Б.Ю. Лемешко, С.Н. Постовалов

Рекомендации ГОСТ Р 50.1-033-2001 и ГОСТ Р 50.1-037-2002 как итог исследований вопросов применения критериев согласия [1]

1. Введение

С 1.07.2002 Госстандарт России взамен устаревших нормативных документов вводит в действие новые рекомендации по прикладной статистике "Правила проверки согласия опытного распределения с теоретическим". ГОСТ Р 50.1-033-2001 регламентирует применение критериев типа χ^2 [1], ГОСТ Р 50.1-037-2002 – применение непараметрических критериев [2]. Рекомендации предназначены для использования в качестве руководства при статистической обработке результатов наблюдений в различных приложениях: анализе данных физических и технологических экспериментов, обработке результатов измерений, задачах исследования надежности и контроля качества.

Необходимость разработки рекомендаций вызвана двумя основными факторами. Во-первых, в нормативных документах практически не оговорены правила применения критериев согласия при проверке сложных статистических гипотез. А именно такие ситуации оказываются наиболее реальными в приложениях. Поэтому практика использования критериев согласия в задачах контроля качества, исследования надежности и в других приложениях зачастую приводит к их некорректному применению и, как следствие, неверным выводам. Во-вторых, в наиболее доступной литературе учебного характера, к сожалению, часто содержатся принципиальные ошибки и неверные рекомендации, приводящие к некорректному или неоптимальному применению критериев и увеличивающие вероятность неверных выводов. Не решены данные вопросы и в популярных программных системах статистического анализа. В основу рекомендаций были положены результаты, полученные нами при исследовании свойств критериев типа χ^2 и непараметрических критериев типа Колмогорова, типа ω^2 и Ω^2 Мизеса.

Основными методами проводимых исследований, наряду с математическим аппаратом, явились развиваемая методика компьютерного моделирования и исследования статистических закономерностей. Благодаря этой методике удалось получить достаточно точные решения для задач, которые не могли быть получены аналитическими средствами.

[1] Выполнена при поддержке Российского фонда фундаментальных исследований (проект є 00-01-00913)

2. Вопросы применения критериев типа χ^2

Проверка статистических гипотез о согласии эмпирических данных с теоретическим законом распределения с применением критериев согласия типа χ^2 обусловлена рядом условий, которые обеспечивают корректное решение задачи. К сожалению, не в каждом источнике, который используется в качестве руководства исследователем, находят отражение эти условия. Вследствие этого, несмотря на кажущуюся простоту, практика использования критериев согласия типа χ^2 изобилует примерами его некорректного или неэффективного применения, особенно при проверке сложных гипотез.

Анализ примеров "неудачного" применения критериев типа χ^2 позволяет выделить две группы причин, которые на практике могут приводить к неверным статистическим выводам. Во-первых, это часто совершаемые принципиальные ошибки, при которых использование в качестве предельного χ^2_{k-m-1}-распределения оказывается неправомерным. Такое происходит, когда, проверяя сложные гипотезы, применяют оценки, в основе которых лежат исходные негруппированные данные. Во-вторых, действия, использующие возможности критерия не наилучшим образом (например нерациональный выбор числа интервалов и способа разбиения выборки на интервалы). В первом случае возрастает вероятность ошибки первого рода α (отклонить верную проверяемую гипотезу), во втором – вероятность ошибки второго рода β (принять проверяемую гипотезу при справедливости альтернативы).

При использовании критериев согласия типа χ^2 неоднозначность при построении и вычислении статистик бывает связана с выбором числа интервалов и тем, каким образом область определения случайной величины разбивается на интервалы. Естественно, что такой произвол отражается на статистических свойствах применяемых критериев, в частности, на их мощности при различении близких конкурирующих гипотез. Очевидно, что выбор числа интервалов и способа разбиения на интервалы следует осуществлять с позиций обеспечения максимальной мощности критерия.

С использованием критериев согласия могут проверяться простые гипотезы вида H_0:$F(x) = F_0(x, \theta)$, где $F_0(x, \theta)$ – функция распределения вероятностей, с которой проверяется согласие наблюдаемой выборки независимых одинаково распределенных величин $X_1, X_2, ..., X_n$, а θ – известное значение параметра (скалярного или векторного), и сложные гипотезы H_0: $F(x) \in \{F_0(x, \theta), \theta \in \Theta\}$, где Θ – пространство параметров. В процессе проверки сложной гипотезы оценка параметра $\hat{\theta}$ вычисля-

ется по этой же самой выборке. Если оценка $\hat{\theta}$ вычислена по другой выборке, то гипотеза простая.

