УДК 519.22 (075.8)
DOI: 10.17277/voprosy.2018.02.pp.180-187

О НЕКОТОРЫХ ХАРАКТЕРИСТИКАХ КРИТЕРИЯ СОГЛАСИЯ

В. И. Фомин

ФГБОУ ВО «Тамбовский государственный технический университет», г. Тамбов, Россия

Рецензент д-р техн. наук, профессор Е. И. Глинкин

Ключевые слова: альтернативная гипотеза; критерий согласия; надежности первого и второго рода; основная гипотеза; ошибочности первого и второго рода; функция надежности критерия согласия; функция ошибочности критерия согласия.

Аннотация: Введены понятия надежности первого рода как вероятности принятия основной гипотезы в случае, когда она верна, и надежности второго рода как вероятности отклонения основной гипотезы в случае, когда она неверна. Указаны взаимосвязи между надежностями и ошибочностями первого и второго рода.

Пусть G – некоторый эксперимент; X_F – случайная величина, связанная с экспериментом G; $F = F(x) = p(X_F < x)$ – функция распределения вероятностей (закон распределения) случайной величины X_F (для непрерывной случайной величины X_F в качестве закона распределения можно рассматривать плотность распределения вероятностей $f(x) = F'(x)$). Пусть из каких-либо теоретических или практических соображений известно, что $F \in \Phi$, где Φ – некоторое заданное семейство распределений. Требуется получить дополнительную информацию о функции распределения F. Рассмотрим, например, в качестве основной гипотезы сложную гипотезу $H_0 : F \in \Phi_0$, где Φ_0 – некоторое конкретное собственное подмножество семейства Φ, содержащее более одного элемента. Альтернативная гипотеза имеет вид $H_1 : F \in \Phi_1$, где $\Phi_1 = \Phi \setminus \Phi_0$.

Статистическая проверка гипотезы H_0 проводится на основе анализа наблюдаемых значений $\vec{x} = (x_1, \ldots, x_n)$ случайной величины X_F при проведении эксперимента G n раз (в n независимых испытаниях) или, други-

Фомин Василий Ильич – кандидат физико-математических наук, доцент кафедры «Техническая механика и детали машин», e-mail: vasiliyfomin@bk.ru, ТамбГТУ, г. Тамбов, Россия.

ми словами, на основе анализа выборки объема *n* (из рассматриваемой генеральной совокупности *W* объема *N*, каждый объект которой характеризуется количественным признаком X_F, выбирают случайным образом *n* объектов и для каждого из них измеряют значение количественного признака X_F; в результате получают выборку $\vec{x} = (x_1, \dots, x_n)$, где x_i – значение количественного признака X_F для *i*-го объекта выборки, $i = \overline{1, n}$).

Статистическая проверка гипотез проводится с помощью критериев согласия.

В целях ясности дальнейшего изложения напомним общую схему построения критерия согласия [1, с. 466].

I. Задается достаточно малое число $\alpha > 0$ (уровень значимости критерия).

II. Подбирается число *n* (число наблюдений случайной величины или объем выборки), некоторая статистика $G_n(\vec{X}) = G_n(X_1, \dots, X_n)$, где X_i – случайная величина X в *i*-м испытании, $i = \overline{1, n}$ (статистика критерия согласия), и некоторое число $\gamma = \gamma(\alpha, n)$ (граница критерия согласия), определяемое заданием чисел α, n, таким образом, что выборочное пространство

$$V = \left\{ \vec{x} = (x_1, \dots, x_n) \in \mathbb{R}^n \mid x_i \in \{x\},\ i = \overline{1, n} \right\}$$

(здесь $\{x\}$ – множество возможных значений случайной величины X) разбивается на два непересекающихся подмножества V_0 и V_1 (V_0 и V_1 – соответственно область принятия гипотезы и критическая область гипотезы; ниже рассматривается правосторонняя критическая область):

$$V_0 = \left\{ \vec{x} \in V \mid g_n(\vec{x}) < \gamma \right\};$$

$$V_1 = \left\{ \vec{x} \in V \mid g_n(\vec{x}) \ge \gamma \right\},$$

где $g_n(\vec{x}) = G_n(\vec{X})\Big|_{\vec{X} = \vec{x}}$ – значение (реализация) статистики $G_n(\vec{X})$ при значении случайного вектора $\vec{X} = (X_1, \dots, X_n)$, равном $\vec{x} = (x_1, \dots, x_n)$. Подбор указанных величин проводят таким образом, чтобы выполнялось условие

$$p(G_n(\vec{X}_\varphi) \ge \gamma) \le \alpha, \forall \varphi \in \Phi_0, \tag{1}$$

