МАТЕМАТИКА

УДК 519.2

И. В. Золотухин, кандидат физико-математических наук, доцент, Санкт-Петербургский филиал Института океанологии имени П. П. Ширшова РАН
Л. А. Золотухина, доктор физико-математических наук, профессор, Санкт-Петербургский государственный морской технический университет

КРИТЕРИЙ ГИРИ В ЗАДАЧАХ, СВЯЗАННЫХ С РАСПРЕДЕЛЕНИЯМИ ГАУССА И ЛАПЛАСА

Исследуется распределение Лапласа как альтернатива распределению Гаусса и критерий Гири. Дано теоретическое обоснование механизма возникновения распределения Лапласа из гауссова распределения. Дан вывод критерия Гири на основе использования метода отношения правдоподобия. Доказаны инвариантность, состоятельность и асимптотическая нормальность критерия. Методом статистического моделирования составлены таблицы и приведены графики, подтверждающие высокую эффективность критерия.

Ключевые слова: распределение Гаусса, распределение Лапласа, критерий Гири, метод отношения правдоподобия, статистическое моделирование.

В последние годы распределение Лапласа нередко используется как альтернатива нормальному (гауссову) распределению. Круг прикладных задач, в которых применяют это распределение, весьма широк. Так, например, это распределение применяется в метрологии, радиоэлектронике, навигации (подробнее см. работу Бенинга и Королёва [1]). Применение распределения Лапласа вместо гауссова распределения основано на лучшем согласии с ним эмпирических данных в тех случаях, когда эксперименты проводятся в случайных условиях. Поэтому выбор из двух альтернатив – нормальному или лапласову распределению подчиняются эмпирические данные – является актуальной прикладной задачей. Такая задача решается с помощью теории проверки статистических гипотез.

Для проверки гипотез разработан целый ряд критериев. На практике чаще всего используются следующие критерии согласия: хи-квадрат, критерий Колмогорова, омега-квадрат (критерий Крамера – Мизеса – Смирнова) [2, 3]. Выбор наилучшего критерия проводят путем сравнения их функций мощности. Ясно, что применять целесообразно критерии, мощности которых достаточно быстро стремятся к 1 при объемах выборок, стремящихся к бесконечности.

В настоящей работе исследуются свойства критерия Гири в задачах статистической обработки экспериментальных данных, когда в качестве альтернативы нормальному распределению рассматривается распределение Лапласа, которое по сравнению с нормальным имеет более «тяжелые хвосты».

В работе показана достаточно высокая мощность критерия Гири уже на выборках сравнительно небольшого объема, что позволяет рекомендовать его для практического применения в подобного рода задачах. В начале исследования способом, отличным от известных [4–6], дается теоретическое обоснование механизма возникновения распределения Лапласа.

Затем рассматривается задача различения распределений Гаусса и Лапласа по выборочным данным. Для случая простых гипотез задача решается с помощью критерия отношения правдоподобия. Для случая сложных гипотез в работе показано, что критерий отношения правдоподобия при подстановке вместо неизвестных параметров их оценок сводится к критерию Гири. Показана инвариантность и исследованы асимптотические свойства критерия при обеих альтернативах. При этом:

- найдены асимптотические математические ожидания и дисперсии;
- показана вырожденность распределения статистики критерия;
- доказана состоятельность критерия;
- показана асимптотическая нормальность критерия.

Далее методом статистического моделирования с помощью программного пакета *Wolfram Mathematica 8* в широкой области изменения уровня значимости критерия Гири и объема выборок проведены расчеты:

- критических точек критерия для выборок из нормального распределения;
- мощности критерия для выборок из распределения Лапласа;
- оценок средних и среднеквадратических отклонений распределения статистики критерия для выборок из распределений Гаусса и Лапласа.