При использовании критериев согласия типа χ^2 область определения случайной величины разбивается на n интервалов граничными точками:

$$x_1 < x_2 < ... < x_{k-1} < x_k.$$

Статистика X_n^2 Пирсона вычисляется в соответствии с соотношением

$$X_n^2 = n \sum_{i=1}^{k} \frac{(n_i/n - P_i(\theta))^2}{P_i(\theta)}, \tag{1}$$

где n_i – количество наблюдений, попавших в i-й интервал, $P_i(\theta) = \int\limits_{x_{i-1}}^{x_i} f_0(x,\theta)dx$ – вероятность попадания наблюдения в i-й интервал, $n = \sum\limits_{i=1}^{k} n_i$, $\sum\limits_{i=1}^{k} P_i(\theta) = 1$. При справедливой простой гипотезе H_0 предельное распределение статистики $G(X_n^2|H_0)$ есть χ^2-распределение с числом степеней свободы $k-1$. Если по выборке оценивалось m параметров закона в результате минимизации статистики X_n^2, статистика подчиняется χ^2-распределению с $k-m-1$ степенями свободы. При справедливой альтернативной гипотезе H_1 предельное распределение $G(X_n^2|H_1)$ представляет собой нецентральное χ^2-распределение с тем же числом степеней свободы и параметром нецентральности

$$s(\theta) = \sum_{i=1}^{k} \frac{c_i^2(\theta)}{P_i}, \tag{2}$$

где $c_i(\theta) = \sqrt{n} \int\limits_{x_{i-1}(\theta)}^{x_i(\theta)} (f_1(x,\theta) - f_0(x,\theta))dx$ и $f_1(x,\theta)$ соответствуют альтернативе.

В случае проверки сложных гипотез и оценивания по выборке параметров наблюдаемого закона использование в качестве предельных χ^2_{k-m-1}-распределений справедливо лишь при определении оценок минимизацией статистики X_n^2 или при вычислении по сгруппированным данным ряда оценок, в том числе оценок максимального правдоподобия (ОМП). Распределения статистик χ^2 Пирсона и отношения правдоподобия при оценивании параметров наблюдаемого закона по точечным наблюдениям и различных способах построения интервалов нами были исследованы методами компьютерного моделирования. Было показано,

что распределения $G(X_n^2|H_0)$ статистики оказываются зависящими от того, как строятся интервалы, и хорошо аппроксимируются семейством гамма-распределений.

Предложенная в работах С.М. Никулина [3, 4] статистика типа χ^2 отличается от X_n^2 только при сложных гипотезах. Предельное распределение этой статистики есть распределение χ^2_{k-1} (количество степеней свободы не зависит от числа оцениваемых параметров). Неизвестные параметры распределения $F(x,\theta)$ в этом случае должны оцениваться по негруппированным данным методом максимального правдоподобия. При этом вектор вероятностей попадания в интервал $P = (P_1, P_2, ..., P_k)^T$ предполагается заданным, и граничные точки интервалов определяются соотношениями

$$x_i(\theta) = F^{-1}(P_1 + P_2 + ... + P_i), \quad i = \overline{1, k-1}.$$

Данная статистика имеет вид

$$Y_n^2(\theta) = X_n^2 + n^{-1} a^T(\theta)\Lambda(\theta)a(\theta), \tag{3}$$

где X_n^2 вычисляются в соответствии с (1). Элементы и размерность матрицы

$$\Lambda(\theta) = \left[J(\theta_l, \theta_j) - \sum_{i=1}^{k} \frac{w_{\theta_l i} w_{\theta_j i}}{P_i} \right]^{-1}_{m \times m}$$

определяются оцениваемыми компонентами вектора параметров θ; $J(\theta_l, \theta_j)$ – элементы информационной матрицы

$$J(\theta) = \left[\int_{-\infty}^{\infty} \left(\frac{\partial \ln f_0(x,\theta)}{\partial \theta_l} \frac{\partial \ln f_0(x,\theta)}{\partial \theta_j} \right) f_0(x,\theta) dx \right]_{m \times m};$$

$a(\theta_l) = w_{\theta_l 1} n_1 / P_1 + ... + w_{\theta_l k} n_k / P_k$ – элементы вектора $a(\theta)$; величины $w_{\theta_l i}$ определяются соотношением

$$w_{\theta_l i} = -f_0[x_i(\theta), \theta] \frac{\partial x_i(\theta)}{\partial \theta_l} + f_0[x_{i-1}(\theta), \theta] \frac{\partial x_{i-1}(\theta)}{\partial \theta_l}.$$

