$$\begin{cases} (a_2\Theta z_0)\circ(x_0^{(2)}\Theta x_0^{(1)})=0 \\ (x_0^{(2)}\Theta x_0^{(1)})\circ(x_j^{(2)}\circ h\circ(\bigoplus_{r=0}^{j-1}(1\Theta h\circ a_1)^r\circ(1\Theta h\circ a_3)^{j-1-r}\circ y_r\oplus(z_j^{(1)}\Theta a_2)\circ(1\Theta h\circ a_1)^j) \quad (j=1,\ldots,l-1) \end{cases}.$$

В заключение отметим, что условия $h\cdot a_1\not\equiv 1(\bmod q)$ и $h\cdot a_1\not\equiv 1(\bmod q)$ эквивалентны тому, что $1\Theta h\cdot a_1$ и $1\Theta h\cdot a_1$ – обратимые элементы кольца $\mathbf{Z}_{q^k}$. Следовательно, $(1\Theta h\cdot a_1)^0=1$ и $(1\Theta h\cdot a_1)^0=1$. Однако, если $\alpha\in\mathbf{Z}_{q^k}$ необратимый элемент кольца $\mathbf{Z}_{q^k}$, то α^0 не определено.

Б.Ю. ЛЕМЕШКО, С.Б. ЛЕМЕШКО, С.Н. ПОСТОВАЛОВ
(Новосибирский государственный технический университет)

СРАВНИТЕЛЬНЫЙ АНАЛИЗ МОЩНОСТИ КРИТЕРИЕВ СОГЛАСИЯ

Введение. Применение критериев согласия до сих пор вызывает разногласия по поводу оценки их мощности. Различие мнений обусловлено, с одной стороны, наличием узких мест в применении критерия χ^2 Пирсона, связанных с произволом в выборе способа группирования и числа интервалов, а с другой, – с отсутствием знаний (или ограниченными знаниями) о распределениях статистик критериев согласия при проверке сложных гипотез. Применяющих статистические критерии всегда интересуют вероятности возможных ошибок. В данном случае нами исследовались распределения и мощность критериев χ^2 Пирсона, типа χ^2 Никулина, непараметрических критериев Колмогорова, ω^2 Крамера-Мизеса, Ω^2 Андерсона-Дарлинга

При проверке статистических гипотез ошибка 1-го рода заключается в том, что в результате проверки отклоняется справедливая проверяемая гипотеза H_0. Ошибка 2-го рода – в признании верной гипотезы H_0, когда на самом деле справедлива некоторая конкурирующая гипотеза H_1.

Процедура проверки гипотезы H_0 предполагает, что известно условное распределение $G(S|H_0)$ статистики S применяемого критерия при справедливости H_0. Для критериев согласия критические области определяются большими значениями статистик. Вероятность ошибки 1-го рода α представляет собой вероятность попадания значения статистики в критическую область: $\alpha=P\{S>S_\alpha|H_0\}=1-G(S_\alpha|H_0)$, где S_α – критическое значение. Величина α, как правило, задается. Если вычисленное по выборке значение статистики $S^*\le S_\alpha$, то проверяемая гипотеза H_0 не отклоняется. Знание $G(S|H_0)$ позволяет по значению S^* найти $P\{S>S^*|H_0\}=1-G(S^*|H_0)$ – достигнутый уровень значимости. Проверяемая гипотеза H_0 не отклоняется при $P\{S>S^*|H_0\}>\alpha$.

Если задана альтернатива H_1, то вероятность ошибки 2-го рода определяется соотношением $\beta=P\{S\le S_\alpha|H_1\}=G(S_\alpha|H_1)$, где $G(S|H_1)$ – распределение статистики критерия при справедливости H_1. Если критерий полностью определен, то задание α однозначно определяет величину β и наоборот. Величина $1-\beta$ называется мощностью критерия при проверке гипотезы H_0 относительно H_1 и является функцией, зависящей от H_0, H_1, объема выборки n и, возможно, от некоторых других факторов, связанных с построением критерия.

Таким образом, чтобы обеспечить при проверке заданную вероятность ошибки 1-го рода α надо знать условное распределение $G(S|H_0)$, а для обеспечения заданной вероятности ошибки 2-го рода необходимо дополнительно знать распределение $G(S_n|H_1)$.

Существующие проблемы. Следует различать проверку простых и сложных гипотез. Простая проверяемая гипотеза имеет вид $H_0: F(x) = F(x,\theta)$, где $F(x,\theta)$ – функция распределения вероятностей, с которой проверяют согласие наблюдаемой выборки, а θ – известное значение параметра (скалярного или векторного).