По результатам моделирования составлены соответствующие таблицы и построены иллюстративные графики функции мощности критерия, гистограммы распределений статистики Гири при обеих гипотезах, а также их аппроксимации нормальными зако-

Получено 22.08.13

нами. Полученные материалы позволяют легко решать различные практические задачи, связанные с альтернативной аппроксимацией эмпирических данных законами Гаусса и Лапласа.

Теоретическое обоснование механизма возникновения распределения Лапласа из нормального (гауссова) распределения

При обработке статистических данных обычно предполагают, что ошибки измерений распределены по нормальному закону. Однако тип распределения ошибок, используемого при обработке, зависит от механизма их возникновения. Если исходить из предположения, что известны среднее и дисперсия, а в остальном характеристики ошибок не определены, и искать плотность распределения ошибок, исходя из максимума энтропии, то получим, что ошибки распределены по нормальному закону [4, 5].

Но во многих реальных ситуациях точностные характеристики ошибок измерения известны лишь в среднем. В этом случае естественно предположить, что дисперсия ошибок σ^2 сама случайна, и известно лишь ее среднее значение $\overline{\sigma}^2$.

На основе концепции максимальной энтропии и в предположении, что σ^2 есть неотрицательная случайная величина с известным средним, получено [4, 5], что σ^2 – показательно распределенная случайная величина.

Рассмотрим следующую модель механизма возникновения ошибок. При фиксированном σ^2 ошибки измерения подчинены нормальному закону со средним μ; σ^2 – случайная величина, подчиненная показательному закону с параметром λ. Нетрудно показать, что тогда безусловное распределение ошибок оказывается распределением Лапласа с параметрами μ, λ.

Действительно, поскольку характеристическая функция нормального закона равна $\exp\left(i\mu t - \frac{\sigma^2 t^2}{2}\right)$, а плотность показательного распределения $f_{\sigma^2}(u) = \lambda e^{-\lambda u}$, $u \ge 0$, характеристическая функция безусловного распределения ошибок измерения вычисляется по формуле

$$\varphi(t) = \int_0^\infty \exp(i\mu t - \frac{ut^2}{2})\lambda e^{-\lambda u} du = \frac{\lambda e^{i\mu t}}{t^2/2+\lambda} = \frac{2\lambda e^{i\mu t}}{t^2+2\lambda}, \quad (1)$$

а это есть характеристическая функция распределения Лапласа с плотностью распределения

$$f(x) = \frac{\sqrt{2\lambda}}{2}\exp(-\sqrt{2\lambda}\,|x-\mu|), \quad -\infty < x < \infty. \quad (2)$$

При этом дисперсия равна $\frac{2}{(\sqrt{2\lambda})^2} = \frac{1}{\lambda}$. Приравнивая дисперсию величине $\overline{\sigma}^2$, найдем $\lambda = \frac{1}{\overline{\sigma}^2}$ и

$$f(x) = \frac{1}{\sqrt{2}\overline{\sigma}}\exp(-\frac{\sqrt{2}}{\overline{\sigma}}|x-\mu|), \quad -\infty < x < \infty. \quad (3)$$

В монографии [6] этот результат получен более сложным способом.

Уместно заметить, что случайность σ^2 физически означает, что эксперименты проводятся в меняющихся (нестабильных) случайных условиях, влияющих на точность их проведения.

Критерий Гири как критерий отношения правдоподобия при различении распределений Гаусса и Лапласа

Рассмотрим задачу: по выборке $X = (x_1, ..., x_n)$ проверить гипотезу нормальности этой выборки $\left(x_i \in N\left(\mu, \sigma^2\right)\right)$ против альтернативной гипотезы: x_i распределены по закону Лапласа $\left(x_i \in L\left(\mu, \sigma^{-2}\right)\right)$.

Пусть сначала гипотезы простые:

$$H_0: \quad f(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right); \quad (4)$$

$$H_1: \quad f(x) = \frac{1}{\sigma\sqrt{2}}\exp\left(-\frac{\sqrt{2}|x-\mu|}{\sigma}\right), \quad (5)$$

где μ, σ – известные параметры.