При верной конкурирующей гипотезе статистика Y_n^2 имеет в качестве предельного $G(Y_n^2|H_1)$ нецентральное χ^2_{k-1}-распределение с параметром нецентральности

$$s(\theta) = \sum_{i=1}^{k} \frac{c_i^2(\theta)}{P_i} + d^T(\theta)\Lambda(\theta)d(\theta), \tag{4}$$

где элементы вектора $d(\theta)$ определяются как

$$d(\theta_l) = \frac{w_{\theta_l 1} c_1(\theta)}{P_1} + ... + \frac{w_{\theta_l k} c_k(\theta)}{P_k}.$$

Зависимость мощности от способа группирования. Способ группирования особенно сильное влияние оказывает на предельное распределение $G(X_n^2|H_1)$. Было показано, что критерии согласия χ^2 Пирсона и отношения правдоподобия при проверке как простых, так и сложных гипотез имеют максимальную мощность против близких альтернатив, если использовать такое разбиение области определения случайной величины на интервалы, при котором потери в информации Фишера о параметрах закона, соответствующего гипотезе H_0, минимальны (асимптотически оптимальное группирование). Чем меньше потери в информации Фишера, связанные с группированием данных, тем больше параметр нецентральности, определяемый соотношением (2). В [1] для конкретных законов распределения представлен достаточно широкий состав (более 50 для примерно 20 законов распределений) построенных таблиц асимптотически оптимального группирования, минимизирующего потери в информации Фишера. Использование асимптотически оптимального группирования при заданном числе интервалов обеспечивает максимальную мощность при близких гипотезах. Причем выигрыш в мощности даже по отношению к разбиению на интервалы равной вероятности очень существенный.

Исследование распределений статистики Y_n^2 Никулина, которая отличается от X_n^2 только при сложных гипотезах, показало, что как $G(Y_n^2|H_0)$, так и $G(Y_n^2|H_1)$ несущественно зависят от способа группирования. Более того, наши исследования методами статистического моделирования показали, что с позиций наибольшей мощности разбиение на интервалы равной вероятности (равновероятное группирование) оказывается наиболее предпочтительным. Подчеркнем, что критерий типа χ^2 Никулина мощнее, чем критерии χ^2 Пирсона и отношения правдоподобия.

Зависимость мощности от числа интервалов k. За всю историю применения критериев типа χ^2 была предложена не одна формула для выбора числа интервалов, но ни одна из представленных в различных рекомендациях не выводилась с позиций максимальной мощности применяемого критерия, а в основном исходя из близости плотности к ее непараметрической оценке, гистограмме. Зная предельные распределения $G(S|H_0)$ и $G(S|H_1)$ статистики S, для любого заданного уровня значимости α можно оценить мощность соответствующего критерия,

что распределения $G(X_n^2|H_0)$ статистики оказываются зависящими от того, как строятся интервалы, и хорошо аппроксимируются семейством гамма-распределений.

Предложенная в работах С.М. Никулина [3, 4] статистика типа χ^2 отличается от X_n^2 только при сложных гипотезах. Предельное распределение этой статистики есть распределение χ^2_{k-1} (количество степеней свободы не зависит от числа оцениваемых параметров). Неизвестные параметры распределения $F(x,\theta)$ в этом случае должны оцениваться по негруппированным данным методом максимального правдоподобия. При этом вектор вероятностей попадания в интервал $P = (P_1, P_2, ..., P_k)^T$ предполагается заданным, и граничные точки интервалов определяются соотношениями

$$x_i(\theta) = F^{-1}(P_1 + P_2 + ... + P_i), \quad i = \overline{1, k-1}.$$

Данная статистика имеет вид

$$Y_n^2(\theta) = X_n^2 + n^{-1} a^T(\theta)\Lambda(\theta)a(\theta), \tag{3}$$

где X_n^2 вычисляются в соответствии с (1). Элементы и размерность матрицы

$$\Lambda(\theta) = \left[J(\theta_l, \theta_j) - \sum_{i=1}^{k} \frac{w_{\theta_l i} w_{\theta_j i}}{P_i} \right]^{-1}_{m \times m}$$