где $\vec{X}_\varphi = (X_1^\varphi, \dots, X_n^\varphi)$, X_i^φ – случайная величина X_φ в *i*-м испытании, $i = \overline{1, n}$. Условие (1) означает следующее: если гипотеза H_0 является верной, то есть $F = \varphi$, где $\varphi \in \Phi_0$, то событие

$$A_1 = \{G_n(\vec{X}_\varphi) \ge \gamma\} \tag{2}$$

является маловероятным (практически невозможным).

III. Берутся результаты конкретной выборки $\vec{x}^* = (x_1^*, \dots, x_n^*)$ объема *n* (напомним, что x_i^* – наблюдаемое значение случайной величины X_F в *i*-м испытании, $i = \overline{1, n}$) и вычисляют

$$g_n(\vec{x}^*) = G_n(\vec{X}_F)\ _{\vec{X}_F = \vec{x}^*},$$

где $G_n(\vec{X}_F) = G_n(X_1^F, ..., X_n^F)$, X_i^F – случайная величина X_F в i-м испытании, $i = \overline{1, n}$.

IV. Сравнивают $g_n(\vec{x}^*)$ с границей критерия γ: если $g_n(\vec{x}^*) \geq \gamma$, то есть $\vec{x}^* \in V_1$, то эмпирические (опытные, экспериментальные) данные не согласуются с гипотезой H_0 (противоречат гипотезе H_0) и ее отклоняют (действительно, если предположить, что гипотеза H_0 верна, то, в силу маловероятности события (2), также маловероятно событие $\{g_n(\vec{x}^*) \geq \gamma\}$, что является основанием для отклонения гипотезы H_0); если $g_n(\vec{x}^*) < \gamma$, т.е. $\vec{x}^* \in V_0$, то эмпирические данные согласуются с гипотезой H_0 и ее принимают (в этом случае нет оснований для отклонения гипотезы H_0).

Статистика $G_n(\vec{X}_F) = G_n(X_1^F, ..., X_n^F)$, будучи функцией случайных величин $X_1^F, ..., X_n^F$, сама является случайной величиной. Следовательно, тот факт, что в результате n испытаний значение статистики окажется не меньше γ, и тот факт, что значение статистики окажется меньше γ, являются случайными событиями: $B_1 = \{g_n(\vec{x}^*) \geq \gamma\}$, $B_0 = \{g_n(\vec{x}^*) < \gamma\}$ – случайные события. По этой причине при статистической проверке гипотезы H_0 с помощью критерия согласия возможны следующие ошибки:

а) выдвинутая гипотеза H_0 верна, но произойдет событие B_1 и гипотеза H_0 будет отклонена (ошибка первого рода);

б) выдвинутая гипотеза H_0 неверна, но произойдет событие B_0 и гипотеза H_0 будет принята (ошибка второго рода).

Пусть G^* – однократное применение критерия согласия $K = K(\alpha; n; G_n(\vec{X}); \gamma)$. В результате эксперимента G^* произойдет одно из следующих событий:

P_1: гипотеза H_0 принимается в случае, когда H_0 верна (правильное решение первого рода);

P_2: гипотеза H_0 отклоняется в случае, когда H_0 неверна (правильное решение второго рода);

N_1: гипотеза H_0 отклоняется в случае, когда H_0 верна, то есть совершается ошибка первого рода (неправильное решение первого рода);

N_2: гипотеза H_0 принимается в случае, когда H_0 неверна, то есть совершается ошибка второго рода (неправильное решение второго рода).

Рассмотрим функции, характеризующие вероятности принятия правильных и неправильных решений. В дальнейшем будем использовать оператор определения :: = (по определению).

Функция надежности критерия согласия $::=$ функция $N_K : \Phi \to [0;1]$, задаваемая формулой

$$N_K(\varphi) = \begin{cases} p(G_n(\vec{X}_\varphi) < \gamma), \ \varphi \in \Phi_0, \\ p(G_n(\vec{X}_\varphi) \ge \gamma), \ \varphi \in \Phi_1. \end{cases}$$

Заметим, что функция $N_K(\varphi)$, рассматриваемая при $\varphi \in \Phi_1$, называется функцией мощности критерия согласия [2, с. 263].