На основании фундаментальной леммы Неймана – Пирсона [7, 8] для проверки двух простых гипотез наиболее мощным критерием является критерий отношения правдоподобия

$$\Phi(X) = \begin{cases} 1, & \frac{L_1(X)}{L_0(X)} \ge h_\alpha, \\ 0, & \frac{L_1(X)}{L_0(X)} < h_\alpha, \end{cases} \quad (6)$$

где h_α находится из условия

$$E_{H_0}\Phi(X) = \alpha, \quad (7)$$

$L_0(X)$, $L_1(X)$ – функции правдоподобия при гипотезах $H_0(X)$, $H_1(X)$ соответственно:

$$L_0(X) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\mu)^2\right); \quad (8)$$

$$L_1(X) = \left(\frac{1}{\sigma\sqrt{2}}\right)^n \exp\left(-\frac{\sqrt{2}}{\sigma}\sum_{i=1}^n |x_i-\mu|\right). \quad (9)$$

Отсюда отношение правдоподобия

$$\frac{L_1(X)}{L_0(X)} = \\ = \left(\sqrt{\pi}\right)^n \exp\left(-\frac{\sqrt{2}}{\sigma}\sum_{i=1}^n |x_i-\mu| + \frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\mu)^2\right). \quad (10)$$

В случае неизвестных параметров μ, σ имеем две сложные гипотезы с двумя мешающими параметрами. Состоятельными оценками этих параметров по выборке являются выборочное среднее

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{11}$$

и выборочное среднеквадратическое отклонение

$$s = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}. \tag{12}$$

При подстановке этих оценок в выражение для отношения правдоподобия (10) нетрудно видеть, что оно инвариантно относительно мешающих параметров:

$$\frac{L_1(X)}{L_0(X)} = \left(\sqrt{\pi}\right)^n \times$$

$$\times \exp\left(-\sqrt{\frac{2}{s^2}}\sum_{i=1}^{n}\left|x_i - \bar{x}\right| + \frac{1}{2s^2}\sum_{i=1}^{n}(x_i - \bar{x})^2\right) =$$

$$= \left(\sqrt{\pi}\right)^n \exp\left(-\sqrt{\frac{2}{s^2}}\sum_{i=1}^{n}\left|x_i - \bar{x}\right| + \frac{n}{2}\right). \tag{13}$$

Вместо отношения правдоподобия удобнее рассматривать его логарифм. После логарифмирования и несложных преобразований (13) найдем, что критическая область

$$W = \left\{X:\ \frac{L_1(X)}{L_0(X)} \ge h_\alpha\right\} \tag{14}$$

эквивалентна области

$$W = \left\{X:\ \frac{1}{n\sqrt{s^2}}\sum_{i=1}^{n}\left|x_i - \bar{x}\right| \le x_\alpha\right\}, \tag{15}$$

где критическая точка x_α ищется из условия

$$P_{H_0}\left(\frac{1}{n\sqrt{s^2}}\sum_{i=1}^{n}\left|x_i - \bar{x}\right| \le x_\alpha\right) = \alpha. \tag{16}$$

При этом статистикой критерия является величина

$$d = \frac{1}{n\sqrt{s^2}}\sum_{i=1}^{n}\left|x_i - \bar{x}\right|. \tag{17}$$

Отметим, что распределение статистики d не зависит от параметров сдвига μ и масштаба σ, так что критерий, основанный на d, инвариантен относительно параметров как при гипотезе H_0, так и при гипотезе H_1. Статистика d – это статистика, на которой основан критерий Гири, предназначенный для проверки нормальной выборки. В работе [9] Гири показал, что распределение статистики d, если верна гипотеза H_0, уже при $n > 50$ удовлетворительно аппроксимируется нормальным. В работе [2] Большев и Смирнов для выборок из нормального распределения приводят формулы для первых двух моментов статистики d:

