*Метрология. 2012. № 5. – С. 3-26.*

УДК 519.24

# О НОРМАЛЬНОСТИ ПОГРЕШНОСТЕЙ ИЗМЕРЕНИЙ В КЛАССИЧЕСКИХ ЭКСПЕРИМЕНТАХ И МОЩНОСТИ КРИТЕРИЕВ, ПРИМЕНЯЕМЫХ ДЛЯ ПРОВЕРКИ ОТКЛОНЕНИЯ ОТ НОРМАЛЬНОГО ЗАКОНА

Лемешко Б.Ю., Рогожников А.П.[1]

Применение совокупности критериев, используемых для проверки отклонения эмпирических распределений от нормального закона, демонстрируется на анализе результатов классических экспериментов, связанных с измерением физических констант. Сравнительный анализ мощности множества рассматриваемых критериев относительно возможных конкурирующих законов проведен при объемах выборок, фигурирующих в рассматриваемых экспериментах.

**Ключевые слова**: мощность критерия, критерии нормальности, непараметрические критерии согласия, критерии типа хи-квадрат, проверка сложных гипотез

**Введение.** При проведении экспериментальных исследований нельзя исключать возможного наличия факторов, которые могут приводить к систематическим ошибкам, к смещенности оценок параметров, к коррелированности результатов измерений, к появлению тренда в той или иной форме. Далеко не всегда реальные измерения (ошибки измерений) представляют собой выборки независимых одинаково распределенных нормальных величин. Само по себе это не является чем-то особенным, но приводит к определенным проблемам при анализе, к появлению вопросов, связанных с точностью измерений и корректностью статистических выводов. Поэтому при статистическом анализе результатов измерений первым из проверяемых предположений обычно является проверка гипотезы о принадлежности ошибок нормальному закону. Если гипотеза о нормальности не отвергается, то дальнейший анализ становится проще, так как в его рамках открывается возможность использования множества классических критериев, применение которых корректно при выполнении предположения о нормальности.

[1] Новосибирский государственный технический университет

Результаты классических экспериментов, связанных с измерениями физических констант, привлекают интерес не только в силу их исторической ценности, но и как примеры высокого уровня организации и проведения соответствующих измерений. Принято считать, что ошибки измерений в этих экспериментах подчиняются нормальному закону. Чтобы убедиться в этом, к проверке гипотез о принадлежности ошибок измерений в таких экспериментах нормальному закону возвращались неоднократно, причем, с использованием различных критериев. В некоторых случаях, анализируя классические результаты, авторы работ сравнивают используемые критерии по мощности [1] относительно определенных конкурирующих гипотез. Приводимые оценки мощности позволяют судить о степени уверенности в том, что ошибки в данных экспериментах действительно подчиняются нормальному закону.

В то же время у специалистов в области математической статистики до сих пор нет однозначного мнения относительно преимуществ какого-либо критерия или группы критериев из множества, применяемых для проверки гипотез о принадлежности выборки случайных величин нормальному закону. Особую остроту ответ на этот вопрос приобретает при достаточно ограниченных объемах выборок, которые характерны не только для классических экспериментов. В последние время интерес к критериям, используемым для проверки нормальности, заметно повысился [2, 3]. Возрос интерес к непараметрическим критериям согласия, применяемым для этих целей [4, 5]. Появились работы, посвященные сравнительному анализу и исследованию мощности модифицированных критериев типа $\chi^2$, используемым для проверки нормальности [6]. При этом в исследованиях возрастает роль методов статистического моделирования.

Сталкиваясь с разрозненными примерами анализа классических результатов, не возникает полноты картины и уверенности в том, что нормальный закон в этих экспериментах наилучшим образом описывает ошибки измерений, и что для проверки гипотез о принадлежности ошибок измерений нормальному закону используются наиболее мощные критерии.

В работах [7, 8] мы подробно исследовали свойства и мощность ряда критериев, предназначенных только для проверки нормальности. Эти исследования позволили, с одной стороны, проранжировать исследуемые критерии по мощности относительно рассмотренных конкурирующих законов, а, с другой стороны, выявили существенные недостатки некоторых критериев. Сравнительный анализ мощности критериев согласия, в том числе при проверке нормальности, осуществлялся в работах [9, 10, 11].

Специалист в конкретной области, анализирующий результаты измерений, может не обладать достаточной квалификацией во всём множестве методов статистического анализа, но ему желательно знать, какие критерии и в каком случае наиболее предпочтительны. Поэтому цель данной работы заключалась в следующем. Во-первых, при использовании различных критериев продемонстрировать результаты проверки нормальности ошибок измерений, проведенных в некоторых классических экспериментах. Во-вторых, оценить мощность применяемых критериев относительно тех возможных конкурирующих законов (конкурирующих гипотез), которые, как и нормальный закон, достаточно хорошо описывают эти ошибки. Анализ мощности провести при тех объемах выборок, которые соответствуют объемам выборок в рассматриваемых экспериментах.

**Анализируемые эксперименты.** В статье [12] отобрано несколько экспериментов, удовлетворяющих довольно строгим требованиям. Мы рассмотрим четыре эксперимента, измерения в которых проводились высококвалифицированными учеными. Измерялись хорошо известные в настоящий момент величины, а обработка характеризовалась скрупулезным устранением погрешностей, поддающихся учету. Измерения в выборках фиксировались полностью и без пропусков (и без цензурирования).

Генри Кавендиш провел эксперимент по определению средней плотности Земли. Результаты 29 его измерений приведены в таблице 1.

Роберт Милликен провел эксперимент с заряженными каплями масла для измерения заряда электрона, и его результаты приведены в таблице 2.

Таблица 3 включает 100 наблюдений Альберта Майкельсона в опыте по измерению скорости света. Измерялось время прохождения света (туда и обратно) между быстро вращающимся и отстоящим от первого неподвижно закрепленным зеркалом. Расстояние между зеркалами составило приблизительно 605 метров (1986.23 фута).

Таблица 1. Измерения средней плотности Земли, полученные Генри Кавендишем, (г/см$^3$).

| | | | | | |
|---|---|---|---|---|---|
| 5.50 | 5.55 | 5.57 | 5.34 | 5.42 | 5.30 |
| 5.61 | 5.36 | 5.53 | 5.79 | 5.47 | 5.75 |
| 4.88 | 5.29 | 5.62 | 5.10 | 5.63 | 5.68 |
| 5.07 | 5.58 | 5.29 | 5.27 | 5.34 | 5.85 |
| 5.26 | 5.65 | 5.44 | 5.39 | 5.46 | |

Таблица 2. Измерения заряда электрона, полученные Робертом Милликеном, (Фр =1/2 997 924 580 Кл)

| | | | | | |
|---|---|---|---|---|---|
| 4.781 | 4.764 | 4.777 | 4.809 | 4.761 | 4.769 |
| 4.795 | 4.776 | 4.765 | 4.790 | 4.792 | 4.806 |
| 4.769 | 4.771 | 4.785 | 4.779 | 4.758 | 4.779 |
| 4.792 | 4.789 | 4.805 | 4.788 | 4.764 | 4.785 |
| 4.779 | 4.772 | 4.768 | 4.772 | 4.810 | 4.790 |
| 4.775 | 4.789 | 4.801 | 4.791 | 4.799 | 4.777 |
| 4.772 | 4.764 | 4.785 | 4.788 | 4.799 | 4.749 |
| 4.791 | 4.774 | 4.783 | 4.783 | 4.797 | 4.781 |
| 4.782 | 4.778 | 4.808 | 4.740 | 4.790 | |
| 4.767 | 4.791 | 4.771 | 4.775 | 4.747 | |

Таблица 3. Измерения Майкельсона (+299000 км/ч).

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 850 | 1000 | 960 | 830 | 880 | 880 | 890 | 910 | 890 | 870 |
| 740 | 980 | 940 | 790 | 880 | 910 | 810 | 920 | 840 | 870 |
| 900 | 930 | 960 | 810 | 880 | 850 | 810 | 890 | 780 | 810 |
| 1070 | 650 | 940 | 880 | 860 | 870 | 820 | 860 | 810 | 740 |
| 930 | 760 | 880 | 880 | 720 | 840 | 800 | 880 | 760 | 810 |
| 850 | 810 | 800 | 830 | 720 | 840 | 770 | 720 | 810 | 940 |
| 950 | 1000 | 850 | 800 | 620 | 850 | 760 | 840 | 790 | 950 |
| 980 | 1000 | 880 | 790 | 860 | 840 | 740 | 850 | 810 | 800 |
| 980 | 960 | 900 | 760 | 970 | 840 | 750 | 850 | 820 | 810 |
| 880 | 960 | 840 | 800 | 950 | 840 | 760 | 780 | 850 | 870 |

Саймон Ньюкомб тремя годами позже провел более обширный эксперимент и уточнил результаты Альберта Майкельсона. Расстояние между зеркалами составило 3721.21 м. В таблице 4 приведены 66 полученных значений.