определяются оцениваемыми компонентами вектора параметров θ; $J(\theta_l, \theta_j)$ – элементы информационной матрицы

$$J(\theta) = \left[\int_{-\infty}^{\infty} \left(\frac{\partial \ln f_0(x,\theta)}{\partial \theta_l} \frac{\partial \ln f_0(x,\theta)}{\partial \theta_j} \right) f_0(x,\theta) dx \right]_{m \times m};$$

$a(\theta_l) = w_{\theta_l 1} n_1 / P_1 + ... + w_{\theta_l k} n_k / P_k$ – элементы вектора $a(\theta)$; величины $w_{\theta_l i}$ определяются соотношением

$$w_{\theta_l i} = -f_0[x_i(\theta), \theta] \frac{\partial x_i(\theta)}{\partial \theta_l} + f_0[x_{i-1}(\theta), \theta] \frac{\partial x_{i-1}(\theta)}{\partial \theta_l}.$$

При верной конкурирующей гипотезе статистика Y_n^2 имеет в качестве предельного $G(Y_n^2|H_1)$ нецентральное χ^2_{k-1}-распределение с параметром нецентральности

$$s(\theta) = \sum_{i=1}^{k} \frac{c_i^2(\theta)}{P_i} + d^T(\theta)\Lambda(\theta)d(\theta), \tag{4}$$

где элементы вектора $d(\theta)$ определяются как

$$d(\theta_l) = \frac{w_{\theta_l 1} c_1(\theta)}{P_1} + ... + \frac{w_{\theta_l k} c_k(\theta)}{P_k}.$$

Зависимость мощности от способа группирования. Способ группирования особенно сильное влияние оказывает на предельное распределение $G(X_n^2|H_1)$. Было показано, что критерии согласия χ^2 Пирсона и отношения правдоподобия при проверке как простых, так и сложных гипотез имеют максимальную мощность против близких альтернатив, если использовать такое разбиение области определения случайной величины на интервалы, при котором потери в информации Фишера о параметрах закона, соответствующего гипотезе H_0, минимальны (асимптотически оптимальное группирование). Чем меньше потери в информации Фишера, связанные с группированием данных, тем больше параметр нецентральности, определяемый соотношением (2). В [1] для конкретных законов распределения представлен достаточно широкий состав (более 50 для примерно 20 законов распределений) построенных таблиц асимптотически оптимального группирования, минимизирующего потери в информации Фишера. Использование асимптотически оптимального группирования при заданном числе интервалов обеспечивает максимальную мощность при близких гипотезах. Причем выигрыш в мощности даже по отношению к разбиению на интервалы равной вероятности очень существенный.

Исследование распределений статистики Y_n^2 Никулина, которая отличается от X_n^2 только при сложных гипотезах, показало, что как $G(Y_n^2|H_0)$, так и $G(Y_n^2|H_1)$ несущественно зависят от способа группирования. Более того, наши исследования методами статистического моделирования показали, что с позиций наибольшей мощности разбиение на интервалы равной вероятности (равновероятное группирование) оказывается наиболее предпочтительным. Подчеркнем, что критерий типа χ^2 Никулина мощнее, чем критерии χ^2 Пирсона и отношения правдоподобия.

Зависимость мощности от числа интервалов k. За всю историю применения критериев типа χ^2 была предложена не одна формула для выбора числа интервалов, но ни одна из представленных в различных рекомендациях не выводилась с позиций максимальной мощности применяемого критерия, а в основном исходя из близости плотности к ее непараметрической оценке, гистограмме. Зная предельные распределения $G(S|H_0)$ и $G(S|H_1)$ статистики S, для любого заданного уровня значимости α можно оценить мощность соответствующего критерия,

рассматривая $e_{\text{ч}}$ как функцию от числа интервалов k при заданном объеме выборки n. Исследование мощности критериев Пирсона и Никулина как функций от n и k проводилось аналитически и методами статистического моделирования. Результаты аналитических вычислений были полностью подтверждены оценками, полученными на основании моделирования.

Мощность критериев типа χ^2 определяется выражением [5]

$$1-\beta = P(s|r,\alpha) = \exp^{-s/2} \sum_{j=0}^{\infty} \frac{s^j}{j!2^{2j-1+r/2}\Gamma(j+r/2)} \times$$

$$\times \int\limits_{\sqrt{x(\alpha,r)}}^{\infty} y^{2j-1+r} \exp^{-y^2/2}\, dy, \qquad (5)$$

где s – параметр нецентральности, определяемый соотношениями (2) и (4), $x(\alpha,r)$ представляет собой $(1-\alpha)$-процентную точку χ^2-распределения с r степенями свободы (α – заданная вероятность ошибки первого рода, β – вероятность ошибки второго рода). Приводимые ниже функции мощности строились при уровне значимости $\alpha = 0{,}1$.