$$Ed_0 = \frac{2}{\sqrt{\pi(n-1)}}\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} =$$

$$= \sqrt{\frac{2}{\pi}}\left[1 + \frac{2}{8n-9} + O\left(\frac{1}{n^3}\right)\right], \tag{18}$$

$$Dd_0 = \frac{1}{n}\left[1 + \frac{2}{\pi}\left(\sqrt{n(n-2)} + \arcsin\frac{1}{n-1}\right)\right] -$$

$$-\frac{n-1}{\pi}\left[\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)}\right]^2 = \frac{1}{n}\left[\left(1 - \frac{3}{\pi}\right) - \frac{1}{4\pi n} + O\left(\frac{1}{n^2}\right)\right]. \tag{19}$$

Здесь случайная величина d_0 – статистика d при основной гипотезе H_0; E – символ математического ожидания; D – символ дисперсии.

Из формул (18), (19) следует, что

$$\lim_{n\to+\infty} Ed_0 = \sqrt{\frac{2}{\pi}} \approx 0,797885, \quad \lim_{n\to+\infty} Dd_0 = 0.$$

Таким образом, при гипотезе H_0 распределение статистики d при $n \to \infty$ становится вырожденным, сосредоточенным в точке $\sqrt{\frac{2}{\pi}}$.

Асимптотические свойства критерия Гири

Поскольку распределение статистики d инвариантно относительно параметров μ и σ, положим μ = 0, σ = 1. Случайную величину с этими значениями параметров обозначим ξ.

Для нахождения асимптотических моментов статистики d воспользуемся методом линеаризации [10–12] и учтем, что при $n \to \infty$ выборочное среднее $\bar{x}$ с вероятностью 1 сходится к $E\xi = 0$.

Имеем при $n \to \infty$

$$Ed = \frac{E|\xi|}{Es} + o\left(\frac{1}{n}\right), \tag{20}$$

$$Dd = \frac{1}{n}D\frac{\xi}{s} + o\left(\frac{1}{n}\right) =$$

$$= \frac{1}{n}\left(\frac{1}{(Es)^2}D|\xi| + \frac{\left(E|\xi|\right)^2}{(s)^4}Ds - 2\frac{E|\xi|}{(Es)^3}\right) + o\left(\frac{1}{n}\right). \tag{21}$$

Для выборок из нормального распределения вычисления по формулам (20) и (21) приводят к уже известным результатам (18) и (19).

Для выборок из распределения Лапласа имеем:

$$f_\xi(x) = \frac{1}{\sqrt{2}}\exp\left(-\sqrt{2}|x|\right), \quad x > 0;$$

$$E\xi = 0, \quad D\xi = 1, \quad \mu_3 = E\xi^3 = 6;$$

$$E|\xi| = \frac{1}{\sqrt{2}}, \quad D|\xi| = E\xi^2 - \left(E\xi\right)^2 = \frac{1}{2};$$

$$Ds = D\sqrt{s^2} = \frac{\mu_4 - \mu_2^2}{4\mu_2} = \frac{5}{4};$$

$$\operatorname{cov}\left(|\xi|, s\right) = \frac{1}{2}\operatorname{cov}\left(|\xi|, s^2\right) = \frac{1}{2}\left(E|\xi|\xi^2 - E|\xi| \cdot E\xi\right) =$$

$$= \frac{1}{2}\left(2\int_0^{+\infty} x^3 \frac{1}{\sqrt{2}} e^{-\sqrt{2}x} dx - \frac{1}{\sqrt{2}}\right) = \frac{\sqrt{2}}{2}.$$

Подставляя найденные значения в (20), (21), получим:

$$Ed_1 = \frac{1}{\sqrt{2}} + o\left(\frac{1}{n}\right) \cong 0,7072, \tag{22}$$

$$Dd_1 = \frac{1}{n}\left(\frac{1}{2} + \frac{1}{2}\frac{5}{4} - \frac{2}{\sqrt{2}}\frac{\sqrt{2}}{2}\right) + o\left(\frac{1}{n}\right) =$$

$$= \frac{0,125}{n} + o\left(\frac{1}{n}\right), \tag{23}$$

где через d_1 обозначена статистика d при гипотезе H_1.