Сложная проверяемая гипотеза может быть записана в виде $H_0: F(x) \in \{F(x,\theta), \theta \in \Theta\}$, где Θ – область определения неизвестного параметра θ. Отличие в применении критериев при проверке сложных гипотез и соответствующие проблемы возникают, если оценку параметра $\hat{\theta}$ теоретического распределения вычисляют по той же самой выборке, по которой проверяют согласие.

Для применяемых критериев согласия в случае простых проверяемых гипотез, как правило, известно либо предельное распределение $G(S|H_0)$ статистики, либо процентные точки. Проблемы могут заключаться в том, что при ограниченных объемах выборок $G(S_n|H_0)$ значимо отличаются от предельного, а, следовательно, вероятности ошибок 1-го рода будут отличаться от заданного α.

При проверке сложных гипотез предельное распределение $G(S|H_0)$ для критерия χ^2 Пирсона известно только при оценивании r компонент вектора параметров закона $F(x,\theta)$ по группированным данным (χ_{k-r-1}-распределение), для критерия Никулина при использовании оценок максимального правдоподобия по негруппированным данным (χ_{k-1}-распределение).

Для непараметрических критериев согласия распределения $G(S|H_0)$ при сложных гипотезах зависят от ряда факторов, определяющих "сложность" гипотезы: вид наблюдаемого закона распределения $F(x,\theta)$, соответствующего истинной гипотезе H_0; тип оцениваемого параметра и число оцениваемых параметров; в некоторых ситуациях конкретное значение параметра; используемый метод оценивания параметров. Поэтому распределения $G(S|H_0)$ статистик или процентные точки (аналитически или численно) получены лишь для некоторых частных случаев сложных гипотез.

О виде распределений $G(S\,H_1)$ известно только для критериев типа χ^2. Однако и в этом случае вид этих распределений, а, следовательно, и вероятности β, зависят от числа интервалов и способа их построения. О виде же распределений статистик $G(S\,H_1)$ для непараметрических критериев ничего не известно.

Методы исследования. Для оценки вероятностей ошибок α и β (для построения приближений $G(S\,H_0)$ и $G(S\,H_1)$) наиболее перспективно применение метода статистических испытаний.

Для построения $G(S|H_i)$, $i = 0, 1$, следует смоделировать N выборок объема n по закону, соответствующему справедливой гипотезе H_i. Далее в случае сложной гипотезы по выборке (соответствующим методом) оцениваются параметры закона, соответствующего проверяемой гипотезе H_0. После этого для каждой из N выборок вычисляется значение статистики S интересующего нас критерия согласия. В результате будет получена выборка значений статистики $S_1, S_2, ..., S_N$ с эмпирическим законом распределения $G_N(S_n|H_i)$. При $N = 10^6$ точность определения достигаемого уровня значимости $P\{S > S^*|H_0\}$ или вероятности ошибки 2-го рода β в соответствии с $G_N(S_n|H_i)$ при проверке простых гипотез будет не хуже чем $\pm 10^{-3}$. Опираясь на $G_N(S_n|H_i)$, можно построить приближенную аналитическую модель условного распределения.

Результаты. Методами статистических испытаний исследованы распределения $G(S|H_0)$ и $G(S|H_1)$, оценены вероятности β и мощности критериев для некоторых пар достаточно близких альтернатив H_0 и H_1.

В качестве примера далее приведены оценки мощности (при проверке простых гипотез) критериев согласия для пары альтернатив, которую составили нормальный и логистический законы: проверяемой гипотезе H_0 соответствовал нормальный закон с плотностью

$f(x) = \dfrac{1}{\theta_0\sqrt{2\pi}}\exp\left\{-\dfrac{(x-\theta_1)^2}{2\theta_0^2}\right\}$, а конкурирующей гипотезе H_1 – логистический с функцией плотности $f(x) = \dfrac{\pi}{\theta_0\sqrt{3}}\exp\left\{-\dfrac{\pi(x-\theta_1)}{\theta_0\sqrt{3}}\right\} \Big/ \left[1+\exp\left\{-\dfrac{\pi(x-\theta_1)}{\theta_0\sqrt{3}}\right\}\right]^2$ и параметрами $\theta_0 = 1$, $\theta_1 = 0$. В случае простой гипотезы H_0 параметры нормального закона имеют те же значения. Эти два закона близки и трудно различаются с помощью критериев согласия.