Результаты измерений Кавендиша, Майкельсона и Ньюкомба приводятся по работе [12], а Милликена – по статье [13].

Таблица 4. Измерения Саймона Ньюкомба ($\times 10^{-3}$ + 24.8, в миллионных долях секунды)

| | | | | | | |
|---|---|---|---|---|---|---|
| 28 | 29 | 24 | 37 | 36 | 26 | 29 |
| 26 | 22 | 20 | 25 | 23 | 32 | 27 |
| 33 | 24 | 36 | 28 | 27 | 32 | 28 |
| 24 | 21 | 32 | 26 | 27 | 24 | 29 |
| 34 | 25 | 36 | 30 | 28 | 39 | 16 |
| -44 | 30 | 28 | 32 | 27 | 28 | 23 |
| 27 | 23 | 25 | 36 | 31 | 24 | |
| 16 | 29 | 21 | 26 | 27 | 25 | |
| 40 | 31 | 28 | 30 | 26 | 32 | |
| -2 | 19 | 29 | 22 | 33 | 25 | |

**Рассматриваемые критерии нормальности.** В данной работе мы не приводим выражений статистик используемых критериев, так как это очень сильно увеличило бы объем статьи, а ограничились лишь ссылками на первоисточники и работы, где можно подробно ознакомиться с описанием критерия или с результатами исследований.

В работе рассмотрены следующие критерии нормальности:

– критерий Шапиро-Уилка [14, 15];
– критерий Шапиро-Уилка для больших выборок (критерий Ройстона [16]);
– критерий Эппса-Палли [17];
– Д’Агостино $z_2$ [3,18];
– критерия Фросини [19, 20];
– критерии Хегази и Грина [21] со статистиками $T_1$ и $T_2$;
– критерий Гири [22, 23, 24];
– критерий Дэвида-Хартли-Пирсона [25];
– критерий Шпигельхальтера [26], статистика которого базируется на комбинации статистик критериев Гири и Дэвида-Хартли-Пирсона.

Сведения о перечисленных критериях наиболее доступны в [7, 8] или в [27]. Критерий Ройстона [16] – это самостоятельный критерий, он является развитием критерия Шапиро-Уилка [14, 15]. В отличие от критерия Шапиро-

Уилка – это правосторонний критерий, и для него отсутствуют ограничения по объему выборок.

Для проверки нормальности ошибок измерений использовались также непараметрические критерии согласия:

- Колмогорова;
- Крамера-Мизеса-Смирнова;
- Андерсона-Дарлинга.

В этом случае непараметрические критерии используются в ситуации проверки сложных гипотез. Напомним, что при проверке сложных гипотез распределения статистик данных критериев зависят от ряда факторов: от вида наблюдаемого закона $F(x,\theta)$, соответствующего справедливой проверяемой гипотезе $H_0$; от типа оцениваемого параметра и числа оцениваемых параметров; в некоторых случаях от конкретного значения параметра (например, в случае семейств гамма-, бета-распределений и т.п.); от метода оценивания параметров [27, 28, 29]. Оценивание параметров распределений проводилось методом максимального правдоподобия, так как в случае этого метода оценивания непараметрические критерии, как правило, обладают большей мощностью.

Для анализа результатов измерений рассматриваемых экспериментов, кроме вышеуказанных критериев, использовались критерий согласия $\chi^2$ Пирсона и модифицированный критерий типа $\chi^2$ Никулина (Никулина-Рао-Робсона) [30, 31, 32, 33].

Статистика $X_n^2$ критерия Пирсона при проверке сложной гипотезы подчиняется $\chi_r^2$-распределению с $r = k - m - 1$ степенями свободы, где $k$ – число интервалов, при оценивании $m$ параметров закона минимизацией этой статистики или методом максимального правдоподобия по сгруппированным наблюдениям. Если параметры оцениваются по негруппированной выборке, то известно [34], что при справедливости проверяемой гипотезы и использовании оценок максимального правдоподобия (ОМП) статистика распределена как сумма независимых слагаемых $\chi^2_{k-m-1} + \sum_{j=1}^{m} \lambda_j \xi_j^2$, где $\xi_1, ..., \xi_m$ – стандартные нормальные случайные величины, независимые одна от другой и от $\chi^2_{k-m-1}$. Однако точный закон распределения статистики неизвестен, и реально найти его можно только при помощи статистического моделирования [35]. Мощность критерия Пирсона зависит от способа группирования и выбранного числа интервалов [27].

Статистика $Y_n^2$ критерия Никулина-Рао-Робсона при проверке сложной гипотезы подчиняется $\chi^2_{k-1}$-распределению с $k-1$ степенями свободы. Подразумевается, что используются ОМП по негруппированной выборке. Мощность критерия Никулина-Рао-Робсона также зависит от способа группирования и выбранного числа интервалов [27].

**Проверка принадлежности ошибок измерений в экспериментах нормальным законам.** В таблице 5 для каждого из анализируемых наборов данных представлены результаты проверки гипотезы о принадлежности выборок нормальному закону по всем применяемым критериям. В таблице приведены значения статистик $S^*$ соответствующих критериев и достигнутый уровень значимости $P = P\{S > S^* | H_0\}$ (p-value). Проверяемая гипотеза $H_0$ по соответствующему критерию не отклоняется, если $P\{S > S^* | H_0\}$ больше заданного уровня значимости $\alpha$.

Таблица 5. Результаты проверки гипотезы о нормальности наблюдений

| Критерий | Кавендиш | | Милликен | | Майкельсон | | Ньюкомб | |
|---|---|---|---|---|---|---|---|---|
| | $S^*$ | $P$ | $S^*$ | $P$ | $S^*$ | $P$ | $S^*$ | $P$ |
| Эппса-Палли | 0.0519 | 0.734 | 0.0296 | 0.872 | 0.0742 | 0.607 | 0.1074 | 0.461 |
| Шапиро-Уилка | 0.9785 | 0.821 | --- | --- | --- | --- | --- | --- |
| Шапиро-Уилка (Royston) | –0.8705 | 0.808 | –0.3686 | 0.644 | –0.0344 | 0.514 | –0.2747 | 0.608 |
| Д’Агостино $z_2$ | 0.5505 | 0.582 | 0.4088 | 0.683 | 1.1359 | 0.256 | 0.6603 | 0.509 |
| Фросини | 0.1224 | 0.905 | 0.1156 | 0.948 | 0.2194 | 0.239 | 0.1906 | 0.408 |
| Хегази-Грина $T_1$ | 0.1072 | 0.871 | 0.0772 | 0.880 | 0.0922 | 0.226 | 0.0977 | 0.450 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Хегази-Грина $T_2$ | | 0.0305 | 0.562 | 0.0169 | 0.586 | 0.0170 | 0.236 | 0.0178 | 0.476 |
| Гири | | 0.8008 | 0.874 | 0.7977 | 0.864 | 0.7790 | 0.320 | 0.7745 | 0.309 |
| Дэвида-Хартли-Пирсона | | 4.3902 | 0.473 | 4.5776 | 0.979 | 5.6954 | 0.190 | 4.7212 | 0.902 |
| Шпигельхальтера | | 1.2272 | 0.633 | 1.2427 | 0.955 | 1.2773 | 0.325 | 1.2810 | 0.168 |
| $X_n^2$ Пирсона (ОМП по негруппированным данным) | $k$=4 | 0.3816 | 0.750 | 0.1782 | 0.927 | 1.2307 | 0.436 | 1.0262 | 0.488 |
| | $k$=5 | 0.8403 | 0.799 | 1.3395 | 0.619 | 2.4182 | 0.370 | 0.0708 | 0.990 |
| | $k$=6 | 1.9787 | 0.897 | 1.6741 | 0.718 | 5.4280 | 0.165 | 0.8710 | 0.902 |
| | $k$=7 | 1.7509 | 0.998 | 6.6350 | 0.163 | 5.9124 | 0.225 | 2.4062 | 0.719 |
| $X_n^2$ Пирсона (ОМП по группированным данным) | $k$=4 | 0.1156 | 0.734 | 1.4201 | 0.233 | 0.6060 | 0.436 | 3.6822 | 0.055 |
| | $k$=5 | 0.1887 | 0.920 | 1.4246 | 0.491 | 3.1733 | 0.205 | 0.8555 | 0.652 |
| | $k$=6 | 0.6699 | 0.880 | 0.5701 | 0.903 | 2.1559 | 0.541 | 3.9529 | 0.267 |
| | $k$=7 | 0.6768 | 0.954 | 2.3917 | 0.664 | 2.3475 | 0.672 | 5.0096 | 0.286 |
| $Y_n^2$ Никулина | $k$=4 | 1.0883 | 0.780 | 0.9730 | 0.808 | 3.6631 | 0.300 | 2.5530 | 0.466 |
| | $k$=5 | 1.2032 | 0.878 | 1.5811 | 0.812 | 10.9737 | 0.027 | 0.2345 | 0.994 |
| | $k$=6 | 5.5553 | 0.352 | 1.9131 | 0.861 | 8.3858 | 0.136 | 1.9755 | 0.853 |
| | $k$=7 | 2.2440 | 0.896 | 6.8988 | 0.330 | 14.4108 | 0.025 | 4.8375 | 0.565 |
| Колмогорова | | 0.5154 | 0.789 | 0.4922 | 0.845 | 0.8443 | 0.092 | 0.7379 | 0.225 |
| Андерсона-Дарлинга | | 0.2029 | 0.879 | 0.2064 | 0.870 | 0.4578 | 0.267 | 0.3783 | 0.411 |
| Крамера-Мизеса-Смирнова | | 0.0254 | 0.913 | 0.0238 | 0.933 | 0.0762 | 0.231 | 0.0623 | 0.356 |