Из (23) следует, что асимптотическое распределение статистики d_1 является вырожденным. Вместе с вырожденностью распределения d_1 это означает, что мощность критерия d стремится к 1 при $n \to \infty$, и d – состоятельный критерий.

Поскольку выборочные моменты сходятся с вероятностью 1 к истинным значениям моментов $\bar{x} \xrightarrow[n\to\infty]{\text{п.в.}} \mu$, $s^2 \xrightarrow[n\to\infty]{\text{п.в.}} \sigma^2$, асимптотическое распределение статистики d при обеих гипотезах совпадает с асимптотическим распределением статистики $d^* = \frac{1}{\sigma n} \cdot \sum_{i=1}^{n} |x_i - \mu|$, так как $\frac{d^*}{d} = s \xrightarrow[n\to\infty]{\text{п.в.}} 1.$

Но по центральной предельной теореме для сумм независимых одинаково распределенных случайных величин распределение $\sqrt{n}d^*$ асимтотически нормально. Тем самым доказана асимптотическая нормальность d_0 и d_1.

Исследование свойств критерия Гири методом статистического моделирования

Методом статистического моделирования были проведены расчеты критических точек x_α для выборок из нормального распределения и мощности π_α критерия Гири для выборок из распределения Лапласа при n = 10 (5) 50 (10) 100 (20) 200 (100) 1000 и α = 0,2; 0,1; 0,05; 0,01; 0,005. При моделировании для каждого объема n соответствующие выборки генерировались k = 10000 раз. Расчеты производились с использованием программного пакета *Wolfram Mathematica 8*. Результаты приведены в табл. 1.

Таблица 1. **Критические точки и мощности критерия Гири в зависимости от объема выборки при различных уровнях значимости**

	α = 0,2		a = 0,1		a = 0,05		a = 0,01		a =0,005	
n	x_a	π_a	x_a	π_a	x_a	π_a	x_a	π_a	x_a	π_a
10	0,7699	0,4510	0,7436	0,3120	0,7218	0,2138	0,6810	0,0866	0,6661	0,0582
15	0,7692	0,5647	0,7467	0,4285	0,7281	0,3214	0,6933	0,1598	0,6805	0,1170
20	0,7704	0,6556	0,7506	0,5301	0,7342	0,4236	0,7034	0,2432	0,6921	0,1893
25	0,7719	0,7115	0,7539	0,5956	0,7391	0,4926	0,7112	0,3059	0,7010	0,2462
30	0,7732	0,7627	0,7567	0,6588	0,7430	0,5622	0,7175	0,3754	0,7081	0,3117
35	0,7744	0,8057	0,7590	0,7106	0,7463	0,6186	0,7225	0,4310	0,7138	0,3640
40	0,7754	0,8407	0,7610	0,7562	0,7491	0,6713	0,7267	0,4894	0,7186	0,4213
45	0,7763	0,8618	0,7627	0,7863	0,7514	0,7089	0,7303	0,5372	0,7226	0,4706
50	0,7772	0,8843	0,7642	0,8164	0,7535	0,7447	0,7334	0,5795	0,7261	0,5132
60	0,7786	0,9203	0,7667	0,8672	0,7569	0,8080	0,7385	0,6610	0,7317	0,5981
70	0,7797	0,9419	0,7687	0,9006	0,7596	0,8530	0,7425	0,7281	0,7363	0,6719
80	0,7807	0,9562	0,7703	0,9233	0,7618	0,8840	0,7458	0,7763	0,7400	0,7258
90	0,7815	0,9700	0,7717	0,9448	0,7637	0,9133	0,7486	0,8215	0,7431	0,7763
100	0,7822	0,9763	0,7729	0,9558	0,7653	0,9297	0,7510	0,8512	0,7457	0,8115
120	0,7834	0,9885	0,7749	0,9768	0,7679	0,9607	0,7548	0,9071	0,7500	0,8776
140	0,7843	0,9942	0,7765	0,9876	0,7700	0,9780	0,7578	0,9433	0,7534	0,9229
160	0,7851	0,9966	0,7778	0,9924	0,7717	0,9860	0,7603	0,9615	0,7561	0,9465
180	0,7857	0,9981	0,7788	0,9957	0,7731	0,9919	0,7624	0,9762	0,7584	0,9661
200	0,7863	0,9992	0,7797	0,9979	0,7743	0,9959	0,7641	0,9865	0,7604	0,9800
300	0,7883	1,0000	0,7829	0,9999	0,7785	0,9997	0,7701	0,9987	0,7671	0,9979
400	0,7895	1,0000	0,7848	1,0000	0,7810	1,0000	0,7737	0,9999	0,7711	0,9998
500	0,7903	1,0000	0,7861	1,0000	0,7827	1,0000	0,7762	1,0000	0,7739	1,0000
600	0,7909	1,0000	0,7871	1,0000	0,7840	1,0000	0,7781	1,0000	0,7759	1,0000
700	0,7914	1,0000	0,7879	1,0000	0,7850	1,0000	0,7795	1,0000	0,7775	1,0000
800	0,7918	1,0000	0,7885	1,0000	0,7858	1,0000	0,7807	1,0000	0,7788	1,0000
900	0,7922	1,0000	0,7890	1,0000	0,7865	1,0000	0,7817	1,0000	0,7799	1,0000
1000	0,7924	1,0000	0,7895	1,0000	0,7871	1,0000	0,7825	1,0000	0,7808	1,0000