Надежность первого рода $::=$ вероятность принятия правильного решения первого рода (обозначение ν_1).

Пусть гипотеза H_0 верна, то есть $F = \varphi_0$, где φ_0 – некоторое конкретное распределение из Φ_0. Значит,

$$\nu_1 = p(G_n(\vec{X}_{\varphi_0}) < \gamma) = N_K(\varphi_0)\,. \tag{3}$$

Мы не знаем, какому именно φ_0 равняется F. Поэтому можно лишь сказать, что

$$\nu_1 \in \{N_K(\varphi) \mid \varphi \in \Phi_0\}\,.$$

Следовательно,

$$\nu_1 \ge \inf_{\varphi \in \Phi_0} N_K(\varphi) \overset{\text{def}}{=} \nu_{\min,1}\,.$$

Назовем число $\nu_{\min,1}$ нижней границей надежности первого рода.

Надежность второго рода $::=$ вероятность принятия правильного решения второго рода (обозначение ν_2).

Заметим, что ν_2 называется мощностью критерия согласия [3, с. 106].

Пусть гипотеза H_0 неверна, то есть $F = \varphi_1$, где $\varphi_1 \in \Phi_1$. Значит,

$$\nu_2 = p(G_n(\vec{X}_{\varphi_1}) \ge \gamma) = N_K(\varphi_1)\,, \tag{4}$$

поэтому

$$\nu_2 \in \{N_K(\varphi) \mid \varphi \in \Phi_1\}\,.$$

Следовательно,

$$\nu_2 \ge \inf_{\varphi \in \Phi_1} N_K(\varphi) \overset{\text{def}}{=} \nu_{\min,2}\,.$$

Назовем число $\nu_{\min,2}$ нижней границей надежности второго рода.

Надежность критерия согласия $::=$ вероятность принятия правильного решения неважно какого рода (обозначение ν).

Таким образом, по определению, $\nu = p(C)$, где событие C – будет принято правильное решение.

Пусть G_0 – выдвижение гипотезы H_0. Для эксперимента G_0 события Q_1 – выдвинутая гипотеза H_0 верна, Q_2 – выдвинутая гипотеза H_0 неверна образуют полную группу событий. По формуле полной вероятности получаем

$$\nu = p(C) = p(Q_1)p_{Q_1}(C) + p(Q_2)p_{Q_2}(C),$$

то есть

$$\nu = p(Q_1)\nu_1 + p(Q_2)\nu_2,$$

или, в силу равенства $p(Q_1) + p(Q_2) = 1$,

$$\nu = p(Q_1)\nu_1 + \bigl(1 - p(Q_1)\bigr)\nu_2. \qquad (5)$$

Из соотношения (5) следует оценка

$$\nu \geq p(Q_1)\nu_{\min,1} + \bigl(1 - p(Q_1)\bigr)\nu_{\min,2}$$

(величину $p(Q_1)$ мы не знаем).

Функция ошибочности критерия согласия $::=$ функция $\Delta_K : \Phi \to [0;1]$, задаваемая равенством

$$\Delta_K(\varphi) = \begin{cases} p(G_n(\vec{X}_\varphi) \geq \gamma), & \varphi \in \Phi_0, \\ p(G_n(\vec{X}_\varphi) < \gamma), & \varphi \in \Phi_1. \end{cases}$$

Ошибочность первого рода $::=$ вероятность принятия неправильного решения первого рода, то есть вероятность ошибки первого рода (обозначение δ_1).

Пусть гипотеза H_0 верна, то есть $F = \varphi_0$, где $\varphi_0 \in \Phi$. Значит,

$$\delta_1 = p(G_n(\vec{X}_{\varphi_0}) \geq \gamma) = \Delta_K(\varphi_0), \qquad (6)$$

поэтому

$$\delta_1 \in \{\Delta_K(\varphi) \mid \varphi \in \Phi_0\}.$$

Следовательно,

$$\delta_1 \leq \sup_{\varphi \in \Phi_0} \Delta_K(\varphi) \overset{\text{def}}{=} \delta_{\max,1}.$$

Назовем $\delta_{\max,1}$ верхней границей ошибочности первого рода.