В стандарте [36] для вычисления статистики критерия Шапиро-Уилка таблицы коэффициентов представлены для объемов выборок не более 50, поэтому в данном случае значения $S^*$ и $P$ приведены только для данных Кавендиша ($n$=29). При вычислении статистик критериев типа $\chi^2$ использовалось асимптотически оптимальное группирование [35] при числе интервалов $k=\overline{4,7}$. При проверке сложных гипотез с использованием критериев согласия параметры нормального распределения оценивались методом максимального правдоподобия.

Представленные в таблице 5 значения достигаемых уровней значимости показывают, что нет оснований для отклонения гипотезы о нормальности ни по одному из критериев.

**Некоторые замечания о вычислении достигаемых уровней значимости**. В случае критериев согласия Колмогорова, Крамера-Мизеса-Смирнова и Андерсона-Дарлинга при таких ограниченных, как в данном случае, объемах выборок следует учитывать, что распределения статистик будут несколько отличаться от предельных распределений, модели которых представлены в [28, 29] или [37] для ситуации проверки сложных гипотез.

Предельное распределение статистики $X_n^2$ критерия Пирсона при проверке сложной гипотезы и оценивании параметров по негруппированным данным методом максимального правдоподобия неизвестно (отличается от $\chi^2_{k-m-1}$-распределения) и зависит от способа группирования [27]. А при небольших объемах выборок и равновероятных интервалах группирования распределение статистики обладает и заметной дискретностью.

Предельным распределением статистики $Y_n^2$ критерия Никулина-Рао-Робсона является $\chi^2_{k-1}$-распределение (критерий обладает свойством «свободы от распределения»). Однако при малых, как в данном случае, объемах выборок наблюдается отличие распределения статистики от предельного.

Для критериев нормальности Шапиро-Уилка, Эппса-Палли, Фросини, Хегази-Грина, Гири, Дэвида-Хартли-Пирсона и Шпигельхальтера характерна сильная зависимость распределений статистик от объема выборки. В то же время в литературных источниках отсутствуют сведения о законах распределения статистик и приводятся лишь таблицы процентных точек. Критерий Шапиро-Уилка – левосторонний, правосторонними критериями являются критерий Шапиро-Уилка для больших выборок (критерий Ройстона) [16], критерии Эппса-Палли, Фросини, оба критерия Хегази-Грина, критерий Шпигельхальтера. Двусторонними критериями являются критерии Гири и Дэвида-Хартли-Пирсона. Распределение статистики критерия Ройстона [16] в области больших значений статистик, на основании которых принимается решение о том, отклонять или нет проверяемую гипотезу, хорошо аппроксимируется стандартным нормальным законом в широком диапазоне объемов выборок. Однако в целом зависимость распределения статистики критерия от объемов выборок является заметной.

В связи с вышеуказанными замечаниями при определении достигаемых уровней значимости во всех случаях, когда распределение статистики применяемого критерия либо неизвестно, либо при данном объеме выборки могло отличаться от асимптотического или предельного, использовались распределения статистик критериев, полученные в результате статистического моделирования. При этом число испытаний составило величину $N$ =1 660 000, что позволяет определить достигаемый уровень значимости с точностью $\varepsilon < 10^{-3}$ при доверительной вероятности $\gamma = 0.99$.

**Конкурирующие законы, пригодные для описания результатов рассматриваемых экспериментов**. Естественно проверить, какие модели законов, кроме нормального, позволяют достаточно хорошо описывать результаты анализируемых экспериментов. Для каждого набора данных с использованием критериев согласия Колмогорова, Крамера-Мизеса-Смирнова, Андерсона-Дарлинга и типа $\chi^2$ Никулина из 38 законов распределения вероятностей, реализованных в программной системе ISW [38], были отобраны законы, относительно которых гипотеза о принадлежности к ним соответствующих выборок не отклонялась и критерии показали наибольший средний достигаемый уровень значимости. При проверке сложных гипотез параметры законов оценивались методом максимального правдоподобия.

Для наблюдений Кавендиша с объемом выборки $n = 29$ наилучшими моделями оказались нормальный закон с плотностью $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ и параметрами $\mu = 5.4479$ и $\sigma = 0.2171$, распределение Лапласа $f(x) = \frac{1}{2\theta_1} e^{-|x-\theta_0|/\theta_1}$ с параметрами сдвига $\theta_0 = 5.46$ и масштаба $\theta_1 = 0.1737$, логарифмически нормальный закон $f(x) = \frac{1}{x\theta_1\sqrt{2\pi}} e^{-(\ln x - \theta_0)^2/2\theta_1^2}$ при $\theta_0 = 1.6944$ и $\theta_1 = 0.0403$ и логистический $f(x) = \frac{1}{\theta_1} \exp\left\{-\frac{(x-\theta_0)}{\theta_1}\right\} \Big/ \left[1 + \exp\left\{-\frac{(x-\theta_0)}{\theta_1}\right\}\right]^2$ при $\theta_0 = 5.456$, $\theta_1 = 0.1233$ (см. таблицу 6).

Далее обозначим конкурирующие гипотезы, соответствующие этим законам, следующим образом: нормальному – $H_0^{29}$; Лапласа – $H_1^{29}$; логарифмически нормальному – $H_2^{29}$; логистическому – $H_3^{29}$.

Таблица 6. Проверка согласия с отобранными моделями для данных Кавендиша

| Критерий | Нормальный | | Лапласа | | Лог-нормальный | | Логистический | |
|---|---|---|---|---|---|---|---|---|
| | $S^*$ | $P$ | $S^*$ | $P$ | $S^*$ | $P$ | $S^*$ | $P$ |
| $Y_n^2$ Никулина | 1.203 | 0.878 | 2.791 | 0.593 | 1.203 | 0.878 | 3.037 | 0.552 |
| Колмогорова | 0.515 | 0.789 | 0.560 | 0.700 | 0.538 | 0.730 | 0.386 | 0.974 |
| Андерсона-Дарлинга | 0.203 | 0.879 | 0.310 | 0.687 | 0.244 | 0.769 | 0.182 | 0.908 |
| Крамера-Мизеса-Смирнова | 0.025 | 0.913 | 0.050 | 0.556 | 0.030 | 0.837 | 0.024 | 0.882 |

Для данных Милликена с объемом выборки $n=58$ среди отобранных моделей оказались нормальный закон с оценками параметров $(\mu = 4.7808, \sigma = 0.0152)$, распределение Лапласа $(\theta_0 = 4.7813, \theta_1 = 0.0125)$ и распределение семейства

$$f(x) = \frac{\theta_2}{2\theta_1\Gamma(1/\theta_2)} \exp\left\{ -\left( \frac{|x-\theta_0|}{\theta_1} \right)^{\theta_2} \right\} \tag{1}$$

(двустороннее экспоненциальное – ДЭ) с параметрами $\theta_0 = 4.7808$, $\theta_1 = 0.0214$, $\theta_2 = 1.9903$ (см. таблицу 7). Обозначим соответствующие гипотезы: $H_0^{58}$ – для нормального закона; $H_1^{58}$ – для распределения Лапласа, $H_2^{58}$ – для распределения семейства (1).