Табл. 1 позволяет при заданных объеме выборки n и уровню значимости α определить мощность критерия. Например, при $n = 100$ и $\alpha = 0,1$ мощность равна 0,9558. Кроме того, табл. 1 позволяет определить минимальный объем выборки, при котором гипотезы различаются с заданной ошибкой первого рода α и заданной мощностью критерия π. Например, при $\alpha = 0,05$ и $\pi = 0,95$ минимальный объем выборки $n = 105$ (использовалась линейная интерполяция между $n = 100$ и $n = 120$).

При моделировании также были вычислены оценки средних Ed_0 и Ed_1, среднеквадратических отклонений $\sqrt{Dd_0}$ и $\sqrt{Dd_1}$ при гипотезах H_0 и H_1 соответственно для тех же объемов выборок. Результаты приведены в табл. 2.

Таблица 2. **Математические ожидания и среднеквадратические отклонения распределений статистик d_0 и d_1 в зависимости от объема выборки**

n	Ed_0	$\sqrt{Dd_0}$	Ed_1	$\sqrt{Dd_1}$	n	Ed_0	$\sqrt{Dd_0}$	Ed_1	$\sqrt{Dd_1}$
10	0,8209	0,0595	0,7788	0,0718	120	0,7995	0,0192	0,7151	0,0300
15	0,8127	0,0509	0,7585	0,0655	140	0,7991	0,0182	0,7132	0,0282
20	0,8079	0,0455	0,7460	0,0611	160	0,7991	0,0166	0,7134	0,0265
25	0,8067	0,0407	0,7401	0,0569	180	0,7990	0,0158	0,7120	0,0254
30	0,8047	0,0377	0,7346	0,0540	200	0,7990	0,0148	0,7116	0,0238
35	0,8040	0,0349	0,7312	0,0501	300	0,7987	0,0122	0,7102	0,0199
40	0,8035	0,0327	0,7280	0,0475	400	0,7986	0,0106	0,7092	0,0174
45	0,8026	0,0312	0,7260	0,0463	500	0,7983	0,0093	0,7090	0,0155
50	0,8015	0,0290	0,7246	0,0439	600	0,7982	0,0087	0,7088	0,0144
60	0,8013	0,0271	0,7217	0,0404	700	0,7982	0,0081	0,7086	0,0131
70	0,8010	0,0248	0,7191	0,0386	800	0,7981	0,0075	0,7085	0,0124
80	0,8007	0,0236	0,7179	0,0367	900	0,7981	0,0071	0,7082	0,0116
90	0,7998	0,0219	0,7170	0,0343	1000	0,7980	0,0067	0,7081	0,0112
100	0,7997	0,0212	0,7164	0,0332					