Анализ функций мощности для различных альтернатив при проверке простых и сложных гипотез показал, что с увеличением числа интервалов мощность критериев типа χ^2 падает. Это соответствует результатам работ [6,7]. Максимальная мощность критериев при заданном объеме выборки n чаще всего достигается или при минимальном числе интервалов, или при некотором оптимальном значении k. Оптимальное число интервалов k зависит от объема выборки n и от конкретной пары конкурирующих гипотез H_0 и H_1. Чаще всего оптимальное k оказывается существенно меньше значений, рекомендуемых различными регламентирующими документами и задаваемых множеством эмпирических формул.

Таким образом, увеличение мощности критериев χ^2 Пирсона и отношения правдоподобия возможно за счет двух факторов: за счет выбора асимптотически оптимального группирования в качестве способа разбиения области определения случайной величины и подбора оптимального числа интервалов k при заданном объеме выборки n. Увеличение мощности критерия типа χ^2 Никулина возможно только за счет выбора оптимального числа интервалов.

Рассматривая пару альтернатив, всегда можно выбрать оптимальное число интервалов и подобрать оптимальное разбиение на интерва-

лы, в результате чего будем иметь критерий максимальной мощности, наилучшим образом различающий данные конкурирующие гипотезы.

3. Применение непараметрических критериев согласия при проверке сложных гипотез

К наиболее используемым критериям согласия относятся непараметрические критерии типа Колмогорова, типа ω^2 и Ω^2 Мизеса. В критерии Колмогорова в качестве расстояния между эмпирическим и теоретическим законом используется величина

$$D_n = \sup_{|x|<\infty} |F_n(x) - F(x,\theta)|,$$

где $F_n(x)$ – эмпирическая функция распределения, $F(x,\theta)$ – теоретическая функция распределения, n – объчм выборки. При проверке гипотез обычно используется статистика вида [5]

$$S_K = \frac{6nD_n + 1}{6\sqrt{n}}, \text{ где } D_n = \max(D_n^+, D_n^-),$$

$$D_n^+ = \max_{1\le i\le n}\left\{\frac{i}{n} - F(x_i,\theta)\right\},\ D_n^- = \max_{1\le i\le n}\left\{F(x_i,\theta) - \frac{i-1}{n}\right\},$$

n – объем выборки, $x_1, x_2, ..., x_n$ – упорядоченные по возрастанию выборочные значения, $F(x,\theta)$ – функция закона распределения, согласие с которым проверяется. Распределение величины S_K при простой гипотезе в пределе подчиняется закону Колмогорова $K(S)$ [5].

В критериях типа ω^2 расстояние между гипотетическим и истинным распределениями рассматривается в квадратичной метрике

$$\int_{-\infty}^{\infty} \{E[F_n(x)] - F(x)\}^2 \psi(F(x)) dF(x),$$

где $E[\star]$ – оператор математического ожидания.

При выборе $\psi(t) = 1$ в критериях типа ω^2 Мизеса пользуются статистикой (статистика Крамера–Мизеса–Смирнова) вида

$$S_\omega = \frac{1}{12n} + \sum_{i=1}^{n}\left\{F(x_i,\theta) - \frac{2i-1}{2n}\right\}^2,$$

при простой гипотезе подчиняющейся распределению $a1(S)$ [5].

При выборе $\psi(t) = 1/t(1-t)$ в критериях типа Ω^2 Мизеса применяемая статистика (статистика Андерсона–Дарлинга) имеет вид

$$S_\Omega = -n - 2\sum_{i=1}^{n}\left\{\frac{2i-1}{2n}\ln F(x_i,\theta) + \left(1 - \frac{2i-1}{2n}\right)\ln(1 - F(x_i,\theta))\right\}.$$

В пределе эта статистика подчиняется распределению $a2(S)$[5].

В случае простых гипотез предельные распределения статистик непараметрических критериев типа Колмогорова, ω^2 и Ω^2 Мизеса известны давно и не зависят от вида наблюдаемого закона распределения и значений его параметров. Говорят, что эти критерии являются "свободными от распределения". Это достоинство предопределяет широкое использование данных критериев в приложениях.