Таблица 7. Проверка согласия с отобранными моделями для данных Милликена

| Критерий | Нормальный | | Лапласа | | ДЭ | |
|---|---|---|---|---|---|---|
| | $S^*$ | $P$ | $S^*$ | $P$ | $S^*$ | $P$ |
| $Y_n^2$ Никулина | 1.581 | 0.812 | 4.450 | 0.349 | 3.406 | 0.307 |
| Колмогорова | 0.492 | 0.845 | 0.690 | 0.361 | 0.494 | 0.806 |
| Андерсона-Дарлинга | 0.206 | 0.870 | 0.392 | 0.508 | 0.205 | 0.828 |
| Крамера-Мизеса-Смирнова | 0.024 | 0.933 | 0.062 | 0.408 | 0.024 | 0.911 |

Для наблюдений Майкельсона с объемом выборки $n=100$ наилучшими моделями оказались нормальный закон $(\mu = 852.3992, \sigma = 78.6145)$, логистический $(\theta_0 = 851.4711, \theta_1 = 44.3633)$ и логарифмически нормальный $(\theta_0 = 6.7437,\ \theta_1 = 0.0935)$ (таблица 8).

Таблица 8. Проверка согласия с отобранными моделями для данных Майкельсона

| Критерий | Нормальный | | Логистический | | Лог-нормальный | |
|---|---|---|---|---|---|---|
| | $S^*$ | $P$ | $S^*$ | $P$ | $S^*$ | $P$ |
| $Y_n^2$ Никулина | 10.973 | 0.027 | 7.673 | 0.104 | 3.619 | **0.460** |
| Колмогорова | 0.844 | 0.092 | 0.697 | 0.169 | 0.710 | **0.277** |
| Андерсона-Дарлинга | 0.458 | 0.267 | 0.434 | 0.215 | 0.446 | **0.285** |
| Крамера-Мизеса-Смирнова | 0.076 | 0.231 | 0.065 | 0.193 | 0.068 | **0.300** |

В данном случае по достигнутым уровням значимости можно говорить о предпочтительности модели логарифмически нормального закона. Обозначим соответствующие гипотезы: $H_0^{100}$ – для нормального закона; $H_1^{100}$ – для логистического; $H_2^{100}$ – для логарифмически нормального.

Наилучшей моделью для полной выборки результатов Ньюкомба оказалось распределение Коши $f(x) = \dfrac{\theta_1}{\pi\left[\theta_1^2 + (x-\theta_0)^2\right]}$ с ОМП параметров (27.2844, 2.4937). Соответствующую гипотезу обозначим $H_1^{66}$. Неробастность ОМП и влияние выбросов (–44 и –2) приводит к построению модели нормального закона с параметрами (26.2121, 10.6366), который не согласуется с эмпирическим распределением (см. рис. 1 и таблицу 9). Гипотезу, соответствующую нормальному закону обозначим $H_0^{66}$. Следует отметить, что Ньюкомб отнёс к аномальным наблюдениям значение –44 и не учитывал его при подведении итогов.

Таблица 9. Проверка согласия с отобранными моделями для данных Ньюкомба

| Критерий | Нормальный | | Коши | |
|---|---|---|---|---|
| | $S^*$ | $P$ | $S^*$ | $P$ |
| $Y_n^2$ Никулина | 71.436 | 0.000 | 6.376 | 0.173 |
| Колмогорова | 1.890 | 0.000 | 0.535 | 0.653 |
| Андерсона-Дарлинга | 5.819 | 0.000 | 0.550 | 0.312 |
| Крамера-Мизеса-Смирнова | 1.003 | 0.000 | 0.058 | 0.410 |

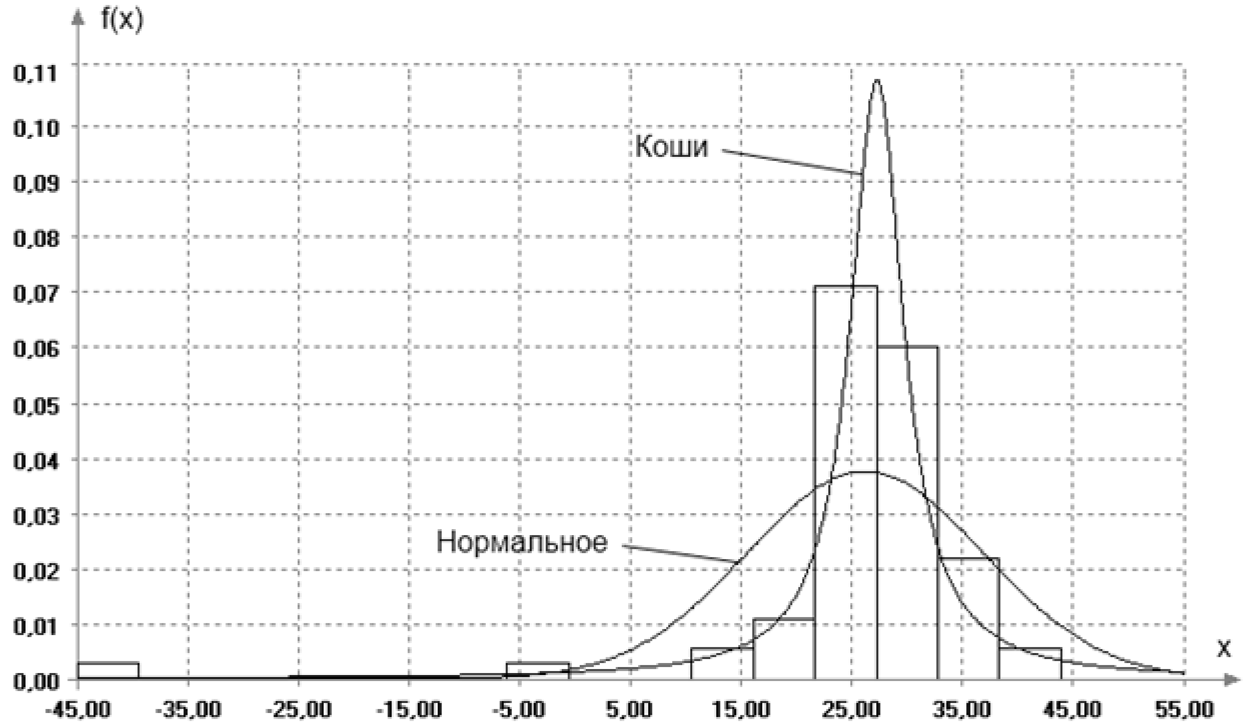


Рис. 1. Гистограмма для выборки Ньюкомба и плотности распределений нормального и Коши с соответствующими ОМП параметров ($\mu = 26.2121,\ \sigma = 10.6636$ и $\theta_0 = 27.2844,\ \theta_1 = 2.9437$).

Применение робастных оценок по группированным данным [39] или оценок с цензурированием выпадающих значений и последующее использование параметрического метода отбраковки приводит к исключению в качестве аномальных обоих значений (–44 и –2). В частности, если провести левостороннее цензурирование I-го типа с точкой цензурирования $C = 15$ и оценить параметры нормального распределения по такой выборке, то полученные ОМП параметров будут равны $\hat{\mu} = 27.3043$ и $\hat{\sigma} = 5.5785$. Для закона с такими параметрами значения –44 и –2 будут отнесены к выбросам.

Проверить адекватность модели, полученной по цензурированной выборке, можно с использованием рандомизации цензурированных наблюдений и преобразования Смирнова [32]. В данном случае в исходной выборке цезурированные наблюдения (значения –44 и –2) заменяются следующим образом. На интервале [0, $\Phi(C;\hat{\mu},\hat{\sigma})$], где $\Phi(C;\hat{\mu},\hat{\sigma})$ – функция распределения нормального закона в точке $C$, по равномерному закону будут сгенерированы псевдослучайные величины $\eta_i$, $i=\overline{1,2}$, которые преобразуются в соответствии с выражением $x_i = \Phi^{-1}(\eta_i, \hat{\mu}, \hat{\sigma})$ и добавляются к остальным 64 элементам выборки. Далее при проверке согласия полученной таким образом выборки с нормальным законом используются модели распределений статистик, применяемые при проверке соответствующих сложных гипотез [27, 28, 29, 37]. Достигаемые уровни значимости по критериям $\chi^2$ Пирсона, Колмогорова, Андерсона-Дарлинга и Крамера-Мизеса-Смирнова в данном случае оказались равными, соответственно, 0.302, 0.367, 0.287, 0.277. Таким образом, для отклонения гипотезы о нормальности оснований нет.