Данные табл. 2 подтверждают, что при больших n распределения статистик d_0 и d_1 также стремятся к вырожденному.

На рис. 1 для наглядности приведены графики функции мощности критерия Гири для $\alpha = 0,2$; 0,1; 0,05; 0,01; 0,005 (сверху вниз соответственно) как функции от n, при этом ось n представлена в логарифмическом масштабе.

На рис. 2 приведены гистограммы распределений статистик d_0 и d_1 при объемах выборок $n = 50$, 100, 200 соответственно, а также их аппроксимации нормальными законами с параметрами, взятыми из табл. 2. Видно, что с увеличением объема выборки кривые плотностей расходятся.

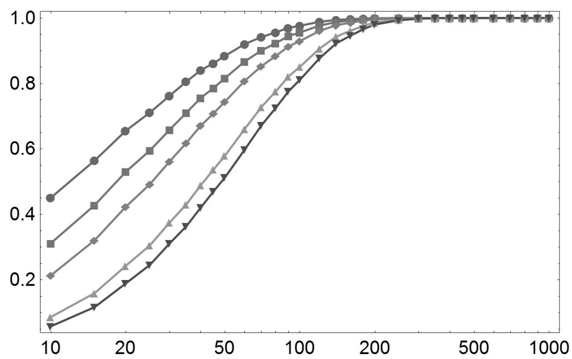

Рис. 1. Мощности критерия Гири

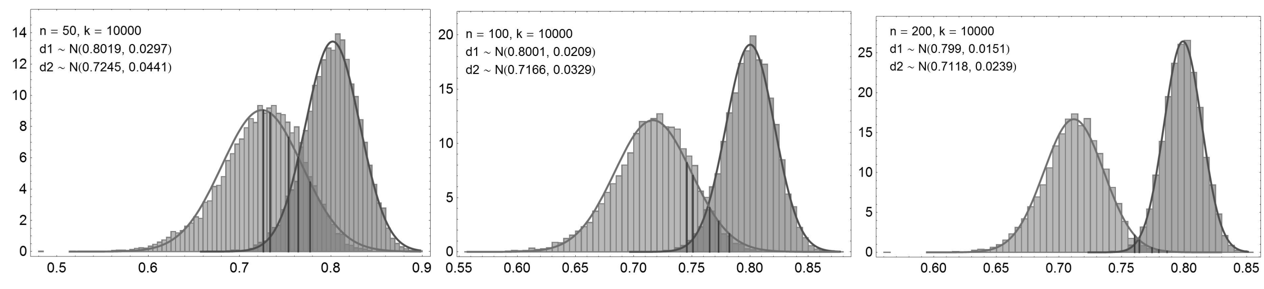

Рис. 2. Гистограммы распределений статистик d_0 и d_1 и их аппроксимации

Заключение

В последние годы при статистической обработке результатов экспериментов в различных областях техники все чаще используется распределение Лапласа в качестве альтернативы нормальному распределению. Это обусловлено лучшим согласием с ним экспериментальных данных в случаях, когда эксперименты проводятся в различных условиях, имеющих случайную природу.

Предметом настоящих исследований являлись распределение Лапласа и критерий Гири при применении его к распределению Лапласа в качестве альтернативы нормальному закону.

Основные результаты исследований заключаются в следующем.