При проверке сложных гипотез, когда по той же самой выборке оцениваются параметры наблюдаемого закона $F(x,\theta)$, непараметрические критерии согласия теряют свойство "свободы от распределения". Однако мощность непараметрических критериев при проверке сложных гипотез при тех же объемах выборок n всегда существенно выше, чем при проверке простых. И если при проверке простых гипотез непараметрические критерии типа Колмогорова, ω^2 и Ω^2 Мизеса уступают по мощности критериям типа χ^2, при условии, что в последних используется асимптотически оптимальное группирование, то при проверке сложных гипотез непараметрические критерии оказываются более мощными. Для того чтобы воспользоваться их преимуществами, надо только знать распределение $G(S|H_0)$ для проверяемой сложной гипотезы.

Различия в предельных распределениях тех же самых статистик при проверке простых и сложных гипотез очень существенны. Поэтому предостережения против неаккуратного применения критериев согласия при проверке сложных гипотез неоднократно поднимались на страницах печати. При проверке сложных гипотез на условный закон распределения статистики $G(S|H_0)$ влияет целый ряд факторов, определяющих "сложность" гипотезы: вид наблюдаемого закона $F(x,\theta)$, соответствующего истинной гипотезе H_0; тип оцениваемого параметра и количество оцениваемых параметров; в некоторых ситуациях конкретное значение параметра (например в случае гамма-распределения); используемый метод оценивания параметров и точность вычисления оценок.

Исходной точкой для исследований предельных распределений статистик непараметрических критериев согласия при сложных гипотезах послужила работа [8]. В литературных источниках изложен ряд подходов к использованию непараметрических критериев согласия в случае

проверки сложных гипотез. При достаточно большом объеме выборки ее можно разбить на две части, и по одной из них оценивать параметры, а по другой проверять согласие [9]. В некоторых частных случаях предельные распределения статистик исследовались аналитическими методами [10], процентные точки распределений строились методами статистического моделирования [11–14]. Для приближенного вычисления вероятностей "согласия" вида $P\{S > S^*\}$ (достигаемого уровня значимости) строились формулы, дающие достаточно хорошие приближения при малых значениях соответствующих вероятностей [15–17]. В наших работах исследование распределений статистик непараметрических критериев согласия и построение моделей этих распределений осуществлялись с использованием методики компьютерного анализа статистических закономерностей.

Наши исследования показали, что распределения статистик критериев согласия существенно зависят от метода оценивания параметров. Вообще говоря, каждому типу оценок при конкретной сложной проверяемой гипотезе соответствует свой предельное распределение $G(S|H_0)$ статистики. В случае метода максимального правдоподобия распределения статистик $G(S|H_0)$ очень сильно зависят от закона, соответствующего гипотезе H_0. В то же время, разброс распределений $G(S|H_0)$ при использовании MD-оценок, минимизирующих статистику критерия, в существенно меньшей степени зависит от вида закона $F(x,\theta)$, соответствующего гипотезе H_0. Это позволяет говорить об определенной "свободе от распределения" для рассматриваемых критериев. Если опираться только на этот факт, то, казалось бы, что только такие методы оценивания и следует применять при проверке сложных гипотез. Однако исследование мощности рассматриваемых критериев при различных методах оценивания показало, что наибольшую мощность данные критерии при близких альтернативах имеют в случае использования ОМП. На рис. 1 иллюстрируется изменение распределений $G(S_K|H_0)$ статистики типа Колмогорова S_K в зависимости от числа оцениваемых параметров закона распределения Su-Джонсона с плотностью

$$f(x) = \frac{\theta_1}{\sqrt{2\pi}\sqrt{(x-\theta_3)^2+\theta_2^2}} \times$$
$$\times \exp\left\{-\frac{1}{2}\left[\theta_0 + \theta_1 \ln\left\{\frac{x-\theta_3}{\theta_2} + \sqrt{\left(\frac{x-\theta_3}{\theta_2}\right)^2 + 1}\right\}\right]^2\right\}.$$