Для редуцированной выборки (без выбросов) лучшими моделями оказались нормальное распределение с параметрами $(\mu = 27.75, \sigma = 5.0436)$, распределение Лапласа $(\theta_0 = 27.75, \theta_1 = 3.0963)$ и логарифмически нормальное распределение $(\theta_0 = 3.3061, \theta_1 = 0.1876)$ (см. таблицу 10). Обозначим соответствующие гипотезы: $H_0^{64}$ – для нормального закона; $H_1^{64}$ – для закона Лапласа; $H_2^{64}$ – для логарифмически нормального.

Таблица 10. Проверка согласия с моделями для данных Ньюкомба (без выбросов)

| Критерий | Нормальный | | Лапласа | | Лог-нормальный | |
|---|---|---|---|---|---|---|
| | $S^*$ | $P$ | $S^*$ | $P$ | $S^*$ | $P$ |
| $Y_n^2$ Никулина | 0.234 | 0.994 | 2.593 | 0.628 | 7.041 | 0.134 |
| Колмогорова | 0.738 | 0.225 | 0.719 | 0.299 | 0.625 | 0.482 |
| Андерсона-Дарлинга | 0.378 | 0.411 | 0.424 | 0.447 | 0.398 | 0.370 |
| Крамера-Мизеса-Смирнова | 0.062 | 0.356 | 0.073 | 0.304 | 0.056 | 0.425 |

**Сравнительный анализ мощности критериев.** Мощность критериев проверки нормальности исследовалась при объемах выборок, соответствующих объемам в настоящих экспериментах, относительно тех конкурирующих законов, использование которых оказалось возможным для описания соответствующих выборок (относительно законов, являющихся достаточно близкими к нормальному). Оценки мощности получены на основании результатов моделирования условных распределений статистик критериев при справедливости соответствующих конкурирующих гипотез. Число испытаний составило N=1 660 000 в каждом случае.

Значения мощности критериев при проверке гипотезы $H_0^{29}$ относительно конкурирующих гипотез $H_1^{29}$, $H_2^{29}$, $H_3^{29}$ при объеме выборки $n = 29$ и задаваемых вероятностях ошибок первого рода $\alpha = 0.15$, $\alpha = 0.05$, $\alpha = 0.01$ приведены в таблице 11.

Таблица 11. Мощность критериев при проверке гипотезы $H_0^{29}$ относительно конкурирующих гипотез $H_1^{29}, H_2^{29}, H_3^{29}$ (для измерений Кавендиша)

| Критерий | | $H_1^{29}$ - Лапласа | | | $H_2^{29}$ - лог.нормальный | | | $H_3^{29}$ - логистический | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | **0.15** | **0.05** | **0.01** | **0.15** | **0.05** | **0.01** | **0.15** | **0.05** | **0.01** |
| Эппса-Палли | | 0.516 | .344 | .181 | .158 | .055 | .012 | .262 | .129 | .046 |
| Шапиро-Уилка | | .480 | .328 | .186 | .158 | .056 | .012 | .255 | .134 | .056 |
| Шапиро-Уилка (Royston) | | .506 | .348 | .196 | **.159** | **.056** | **.012** | .269 | .143 | .060 |
| Д'Агостино $z_2$ | | .408 | .243 | .113 | .150 | .050 | .010 | .235 | .106 | .034 |
| Фросини | | .505 | .330 | .172 | .156 | .053 | .011 | .232 | .105 | .034 |
| Хегази-Грина $T_1$ | | .558 | .380 | .204 | .157 | .054 | .011 | .273 | .133 | .047 |
| Хегази-Грина $T_2$ | | .622 | **.437** | **.243** | .158 | .055 | .012 | **.342** | **.185** | **.074** |
| Гири | | .571 | .397 | .221 | .151 | .051 | .010 | .264 | .129 | .048 |
| Дэвида-Хартли-Пирсона | | .438 | .272 | .131 | .150 | .050 | .010 | .255 | .123 | .042 |
| Шпигельхальтера | | **.625** | .425 | .227 | .152 | .051 | .010 | .299 | .143 | .050 |
| $X_n^2$ Пирсона (ОМП по негруппированным данным) | $k$=4 | .234 | .103 | .034 | **.155** | **.052** | **.011** | .162 | .056 | .013 |
| | $k$=5 | .318 | .154 | .045 | .153 | .052 | .010 | .172 | .063 | .013 |
| | $k$=6 | .312 | .146 | .052 | .153 | .051 | .011 | .179 | .065 | .016 |
| | $k$=7 | **.368** | **.196** | **.064** | .153 | .052 | .011 | **.198** | **.075** | **.017** |
| $X_n^2$ Пирсона (ОМП по группированным данным) | $k$=4 | .226 | .098 | .028 | .167 | .059 | .013 | .172 | .062 | .014 |
| | $k$=5 | .357 | .181 | .059 | .170 | .059 | .012 | .189 | .070 | .015 |
| | $k$=6 | **.366** | **.208** | **.091** | **.180** | **.064** | **.014** | **.220** | **.098** | **.031** |
| | $k$=7 | .352 | .180 | .063 | .173 | .060 | .012 | .206 | .087 | .024 |
| $Y_n^2$ Никулина | $k$=4 | .262 | .111 | .044 | .153 | .052 | .010 | .174 | .059 | .015 |
| | $k$=5 | .356 | .167 | .066 | .153 | .051 | .010 | .197 | .071 | .020 |
| | $k$=6 | .342 | .181 | .075 | .153 | .052 | .011 | .200 | .081 | .023 |
| | $k$=7 | **.399** | **.223** | **.097** | **.153** | **.052** | **.011** | **.215** | **.088** | **.026** |
| Колмогорова | | .460 | .272 | .118 | .155 | .053 | .011 | .216 | .089 | .024 |
| Андерсона-Дарлинга | | **.517** | **.342** | **.180** | **.157** | **.054** | **.011** | **.247** | **.115** | **.039** |
| Крамера-Мизеса-Смирнова | | .507 | .330 | .170 | .155 | .053 | .011 | .231 | .103 | .033 |

Значения мощности критериев при проверке гипотезы $H_0^{58}$ относительно конкурирующей гипотезы $H_1^{58}$ при объеме выборки $n = 58$ представлены в таблице 12. Оценки мощности относительно $H_2^{58}$ не приводятся, так как в этом случае распределение (1) практически не отличается от нормального и оценки мощности с точностью до погрешности моделирования совпадают задаваемыми уровнями значимости.

Таблица 12. Мощность критериев при проверке гипотезы $H_0^{58}$ относительно конкурирующей гипотезы $H_1^{58}$ (для измерений Милликена)

| Критерий | | $H_1^{58}$ - Лапласа | | |
|---|---|---|---|---|
| | | **0.15** | **0.05** | **0.01** |
| Эппса-Палли | | 0.740 | .574 | .365 |
| Шапиро-Уилка | | --- | --- | --- |
| Шапиро-Уилка (Royston) | | .727 | .577 | .391 |
| Д'Агостино $z_2$ | | .610 | .435 | .250 |
| Фросини | | .742 | .578 | .375 |
| Хегази-Грина $T_1$ | | .783 | .631 | .427 |
| Хегази-Грина $T_2$ | | .823 | .675 | .467 |
| Гири | | .827 | .694 | .501 |
| Дэвида-Хартли-Пирсона | | .606 | .434 | .254 |
| Шпигельхальтера | | **.890** | **.761** | **.553** |
| $X_n^2$ Пирсона (ОМП по негруппированным данным) | $k$=4 | .281 | .126 | .041 |
| | $k$=5 | .528 | .309 | .136 |
| | $k$=6 | .437 | .248 | .095 |
| | $k$=7 | **.558** | **.362** | **.171** |
| $X_n^2$ Пирсона (ОМП по группированным данным) | $k$=4 | .248 | .112 | .036 |
| | $k$=5 | **.562** | **.362** | **.169** |
| | $k$=6 | .518 | .329 | .154 |
| | $k$=7 | .551 | .340 | .148 |
| $Y_n^2$ Никулина | $k$=4 | .322 | .176 | .068 |
| | $k$=5 | .555 | .372 | .205 |
| | $k$=6 | .490 | .309 | .156 |
| | $k$=7 | **.605** | **.427** | **.250** |
| Колмогорова | | .679 | .474 | .249 |
| Андерсона-Дарлинга | | .755 | .595 | .389 |
| Крамера-Мизеса-Смирнова | | .750 | .585 | .375 |