• Дано теоретическое обоснование механизма возникновения распределения Лапласа из нормального распределения.

• Показано, что критерий отношения правдоподобия, примененный к различению законов Гаусса и Лапласа, сводится к критерию Гири.

• Доказаны инвариантность, состоятельность и асимптотическая нормальность критерия Гири при обеих альтернативах.

• Методом статистического моделирования в широкой области изменения уровней значимости и объемов вычислены критические точки и мощности критерия для обоих распределений (Гаусса и Лапласа), построены гистограммы статистики критерия и их аппроксимации нормальными законами, вычислены математические ожидания и среднеквадратические отклонения указанной статистики при обеих гипотезах.

Полученные теоретические результаты и результаты статистического моделирования свидетельствуют о высокой эффективности критерия Гири при различении распределений Гаусса и Лапласа даже по выборкам сравнительно небольшого объема. Это позволяет рекомендовать его к более широкому практическому применению, а приведенные в статье таблицы и графики существенно упрощают решение подобного рода задач.

В заключение приносим благодарность выпускнику СПбГМТУ Киму Адамейко за проведение численных расчетов.

Библиографические ссылки

1. *Бенинг В. Е.*, *Королёв В. Ю.* Некоторые статистические задачи, связанные с распределением Лапласа // Информатика и ее применения. – 2008. – Т. 2. – № 2. – С. 19–34.

2. *Большев Л. Н.*, *Смирнов Н. В.* Таблицы математической статистики. – Изд. 3-е. – М. : Наука, 1983. – 415 с.

3. *Орлов А. И.* Распространенная ошибка при использовании критериев Колмогорова и омега-квадрат // Заводская лаборатория. – 1985. – Т. 51. – № 1. – С. 60–62.

4. *Каган А. М.*, *Линник Ю. В.*, *Рао С. Р.* Характеризационные задачи математической статистики. – М. : Наука, 1972. – 656 с.

5. Эффективность и надежность сложных систем / И. Л. Плетнев, А. И. Рембеза, Ю. А. Соколов, В. А. Чалый-Прилуцкий. – М. : Машиностроение, 1977. – 215 с.

6. *Мудров В. И.*, *Кушко В. Л.* Методы обработки наблюдений. – М. : Радио и связь, 1983. – 303 с.

7. *Боровков А. А.* Математическая статистика : Оценка параметров, проверка гипотез. – М. : Наука, 1984. – 472 с.

8. *Леман Э.* Проверка статистических гипотез. – М. : Наука, 1979. – 408 с.

9. *Geary R. C.* The Ratio of the Mean Deviation to the Standard Deviation as a Test of Normality // Biometrika. – 1935. – No. 27(3/4). – С. 310–332.

10. *Вентцель Е. С.*, *Овчаров Л. А.* Теория вероятностей и ее инженерные приложения. – М. : Наука, 1979. – 480 с.

11. *Кендалл М. Дж.*, *Стьюарт А.* Теория распределений. – М. : Наука,1966. – 587 с.

12. *Орлов А. И.* Прикладная статистика. – М. : Экзамен, 2006. – 671 с.

I. V. Zolotukhin, PhD (Physics and Mathematics), Associate Professor, St.-Petersburg Department of the P. P. Shirshov Institute of Oceanology
L. A. Zolotukhina, DSc (Physics and Mathematics), Professor, State Marine Technical University of St. Petersburg

Geary Test in the Problems Connected with the Gauss and Laplace Distributions

Laplace distribution as an alternative to Gaussian distribution as well as Geary test are under consideration. The mechanism for obtaining Laplace distribution from Gaussian one is given. It is proved that for our case the Geary test is the likelihood ratio test. Also invariance, consistency and asymptotic normality of the test are proved. The statistical simulation methods were used to compile the tables and graphs that confirm the test high efficiency are presented.

Key words: Gaussian distribution, Laplace distribution, Geary test, likelihood ratio method, statistical simulation.