На рисунке кривая *1* – функция распределения Колмогорова, ко-

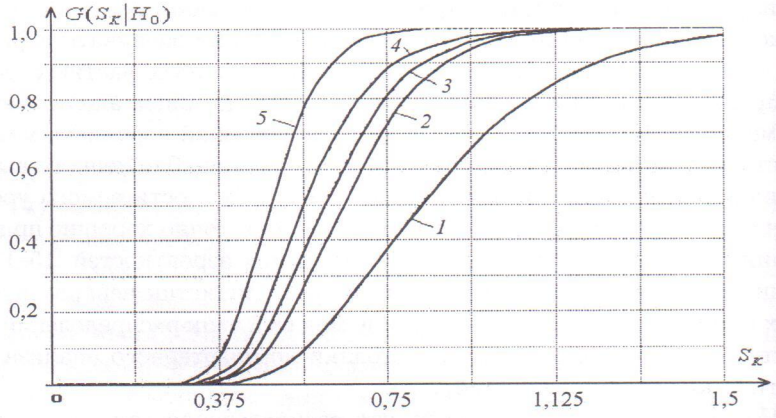

Рис. 1. Изменение распределений статистики типа Колмогорова в зависимости от числа оцениваемых параметров наблюдаемого распределения Su-Джонсона

торому подчиняется статистика при проверке простой гипотезы, *2* – распределение статистики при проверке сложной гипотезы и оценивании по данной выборке методом максимального правдоподобия только параметра θ_0, *3* – при оценивании только параметров θ_0 и θ_1, *4* – при оценивании параметров θ_0, θ_1 и θ_2, *5* – при оценивании всех четырех параметров θ_0, θ_1, θ_2 и θ_3.

Распределения статистик непараметрических критериев согласия существенно зависят от вида и числа оцениваемых параметров. И даже при оценивании единственного параметра предельное распределение статистики резко отличается от предельного распределения той же самой статистики в случае проверки простой гипотезы. Различие возрастает с увеличением числа оцениваемых параметров. Пренебрежение этим фактом в практике применения критериев согласия приводит к большим ошибкам в вычислении вероятности вида $P\{S > S^*\}$ и неоправданному принятию проверяемой гипотезы.

При малых объемах выборки n распределения $G(S_n|H_0)$ зависят от n. Однако существенная зависимость распределения статистик от n наблюдается только при небольших объемах выборки. Как показали наши исследования, при $n \geq 15 - 20$ распределения $G(S_n|H_0)$ достаточно близки к предельным $G(S|H_0)$, и зависимостью от n можно пренебречь.

Построенные на настоящий момент другими исследователями таблицы процентных точек и предельные распределения статистик непараметрических критериев ограничены относительно узким кругом сложных гипотез. Это объясняется тем, что построение предельного распределения аналитическими методами выливается в чрезвычайно непростую задачу. В то же время методика компьютерного анализа статистических закономерностей, хорошо зарекомендовавшая себя при моделировании распределений статистик критериев, позволяет при необходимости без существенных трудностей расширить этот круг.

В рекомендациях [2] нами были построены модели предельных распределений статистик рассматриваемых критериев при проверке сложных гипотез относительно 17 различных законов наблюдаемых случайных величин при использовании ОМП и MD-оценок. Это в общей сложности 350 моделей предельных распределений для такого же количества вариантов сложных гипотез.

В соответствии с этой методикой следует по закону $F(x, \hat{\theta})$ смоделировать N выборок того же объема n, что и выборка, для которой необходимо проверить гипотезу H_0: $F(x) \in \{F(x, \theta), \theta \in \Theta\}$. Далее для каждой из N выборок вычислить оценки тех же параметров закона, а затем значение статистики S соответствующего критерия. В результате будет получена выборка значений статистики S_1, S_2,...,S_N с законом распределения $G(S_n|H_0)$ для проверяемой гипотезы H_0. По этой выборке при значительных N можно построить достаточно гладкую эмпирическую функцию распределения $G_N(S_n|H_0)$, которой можно непосредственно воспользоваться для вывода, следует ли принимать гипотезу H_0. При необходимости можно по $G_N(S_n|H_0)$ построить приближенную аналитическую модель, аппроксимирующую $G_N(S_n|H_0)$, и тогда уже, опираясь на эту модель, принимать решение относительно проверяемой гипотезы.

Как показали исследования, хорошей аналитической моделью для $G_N(S_n|H_0)$ часто оказывается один из следующих четырех законов: гамма-распределение, распределение Su-Джонсона, распределение Sl-Джонсона или логарифмически нормальное. В крайнем случае, всегда можно, опираясь на ограниченное множество законов распределения, построить модель в виде смеси законов.