Таблица 13. Мощность критериев при проверке гипотезы $H_0^{66}$ относительно конкурирующей гипотезы $H_1^{66}$ (для измерений Ньюкомба)

| Критерий | | $H_1^{66}$ - Коши | | |
|---|---|---|---|---|
| | | **0.15** | **0.05** | **0.01** |
| Эппса-Палли | | 0.999 | .999 | 0.999 |
| Шапиро-Уилка | | --- | --- | --- |
| Шапиро-Уилка (Royston) | | .999 | .999 | .999 |
| Д'Агостино $z_2$ | | .948 | .927 | .900 |
| Фросини | | .999 | .999 | .999 |
| Хегази-Грина $T_1$ | | .999 | .999 | .999 |
| Хегази-Грина $T_2$ | | .999 | .999 | .999 |
| Гири | | .999 | .999 | .999 |
| Дэвида-Хартли-Пирсона | | .996 | .990 | .974 |
| Шпигельхальтера | | .999 | .999 | .999 |
| $X_n^2$ Пирсона (ОМП по негруппированным данным) | $k$=4 | .953 | .899 | .804 |
| | $k$=5 | .999 | .997 | .990 |
| | $k$=6 | .995 | .985 | .959 |
| | $k$=7 | .999 | .997 | .993 |
| $X_n^2$ Пирсона (ОМП по группированным данным) | $k$=4 | .526 | .389 | .257 |
| | $k$=5 | .995 | .987 | .967 |
| | $k$=6 | .996 | .991 | .979 |
| | $k$=7 | .997 | .993 | .981 |
| $Y_n^2$ Никулина | $k$=4 | .971 | .948 | .909 |
| | $k$=5 | .999 | .998 | .995 |
| | $k$=6 | .997 | .993 | .983 |
| | $k$=7 | .999 | .998 | .996 |
| Колмогорова | | .999 | .999 | .996 |
| Андерсона-Дарлинга | | .999 | .999 | .999 |
| Крамера-Мизеса-Смирнова | | .999 | .999 | .999 |

Оценки мощности критериев при проверке гипотезы $H_0^{66}$ относительно конкурирующей гипотезы $H_1^{66}$ при объеме выборки $n = 66$ приведены в таблице 13, при проверке гипотезы $H_0^{64}$ относительно конкурирующих гипотез $H_1^{64}$, $H_2^{64}$ при объеме выборки $n = 64$ – в таблице 14, при проверке гипотезы $H_0^{100}$ относительно конкурирующих гипотез $H_1^{100}$ и $H_2^{100}$ при объеме выборки $n = 100$ – в таблице 15.

Таблица 14. Мощность критериев при проверке гипотезы $H_0^{64}$ относительно конкурирующих гипотез $H_1^{64}, H_2^{64}$ (для измерений Ньюкомба без выбросов)

| Критерий | | $H_1^{64}$ - Лапласа | | | $H_2^{64}$ - лог.нормальный | | |
|---|---|---|---|---|---|---|---|
| | | **0.15** | **0.05** | **0.01** | **0.15** | **0.05** | **0.01** |
| Эппса-Палли | | 0.774 | .615 | .404 | .501 | .309 | .141 |
| Шапиро-Уилка | | --- | --- | --- | --- | --- | --- |
| Шапиро-Уилка (Royston) | | .760 | .617 | .431 | **.526** | **.342** | **.172** |
| Д’Агостино $z_2$ | | .643 | .470 | .278 | .224 | .102 | .036 |
| Фросини | | .777 | .622 | .419 | .420 | .237 | .098 |
| Хегази-Грина $T_1$ | | .814 | .672 | .471 | .467 | .282 | .128 |
| Хегази-Грина $T_2$ | | .848 | .712 | .507 | .491 | .307 | .146 |
| Гири | | .858 | .737 | .552 | .203 | .086 | .027 |
| Дэвида-Хартли-Пирсона | | .632 | .458 | .274 | .178 | .069 | .018 |
| Шпигельхальтера | | **.914** | **.803** | **.610** | .232 | .104 | .034 |
| $X_n^2$ Пирсона (ОМП по негруппированным данным) | $k$=4 | .283 | .134 | .041 | **.360** | **.191** | **.066** |
| | $k$=5 | .550 | .332 | .147 | .347 | .156 | .040 |
| | $k$=6 | .459 | .260 | .106 | .330 | .145 | .046 |
| | $k$=7 | **.589** | **.384** | **.195** | .313 | .149 | .049 |
| $X_n^2$ Пирсона (ОМП по группированным данным) | $k$=4 | .232 | .103 | .032 | .418 | .225 | .081 |
| | $k$=5 | .598 | **.404** | **.206** | .441 | .239 | .087 |
| | $k$=6 | .563 | .379 | .204 | **.492** | **.282** | **.115** |
| | $k$=7 | **.600** | .401 | .204 | .441 | .235 | .086 |
| $Y_n^2$ Никулина | $k$=4 | .328 | .170 | .067 | .323 | .150 | .049 |
| | $k$=5 | .583 | .402 | .219 | **.327** | .143 | .041 |
| | $k$=6 | .506 | .324 | .166 | .319 | .151 | .050 |
| | $k$=7 | **.631** | **.452** | **.271** | .318 | **.157** | **.056** |
| Колмогорова | | .705 | .505 | .273 | .373 | .191 | .067 |
| Андерсона-Дарлинга | | **.781** | **.628** | **.423** | **.451** | **.263** | **.113** |
| Крамера-Мизеса-Смирнова | | .775 | .617 | .410 | .414 | .231 | .094 |

Таблица 15. Мощность критериев при проверке гипотезы $H_0^{100}$ относительно конкурирующих гипотез $H_1^{100}, H_2^{100}$ (для измерений Майкельсона)

| Критерий | | $H_1^{100}$ - логистический | | | $H_2^{100}$ - лог.нормальный | | |
|---|---|---|---|---|---|---|---|
| | | **0.15** | **0.05** | **0.01** | **0.15** | **0.05** | **0.01** |
| Эппса-Палли | | 0.429 | .250 | .108 | .302 | .148 | .050 |
| Шапиро-Уилка | | --- | --- | --- | --- | --- | --- |
| Шапиро-Уилка (Royston) | | .461 | .305 | .167 | **.318** | **.165** | **.061** |
| Д'Агостино $z_2$ | | .459 | .286 | .140 | .167 | .062 | .015 |
| Фросини | | .374 | .205 | .083 | .259 | .117 | .035 |
| Хегази-Грина $T_1$ | | .456 | .276 | .127 | .283 | .135 | .045 |
| Хегази-Грина $T_2$ | | .582 | .398 | .218 | .301 | .150 | .052 |
| Гири | | .486 | .313 | .161 | .165 | .059 | .014 |
| Дэвида-Хартли-Пирсона | | .428 | .264 | .129 | .160 | .056 | .012 |
| Шпигельхальтера | | **.613** | **.408** | **.212** | .177 | .066 | .016 |
| $X_n^2$ Пирсона (ОМП по негруппированным данным) | $k$=4 | .192 | .067 | .014 | **.230** | **.099** | **.028** |
| | $k$=5 | .259 | .111 | .030 | .226 | .092 | .023 |
| | $k$=6 | .244 | .103 | .028 | .227 | .089 | .022 |
| | $k$=7 | **.270** | **.121** | **.036** | .219 | .087 | .023 |
| $X_n^2$ Пирсона (ОМП по группированным данным) | $k$=4 | .175 | .065 | .015 | .276 | .125 | .036 |
| | $k$=5 | .262 | .117 | .034 | .274 | .120 | .034 |
| | $k$=6 | .302 | .150 | .054 | .307 | .142 | .043 |
| | $k$=7 | **.338** | **.183** | **.078** | **.307** | **.143** | **.047** |
| $Y_n^2$ Никулина | $k$=4 | .219 | .092 | .027 | .208 | .080 | .020 |
| | $k$=5 | .304 | .153 | .061 | .213 | .083 | .020 |
| | $k$=6 | .295 | .147 | .057 | .217 | .085 | .022 |
| | $k$=7 | **.327** | **.173** | **.072** | **.217** | **.086** | **.023** |
| Колмогорова | | .316 | .148 | .046 | .242 | .102 | .028 |
| Андерсона-Дарлинга | | **.402** | **.227** | **.095** | **.274** | **.127** | **.040** |
| Крамера-Мизеса-Смирнова | | .368 | .197 | .076 | .258 | .115 | .035 |

В таблицах 11-15 полужирным шрифтом выделены максимальные значения мощностей для соответствующих групп критериев.