Заметим, что условие (1) можно записать в виде

$$\Delta_K(\varphi) \leq \alpha, \ \forall \varphi \in \Phi_0,$$

в частности,

$$\Delta_K(\varphi_0) \leq \alpha. \qquad (7)$$

Из соотношений (6), (7) получаем оценку

$$\delta_1 \leq \alpha. \qquad (8)$$

Ошибочность второго рода $::=$ вероятность принятия неправильного решения второго рода, то есть вероятность ошибки второго рода (обозначение δ_2).

Пусть гипотеза H_0 неверна, то есть $F=\varphi_1$, где $\varphi_1 \in \Phi_1$. Значит,

$$\delta_2 = p(G_n(\vec{X}_{\varphi_1})<\gamma)=\Delta_K(\varphi_1), \tag{9}$$

поэтому

$$\delta_2 \in \{\Delta_K(\varphi) \mid \varphi \in \Phi_1\}.$$

Следовательно,

$$\delta_2 \le \sup_{\varphi\in\Phi_1} \Delta_K(\varphi) \overset{\text{def}}{=} \delta_{\max,2}.$$

Назовем $\delta_{\max,2}$ верхней границей ошибочности второго рода.

Ошибочность критерия согласия $::=$ вероятность принятия неправильного решения неважно какого рода (обозначение δ).

По определению, $\delta = p(L)$, где событие L – будет принято неправильное решение. Получаем

$$\delta = p(L) = p(Q_1)p_{Q_1}(L) + p(Q_2)p_{Q_2}(L),$$

то есть

$$\delta = p(Q_1)\delta_1 + \bigl(1-p(Q_1)\bigr)\delta_2. \tag{10}$$

Из соотношения (10) получаем оценку

$$\delta \le p(Q_1)\delta_{\max,1} + \bigl(1-p(Q_1)\bigr)\delta_{\max,2}.$$

Заметим, что $\nu+\delta=1$.

Из вышесказанного следует, что в результате однократного применения критерия согласия K:

1) вероятность принятия правильного решения первого рода не меньше, чем $\nu_{\min,1}$;

2) вероятность принятия правильного решения второго рода не меньше, чем $\nu_{\min,2}$;

3) вероятность принятия неправильного решения первого рода не превосходит $\delta_{\max,1}$;

4) вероятность принятия неправильного решения второго рода не превосходит $\delta_{\max,2}$.

Укажем соотношение, связывающее функции $N_K(\varphi)$ и $\Delta_K(\varphi)$. События $\{G_n(\vec{X}_\varphi)<\gamma\}$ и $\{G_n(\vec{X}_\varphi)\ge\gamma\}$ противоположны, а для противоположных событий A и $\overline{A}$ справедлива формула $p(A)+p(\overline{A})=1$. Следовательно,

$$N_K(\varphi)+\Delta_K(\varphi)=1,\ \forall\varphi\in\Phi,$$

или

$$N_K(\varphi)=1-\Delta_K(\varphi),\ \forall\varphi\in\Phi. \tag{11}$$

Пусть гипотеза H_0 верна, то есть $F=\varphi_0$, где $\varphi_0\in\Phi_0$. В силу равенства (11)

$$N_K(\varphi_0)=1-\Delta_K(\varphi_0),$$

или, в силу равенств (3), (6),

$$\nu_1 = 1 - \delta_1 . \tag{12}$$

Из соотношений (8), (12) получаем оценку

$$\nu_1 \geq 1 - \alpha . \tag{13}$$

Из равенства (11) следует, что

$$\inf_{\varphi \in \Phi_0} N_K(\varphi) = 1 - \sup_{\varphi \in \Phi_0} \Delta_K(\varphi) ,$$

то есть

$$\nu_{\min,1} = 1 - \delta_{\max,1} .$$

Пусть гипотеза H_0 неверна, то есть $F = \varphi_1$, где $\varphi_1 \in \Phi_1$. В силу равенства (11)

$$N_K(\varphi_1) = 1 - \Delta_K(\varphi_1) ,$$

или, в силу соотношений (4), (9),

$$\nu_2 = 1 - \delta_2 .$$

Из равенства (11) получаем

$$\inf_{\varphi \in \Phi_1} N_K(\varphi) = 1 - \sup_{\varphi \in \Phi_1} \Delta_K(\varphi) ,$$