Построенные и представленные в [2] аппроксимации предельных распределений статистик непараметрических критериев согласия расширяют область корректного применения этих критериев и рекомендуются широкому кругу исследователей. Расширение множества моделей распределений статистик, то есть расширение множества корректно про-

веряемых сложных гипотез, не является самоцелью. Это имеет смысл только для тех законов распределения случайных величин, которые наиболее часто используются в приложениях. Именно для этих законов следует продолжить исследования. В целом же мы имеем дело с бесконечными множествами существующих случайных величин, их законов, возможных формулировок сложных гипотез. Следовательно, речь должна идти о постепенном *создании информационных технологий* исследования статистических закономерностей, технологий обеспечивающих статистический анализ и обработку экспериментальных наблюдений в различных областях знаний и человеческой жизнедеятельности.

4. Заключение

Основу ГОСТ Р 50.1-033-2001 составляют полученные таблицы асимптотически оптимального группирования, обеспечивающие максимальную мощность критериев типа χ^2 при простых и сложных гипотезах, и рекомендации по их применению.

Содержание ГОСТ Р 50.1-037-2002 определяют построенные модели распределений статистик непараметрических критериев согласия и соответствующие таблицы процентных точек, обеспечивающие корректные статистические выводы при проверке сложных гипотез.

СПИСОК ЛИТЕРАТУРЫ

1. *ГОСТ Р 50.1-033-2001* Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Часть I. Критерии типа χ^2.
2. *ГОСТ Р 50.1-037-2002* Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим. Часть II. Непараметрические критерии.
3. *Никулин М.С.* Критерий хи-квадрат для непрерывных распределений с параметрами сдвига и масштаба // Теория вероятностей и ее применение. 1973. Т. XVIII. № 3. С. 583-591.
4. *Никулин М.С.* О критерии хи-квадрат для непрерывных распределений // Теория вероятностей и ее применение. 1973. Т. XVIII. № 3. С. 675-676.
5. *Большев Л.Н., Смирнов Н.В.* Таблицы математической статистики. М.: Наука, 1983. 416 с.
6. *Чибисов Д.М., Гванцеладзе Л.Г.* О критериях согласия, основанных на группированных данных // III советско-японский симпозиум по теории вероятностей. Ташкент: Фан, 1975. С. 183-185.
7. *Боровков А.А.* О мощности критерия χ^2 при увеличении числа групп // Теория вероятностей и ее применение. 1977. Т. XXII. № 2. С. 375-378.

8. *Kac M., Kiefer J., Wolfowitz J.* On tests of normality and other tests of goodness of fit based on distance methods // Ann. Math. Stat. 1955. Vol. 26. P. 189-211.
9. *Durbin J.* Kolmogoriv–Smirnov test when parameters are estimated // Lect. Notes Math. 1976. Vol. 566. P. 33-44.
10. *Мартынов Г.В.* Критерии омега-квадрат. М.: Наука, 1978. 80 с.
11. *Pearson E.S., Hartley H.O.* Biometrica tables for statistics. Vol. 2. Cambridge: University Press, 1972. 634 p.
12. *Stephens M.A.* Use of Kolmogorov–Smirnov, Cramer – von Mises and related statistics – vithout extensive table // J. R. Stat. Soc. 1970. Vol. 32. P. 115-122.
13. *Stephens M.A.* EDF statistics for goodness of fit and some comparisons // J. Am. Statist. Assoc. 1974. Vol. 69. P. 730-737.
14. *Chandra M., Singpurwalla N.D., Stephens M.A.* Statistics for test of fit for the extrem-value and weibull distribution // J. Am. Statist. Assoc. 1981. Vol. 76. P. 375.
15. *Тюрин Ю.Н.* О предельном распределении статистик Колмогорова–Смирнова для сложной гипотезы // Изв. АН СССР. Сер. Математич. 1984. Т. 48, № 6. С. 1314-1343.
16. *Тюрин Ю.Н., Саввушкина Н.Е.* Критерии согласия для распределения Вейбулла–Гнеденко // Изв. АН СССР. Сер. Техн. кибернетика. 1984. № 3. С. 109-112.
17. *Саввушкина Н.Е.* Критерий Колмогорова–Смирнова для логистического и гамма-распределения // Сб. тр. ВНИИ систем. исслед. 1990. № 8. С. 50-56.

Г. А. Медведев

Непараметрические методы оценивания нейтральных к риску вероятностей по данным рыночных наблюдений

Настоящая статья представляет собой краткий обзор результатов по непараметрическим методам оценивания нейтральных к риску распределений вероятностей с помощью наблюдения цен финансовых производных (в основном опционов европейского типа) в дату их истечения. Статья содержит в основном не описание методов, а информацию о задачах, которые решаются с их помощью, с соответствующими ссылками.