Конкурирующие модели законов для данных Кавендиша можно упорядочить по близости к нормальному закону следующим образом: наиболее далеким законом является распределение Лапласа ($H_1^{29}$); ближе – логистический закон ($H_3^{29}$); наиболее близкий – логарифмически нормальный ($H_2^{29}$) (см. рис. 2).

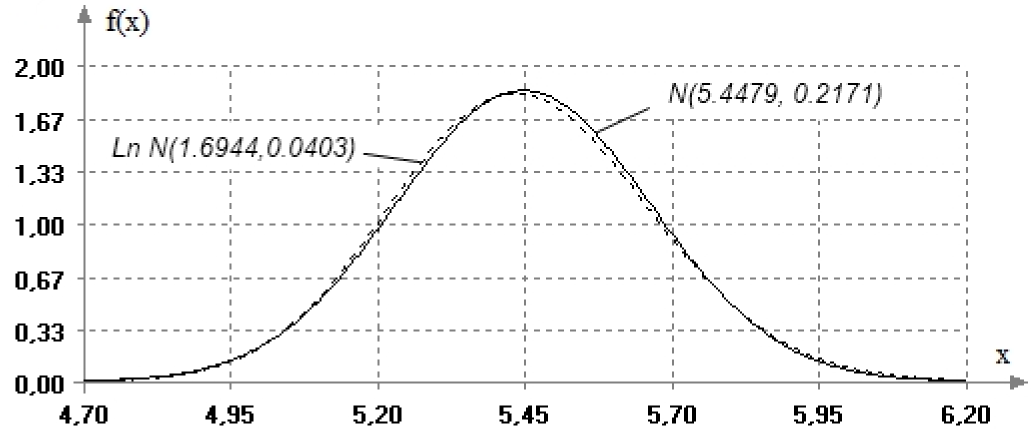


Рис. 2. Плотности распределений нормального и логарифмически нормального законов, соответствующие гипотезам $H_0^{29}$ и $H_2^{29}$.

То же самое можно сказать относительно других экспериментов, когда эти законы оказываются среди моделей, пригодных для описания результатов измерений.

**Заключение**

Подводя итоги, можно констатировать, что нет оснований для отклонения гипотезы о нормальности ошибок измерений в рассмотренных экспериментах. С другой стороны, очевидно, что для описания ошибок измерений в данных экспериментах можно подобрать и другие параметрические модели законов, не менее хорошо согласующиеся с результатами измерений.

Например, при использовании для описания распределений семейства (1), частным случаем которого является нормальный закон, всегда можно построить модель, которая в определенном смысле (в зависимости от меры близости и метода оценивания) оказывается ближе к эмпирическому распределению, чем нормальный закон. Однако ответ о наиболее предпочтительной модели, опирающийся на результаты проверки с использованием критериев согласия, необязательно будет в пользу модели семейства (1). В частности, для описания данных Кавендиша получаем распределение вида (1) с ОМП параметров $\theta_0 = 5.4486$, $\theta_1 = 0.3024$, $\theta_2 = 1.9440$, для данных Милликена – $\theta_0 = 4.7808$, $\theta_1 = 0.0214$, $\theta_2 = 1.9903$, для данных Майкельсона – $\theta_0 = 851.315$,

$\theta_1 = 100.047$, $\theta_2 = 1.6801$, для данных Ньюкомба – $\theta_0 = 27.652$, $\theta_1 = 6.2432$, $\theta_2 = 1.6160$. Результаты проверки согласия с этими законами с использованием непараметрических критериев согласия приведены в таблице 16. Их можно сравнить с результатами проверки нормальности по этим критериям (со значениями статистик и достигаемыми уровнями значимости), приводимыми в таблице 5.

Таблица 16. Проверка согласия данных с распределением семейства (1)

| Критерий | Кавендиш | | Милликен | | Майкельсон | | Ньюкомб | |
|---|---|---|---|---|---|---|---|---|
| | $S^*$ | $P$ | $S^*$ | $P$ | $S^*$ | $P$ | $S^*$ | $P$ |
| Колмогорова | 0.501 | 0.765 | 0.494 | 0.806 | 0.682 | 0.191 | 0.648 | 0.268 |
| Андерсона-Дарлинга | 0.198 | 0.845 | 0.205 | 0.828 | 0.417 | 0.205 | 0.301 | 0.458 |
| Крамера-Мизеса-Смирнова | 0.025 | 0.882 | 0.024 | 0.911 | 0.063 | 0.189 | 0.045 | 0.400 |

Анализируя мощность рассмотренных критериев относительно трех конкурирующих законов, оказавшихся подходящими для описания измерений, можно однозначно утверждать следующее. Специальные критерии нормальности, предназначенные для проверки гипотезы только о принадлежности выборки нормальному закону, по крайней мере, лучшие их представители, имеют преимущество перед непараметрическими критериями согласия. В свою очередь, непараметрические критерии согласия при проверке нормальности обладают более высокой мощностью по сравнению с критериями типа $\chi^2$.

Специальные критерии нормальности можно упорядочить по мощности относительно конкурирующей гипотезы в виде распределения Лапласа следующим образом:

*Шпигельхальтера* $\succ$ *Гири* ~ *Хегази-Грина* $T_2$ $\succ$ *Хегази-Грина* $T_1$ $\succ$ *Фросини* ~ *Эппса-Палли* $\succ$ *Шапиро-Уилка* $\succ$ *Дэвида-Хартли-Пирсона*.

Относительно конкурирующей гипотезы в виде логистического распределения в следующем порядке:

*Шпигельхальтера* $\sim$ *Хегази-Грина* $T_2$ $\succ$ *Гири* $\sim$ *Хегази-Грина* $T_1$ $\succ$ *Эппса-Палли* $\succ$ *Шапиро-Уилка* $\succ$ *Дэвида-Хартли-Пирсона* $\succ$ *Фросини*.

Относительно конкурирующего логарифмически нормального закона критерии упорядочиваются следующим образом:

*Эппса-Палли* $\sim$ *Шапиро-Уилка* $\succ$ *Хегази-Грина* $T_2$ $\succ$ *Хегази-Грина* $T_1$ $\succ$ *Фросини* $\succ$ *Шпигельхальтера* $\succ$ *Гири* $\succ$ *Дэвида-Хартли-Пирсона*.

Если отталкиваться от полученных в данной работе оценок мощности критериев, то к наиболее предпочтительным следует отнести критерии Шпигельхальтера, Хегази-Грина $T_2$, Эппса-Палли, Шапиро-Уилка, Гири. Однако не следует забывать, что не для любой конкурирующей гипотезы критерии Шпигельхальтера, Хегази-Грина $T_2$, Эппса-Палли, Шапиро-Уилка имеют преимущество в мощности. Если рассматривать в качестве конкурирующего закона, например, распределение семейства (1) с параметром формы $\theta_2 = 4$, то при малых объемах выборок критерии Шапиро-Уилка, Эппса-Палли [7], Хегази-Грина $T_2$, Хегази-Грина $T_1$ [8] не могут отличить этот закон от нормального вследствие смещенности распределений статистик при справедливости конкурирующей гипотезы. А критерий Шпигельхальтера, статистика которого построена как комбинация статистик критериев Гири и Дэвида-Хартли-Пирсона, не способен этого сделать и с ростом объема выборок [8].

Критерии согласия типа $\chi^2$ (Пирсона, Никулина-Рао-Робсона) при проверке сложных гипотез уступают по мощности непараметрическим критериям. Это касается и проверки гипотез о нормальности [10, 11]. Использованное асимптотически оптимальное группирование максимизирует мощность критерия $\chi^2$ Пирсона относительно “очень близких” конкурирующих гипотез [27]. Для конкретной конкурирующей гипотезы за счет выбора оптимального числа интервалов и способа группирования можно построить критерий типа $\chi^2$ для проверки нормальности, который будет обладать максимальной мощностью. Такая задача не является тривиальной [40], но позволяет добиться мощности большей, чем у непараметрических критериев согласия [27]. Сказанное может быть также отнесено и к критерию Никулина: для случая конкретной альтернативы асимптотически оптимальное группирование не будет тем методом, который максимизирует его мощность [9]. А повсеместное использование на практике интервалов равной длины или равновероятного группирования является скорее выбором от неимения лучших вариантов.

Работа выполнена при поддержке РФФИ (проект № 09-01-00056а) и Федеральной целевой программы Минобрнауки РФ “Научные и научно-педагогические кадры инновационной России”.