то есть

$$\nu_{\min,2} = 1 - \delta_{\max,2} . \tag{14}$$

Пусть K_α – класс критериев согласия заданного уровня значимости α (α фиксировано). Из оценок (8), (13) видно, что при любом фиксированном критерии согласия $K \in K_\alpha$ величины ν_1, δ_1 предопределены уровнем значимости α. Следовательно, улучшать качество критерия согласия можно лишь за счет увеличения надежности второго рода (мощности) критерия (то есть в силу равенства (14), за счет уменьшения ошибочности второго рода): взять вместо критерия K такой критерий $K_1 \in K_\alpha$, что $\nu_{\min,2}^{(K_1)} > \nu_{\min,2}^{(K)}$ (то есть $\delta_{\max,2}^{(K_1)} < \delta_{\max,2}^{(K)}$). Желательно выбрать (если такой выбор возможен) оптимальный по надежности второго рода (по мощности), другими словами, оптимальный по ошибочности второго рода критерий $K_* \in K_\alpha$:

$$\nu_{\min,2}^{(K_*)} \geq \nu_{\min,2}^{(K)} ,\ \forall K \in K_\alpha ; \tag{15}$$

$$\delta_{\max,2}^{(K_*)} \leq \delta_{\max,2}^{(K)} ,\ \forall K \in K_\alpha . \tag{16}$$

Соотношения (15), (16) можно записать в виде

$$\nu_{\min,2}^{(K_*)} = \sup_{K \in K_\alpha} \nu_{\min,2}^{(K)} ;$$

$$\delta_{\max,2}^{(K_*)} = \inf_{K \in K_\alpha} \delta_{\max,2}^{(K)} .$$

Заметим, что оптимальные по мощности критерии согласия найдены лишь для класса сложных гипотез специального вида (см., например, [4, с. 82]). Простейшие примеры таких критериев приведены в [5, с. 175].

Список литературы

1. Справочник по теории вероятностей и математической статистике / В. С. Королюк [и др.]. – Киев : Наукова Думка, 1978. – 584 с.
2. Коваленко, И. Н. Теория вероятностей / И. Н. Коваленко, Б. В. Гнеденко. – Киев : Выща шк., 1990. – 328 с.
3. Ивченко, Г. И. Математическая статистика / Г. И. Ивченко, Ю. И. Медведев. – М. : Высш. шк., 1984. – 248 с.
4. Леман, Э. Проверка статистических гипотез / Э. Леман. – М. : Наука, 1979. – 408 с.
5. Математическая статистика / В. Б. Горяинов [и др.]. – М. : Изд-во МГТУ им. Н. Э. Баумана, 2002. – 424 с.

References

1. Korolyuk V.S., Portenko N.I., Skorohod A.V., Turbin A.F. *Spravochnik po teorii veroyatnostey i matematicheskoy statistike* [Probability Theory and Mathematical Statistics handbook], Kiev: Naukova Dumka, 1978, 584 p. (In Russ.).
2. Kovalenko I.N., Gnedenko B.V. *Teoriya veroyatnostey* [Probability Theory], Kiev: Vysshaya shkola, 1990, 328 p. (In Russ.).
3. Ivchenko G.I., Medvedev Yu. I. *Matematicheskaya statistika* [Mathematical Statistics], Moscow: Vysshaya shkola, 1984, 248 p. (In Russ.).
4. Leman E. *Proverka statisticheskih gipotez* [Statistical hypotheses testing], Moscow: Nauka, 1979, 408 p. (In Russ.).
5. Goryainov V.B., Pavlov I.V., Cvetkova G.M., Teskin O.I. *Matematicheskaya statistika* [Mathematical Statistics], Moscow: Izdatal’stvo MGTU im. N.E. Baumana, 2002, 424 p. (In Russ.).

About Some Goodness-of-Fit Test Characteristics

V. I. Fomin

Tambov State Technical University, Tambov, Russia

Keywords: base hypothesis; alternative hypothesis; goodness-of-fit test; reliability function of goodness-of-fit test; reliabilities of first and second genus; inaccuracy function of goodness-of-fit test; inaccuracies of first and second genus.

Abstract: The notions of reliability of the first genus are introduced as the probability of accepting the base hypothesis in the case when it is true, and the reliability of the second genus as the probability of rejecting the base hypothesis in the case when it is incorrect. The interrelations between reliability and inaccuracy of the first and second genus are indicated.
