероятностные модели: справочник. В 2-х частях. Часть 1. [Текст] / В.В. Губарев. – Новосибирск: Новосибирский электротехнический институт, 1992. – 198 с.
8. Программа статистического анализа интервальных наблюдений одномерных непрерывных случайных величин «Интервальная статистика ISW». Версия: 4.4.1.105. [Электронный ресурс] / Б.Ю. Лемешко, С.Н. Постовалов. – Режим доступа: <https://ami.nstu.ru/~headrd/ISW.htm>

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.

Черепанова Анна Дмитриевна – к.х.н., научный сотрудник отдела нефтяных и альтернативных топлив ФАУ «25 ГосНИИ химмотологии Минобороны России», тел. (499)141-97-51, e-mail: kulazhskaya_92@inbox.ru.

A DISTRIBUTION LAW OF THE RESULTS OF MEASUREMENT OF COMPOSITION AND PROPERTIES OF PETROLEUM PRODUCTS

K.V. Shatalov, A.D. Cherepanova

Federal Autonomous Enterprise «The 25-th State Research Institute of Himmotology, Ministry of Defence of Russian Federation», Moscow

The identification of a distribution law of the results of measurements of the composition and properties of petroleum products was carried out by testing a complex hypothesis using five goodness-of-fit tests: χ^2 -Pearson, Kolmogorov, Smirnov, ω^2 Cramer-Mises-Smirnov, Ω^2 Anderson-Darling. Twelve symmetric unimodal distribution laws and 66 mixtures of the same distribution laws were considered as possible probability distribution functions. The purpose of the identification was to find a universal distribution law (a mixture of distribution laws) that is valid for all considered quantities. Testing the complex hypothesis of compliance with any symmetric unimodal distribution law showed that there is no universal distribution law that is valid for all measurement techniques of the composition and properties of petroleum products; most often the hypothesis of the correspondence the data to the generalized logistic distribution, the Laplace distribution and the two-sided exponential distribution was not rejected. Testing a complex hypothesis about the correspondence of any mixture of symmetric unimodal distribution laws showed that the empirical distribution function of the results of measurements of the composition and properties of petroleum products can be represented as a mixture of two normal distributions with different values of the position and scale parameters. At the same time, for the same sample, the values of the achieved significance level of the hypothesis about the correspondence to the mixture of distribution laws is several times higher than the average value of the achieved significance level of the hypothesis about the correspondence to one distribution law. Based on this study, a chance model of the process of testing petroleum products was substantiated, within which the result of testing petroleum products is considered as a random variable with a distribution function in the form of a mixture of normal distribution laws: "basic" with a variance not exceeding the established requirements (with a statistically controlled state of the test process), and "fouling" with a variance significantly exceeding the established requirements (with a statistically uncontrolled state of the test process).

Key words: petroleum products, measurements of composition and properties, distribution law, identification, goodness of fit of a distribution, chance model

REFERENCES

1. Bocharov, P.P., Pechinkin A.V. *Teoriya veroyatnostej. Matematicheskaja statistika.* – 2-e izd. – M.: Fizmatlit, 2005. – 296 s.
2. Gnedenko B.V. *Kurs teorii veroyatnostej: uchebnik.* – 8-e izd., ispr. i dop. – M.: Izdatel'stvo URSS, 2005. – 448 s.
3. K.V. Shatalov, A.D. Cherepanova. Verification of the hypothesis of compliance of the errors of the results of analyses of oil products to the normal distribution law for a random variable // *Measurement Techniques*, Vol. 62, No. 10, 2020.
4. Lemesenko B. Yu. O zadache identifikacii zakona raspredeleniya sluchajnoj sostavlyayushej pogreshnosti izmerenij // *Metrologiya.* – 2004. №7. – 8-17 p.
5. Borovkov A. A. *Matematicheskaya statistika. Ocenka parametrov. Proverka gipotez.* – M.: Nauka. Glavnaya redakciya fiziko-matematicheskoy literature, 1984. – 472 p.
6. Vadzinskij R. N. *Spravochnik po veroyatnostnim raspredeleniyam.* – SPb.: Nauka, 2001. – 295 p.
7. Gubarev, V.V. *Veroyatnostnye modeli: spravochnik. V 2-h chastyah. Chast' 1.* – Novosibirsk: Novosibirskij jelektrotehnicheskij institut, 1992. – 198 s.
8. Programma statisticheskogo analiza interval'nikh nablyudenij odnovernikh neprerivnikh sluchajnikh velichin "Interval'naya statistica ISW" Versiya 4.4.1.105 [Elektronnyj resurs] / B. Yu. Lemesenko, S.N. Postovalov. – <https://ami.nstu.ru/~headrd/ISW.htm>

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.

Cherepanova Anna Dmitrievna - Candidate of Chemical Sciences, Researcher of the Fuel Department Federal Autonomous Enterprise «The 25-th State Research Institute of Himmotology, Ministry of Defence of Russian Federation», tel. (499)141-97-51, e-mail: kulazhskaya_92@inbox.ru.