**Список литературы**


1. Voinov V., Voinov E. A statistical reanalysis of the classical Rutherford’s experiment // Communications in Statistics – Simulation and Computation, 2010. – V.39. – № 1. – P.157–171.
2. Dong, Lauren Bin and Giles, David E. A. An Empirical Likelihood Ratio Test for Normality // Communications in Statistics - Simulation and Computation, 2007. – Vol.36. – № 1. – P.197 – 215.
3. Doornik J.A. and Hansen H. An Omnibus Test for Univariate and Multivariate Normality // Oxford Bulletin of Economics and Statistics, 2008. – Vol.70, – P.927-939.
4. Scott W.F., Stewart B. Tables for the Lilliefors and Modified Cramer–von Mises tests of normality // Communications in Statistics – Theory and Methods, 2011. – V.40. – № 4. – P.726-730.
5. Martynov G. Weighted Cramer-von Mises Test with Estimated Parameters // Communications in Statistics – Theory and Methods, 2011. – V.40. – № 19-20. – P.3569-3586.
6. Voinov V., Pya N., Alloyarova R. A comparative study of some modified chi-squared tests // Communications in Statistics – Simulation and Computation, 2009. – V.38. – № 3. – P.355–367.

7. Лемешко Б.Ю., Лемешко С.Б. Сравнительный анализ критериев проверки отклонения распределения от нормального закона // Метрология. 2005. – № 2. – С.3-24.
8. Лемешко Б.Ю., Рогожников А.П. Исследование особенностей и мощности некоторых критериев нормальности // Метрология. 2009. – № 4. – С.3-24.
9. Лемешко Б.Ю., Лемешко С.Б., Постовалов С.Н. Мощность критериев согласия при близких альтернативах // Измерительная техника. 2007. № 2. – С.22-27.
10. Лемешко Б.Ю., Лемешко С.Б., Постовалов С.Н. Сравнительный анализ мощности критериев согласия при близких конкурирующих гипотезах. I. Проверка простых гипотез // Сибирский журнал индустриальной математики. 2008. – Т.11. – № 2(34). – С.96-111.
11. Лемешко Б.Ю., Лемешко С.Б., Постовалов С.Н. Сравнительный анализ мощности критериев согласия при близких альтернативах. II. Проверка сложных гипотез // Сибирский журнал индустриальной математики. 2008. – Т.11. – № 4(36). – С.78-93.
12. Stigler S.M. Do robust estimators work with real data? // The Annals of Statistics, 1977. – V.5. № 6. – P.1055–1098.
13. Millikan R. A. On the elementary electrical charge and the Avogadro constant // The Physical Review, 1913. Series II 2. – P.109–143.
14. Shapiro S.S., Wilk M.B. An analysis of variance test for normality (complete samples) // Biometrika, 1965. – V.52 – P.591-611.
15. Shapiro S.S., Francia R.S. An approximate analysis of variance test for normality // J. Amer. Statist. Assoc., 1972. – V.337. – P.215-216.
16. Royston J. P. Approximating the Shapiro-Wilk W-test for non-normality // Statistics and Computing, 1992. – V. 2(3). – P.117-119.
17. Epps T.W., Pulley L.B. A test for normality based on the empirical characteristic function // Biometrika, 1983. – V.70. – P.723-726.

18. D’Agostino R.B. Transformation to normality of the null distribution of $g_1$ // Biometrika, V. 57, 1970. – P.679-681.
19. Frosini B.V. A survey of a class of goodness-of-fit statistics // Metron. – 1978. – V.36. – №1-2. – P.3-49.
20. Frosini B.V. On the distribution and power of goodness-of-fit statistic with parametric and nonparametric applications, “Goodness-of-fit” / Ed. by Revesz P., Sarkadi K., Sen P.K. – Amsterdam-Oxford-New York: North Holland Publ. Comp, 1987. – P.133-154.
21. Hegazy Y.A.S., Green J.R. Some new goodness-of-fit tests using order statistics //Applied Statistics. – 1975. – V.24. – №3. – P.299-308.
22. Geary R.C. The ratio of the mean deviation to the standard deviation as a test of normality // Biometrica. – 1935. – V.27. – P.310-322.
23. Geary R.C. Moments of the ratio of the mean deviation to the standard deviation for normal samples // Biometrika. 1936. – V.28. – P.295-307.
24. Geary R.C. Testing for normality // Biometrika. 1937. – V.34. – P.209-242.
25. David H.A., Hartley H.O., Pearson E.S. The distribution of the ratio, in a single normal sample, of range to standard deviation // Biometrika. 1964. – V.512. – №3-4. – P.484-487.
26. Spiegelhalter D.J. A test for normality against symmetric alternatives // Biometrika. 1977. – V.64. – № 2. – P.415-418.
27. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход : монография / Б.Ю. Лемешко, С.Б. Лемешко, С.Н. Постовалов, Е.В. Чимитова. – Новосибирск : Изд-во НГТУ, 2011. – 888 с.
28. Лемешко Б.Ю., Лемешко С.Б. Модели распределений статистик непараметрических критериев согласия при проверке сложных гипотез с использованием оценок максимального правдоподобия. Ч.I // Измерительная техника, 2009. – № 6. – С.3-11.
29. Лемешко Б.Ю., Лемешко С.Б. Модели распределений статистик непараметрических критериев согласия при проверке сложных гипотез с ис-

пользованием оценок максимального правдоподобия. Ч.II // Измерительная техника, 2009. – № 8. – С.17-26.

30. Никулин М. С. Критерий хи-квадрат для непрерывных распределений с параметрами сдвига и масштаба / М. С. Никулин // Теория вероятностей и ее применение. – 1973. – Т. XVIII, № 3. – С. 583–591.
31. Никулин М. С. О критерии хи-квадрат для непрерывных распределений / М. С. Никулин // Теория вероятностей и ее применение. – 1973. – Т. XVIII. – № 3. – С. 675–676.
32. Greenwood P. E. A guide to chi-squared testing / P. E. Greenwood, M. S. Nikulin. – New York : John Wiley & Sons, 1996. – 280 p.
33. Rao K. C. A chi-squared statistic for goodness-of-fit tests within the exponential family / K. C. Rao, D. S. Robson // Commun. Statist. – 1974. – Vol. 3. – P. 1139–1153.
34. Chernoff H. The use of maximum likelihood estimates in $\chi^2$ test for goodness of fit / H. Chernoff, E. L. Lehmann // Ann. Math. Stat. – 1954. – Vol. 25. – P. 579–586.
35. Р 50.1.033-2001. Рекомендации по стандартизации. Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Часть I. Критерии типа хи-квадрат. - М.: Изд-во стандартов. 2002. – 87 с.
36. ГОСТ Р ИСО 5479-2002. Статистические методы. Проверка отклонения распределения вероятностей от нормального распределения. – М.: Изд-во стандартов. 2002. – 30 с.
37. Lemeshko B.Yu., Lemeshko S.B. and Postovalov S.N. Statistic Distribution Models for Some Nonparametric Goodness-of-Fit Tests in Testing Composite Hypotheses // Communications in Statistics - Theory and Methods, 2010. – Vol. 39. – No. 3. – P. 460-471.
38. Лемешко Б.Ю., Постовалов С.Н. Компьютерные технологии анализа данных и исследования статистических закономерностей: Учебное пособие. – Новосибирск: Изд-во НГТУ, 2004. – 119 с.

39. Лемешко Б.Ю. Группирование наблюдений как способ получения робастных оценок // Надежность и контроль качества. – 1997. – № 5. – С.26-35.
40. Бушакова А.Д., Лемешко Б.Ю. Исследование влияния вариантов асимптотической оптимальности группирования на мощность критериев типа $\chi^2$ // Материалы Российской НТК "Информатика и проблемы телекоммуникаций", Новосибирск. 2009. – Т.1. – С.34-37.

UDC 519.24

# ON NORMALITY OF ERRORS IN MEASUREMENTS IN CLASSICAL EXPERIMENTS AND POWER OF TESTS FOR DEVIATION FROM THE NORMAL LAW

Boris Yu. Lemeshko, Andrey P. Rogozhnikov[2]

Application of a set of statistical tests that are used to test empirical distributions for deviation from the normal law is demonstrated by an analysis of results of classical experiments related to measuring physical constants. A comparative analysis of power of the set of tests with respect to acceptable competing hypotheses is conducted with sample sizes equal to those in the experiments under study.

**Keywords:** power of test, tests for normality, nonparametric goodness-of-fit tests, chi-squared tests, testing complex hypotheses

[2]Novosibirsk State Technical University