Таблица

Мощность критериев при проверке простой гипотезы H_0 (нормальное распределение) против альтернативы H_1 (логистическое)

α	n=100	n=200	n=300	n=500	n=1000	n=2000
	Мощность критерия χ^2 Пирсона при $k=15$ и АОГ					
0,15	0.349	0.459	0.565	0.737	0.946	0.999
0,1	0.290	0.388	0.490	0.671	0.922	0.998
0,05	0.210	0.292	0.385	0.565	0.871	0.996
0,01	0.107	0.159	0.221	0.369	0.729	0.983
	Мощность критерия χ^2 Пирсона при $k=9$ и АОГ					
0,15	0.269	0.381	0.488	0.670	0.917	0.998
0,1	0.204	0.302	0.403	0.589	0.880	0.995
0,05	0.129	0.203	0.287	0.464	0.806	0.989
0,01	0.050	0.081	0.127	0.249	0.608	0.957
	Мощность критерия χ^2 Пирсона при $k=9$ и РВГ					
0,15	0.210	0.282	0.349	0.483	0.747	0.960
0,1	0.152	0.208	0.270	0.392	0.673	0.938
0,05	0.083	0.123	0.170	0.273	0.547	0.890
0,01	0.020	0.036	0.056	0.109	0.310	0.734
	Мощность критерия Ω^2 Андерсона-Дарлинга					
0,15	0.194	0.258	0.328	0.472	0.776	0.982
0,1	0.125	0.169	0.222	0.343	0.654	0.957
0,05	0.057	0.079	0.107	0.181	0.439	0.869
0,01	0.010	0.013	0.017	0.031	0.114	0.491
	Мощность критерия χ^2 Пирсона при $k=15$ и РВГ					
0,15	0.192	0.257	0.312	0.432	0.690	0.941
0,1	0.139	0.187	0.237	0.343	0.607	0.911
0,05	0.073	0.106	0.144	0.227	0.477	0.848
0,01	0.018	0.029	0.043	0.083	0.247	0.662
	Мощность критерия Колмогорова					
0,15	0.190	0.246	0.303	0.415	0.662	0.922
0,1	0.127	0.170	0.215	0.309	0.544	0.861
0,05	0.062	0.088	0.116	0.179	0.365	0.721
0,01	0.012	0.018	0.026	0.044	0.119	0.366
	Мощность критерия ω^2 Крамера-Мизеса-Смирнова					
0,15	0.178	0.228	0.283	0.401	0.680	0.947
0,1	0.114	0.147	0.186	0.277	0.542	0.892
0,05	0.052	0.067	0.086	0.136	0.324	0.742
0,01	0.010	0.011	0.014	0.021	0.065	0.307

Результаты анализа позволяют в случае простых проверяемых гипотез упорядочить критерии по мощности следующим образом:

χ^2 Пирсона $\succ \Omega^2$ Андерсона-Дарлинга $\succ$ ω^2 Мизеса $\succ$=Колмогорова.

Такая шкала справедлива **при использовании** в критерии χ^2 Пирсона **асимптотически оптимального группирования**, при котором минимизируются потери в информации Фишера. При очень близких гипотезах возможно: Колмогорова $\succ$ ω^2 Мизеса.

При проверке сложных гипотез критерии упорядочиваются следующим образом:

Ω^2 Андерсона-Дарлинга $\succ \omega^2$ Мизеса $\succ$ Никулина $\succ \chi^2$ Пирсона $\succ$ Колмогорова.

При очень близких гипотезах возможно:

Ω^2 Андерсона-Дарлинга $\succ$ Никулина $\succ \omega^2$ Мизеса $\succ \chi^2$ Пирсона $\succ$ Колмогорова.

Эти выводы носят интегрированный характер. Такое упорядочение не является абсолютно строгим: иногда конкретный критерий имеет преимущества в мощности при одних значениях α и объемах выборок n и уступает при других.

А.Б. ЩЕРБАНЬ
(Пензенская государственная технологическая академия)

ЗАДАЧИ IS – АНАЛИЗА ИНФОРМАЦИОННЫХ СТРУКТУР

Если упрощенную модель оцениваемой информационной структуры A_s^p обобщенно представить в виде кортежа $M_s^p = \langle E_s^p, V_s^p \rangle$, в котором E_s^p - множество элементов структуры A_s^p; $V_s^p = \left\{V_{s_1}^p, V_{s_2}^p, ... V_{s_e}^p\right\}$ - множество предикатов, описывающих множество видов отношений на множестве $E_s^p \times E_s^p$ и определяющих сигнатуру M_s^p, то в соответствии с принципом ситуационно-структурного управления (S - принципом) задача оценивания A_s^p может быть поставлена, как задача структурной идентификации M_s^p в пространстве структурных образов $p^0 = \left\{p_1^0, p_2^0, ... p_j^0, ... p_n^0\right\}$, которое представляет собой класс множеств структурных ситуаций $p_j^0 \subset p^0$, $p_j^0 = \left\{p_{j1}^0, p_{j2}^0, ... p_{jk}^0 ...\right\}$, объединяющих структурные состояния $p_{jk}^0 \in p_j^0$, детерминированные (параметрически идентифицированные) с точки зрения оценки, т.е. в пространстве образов, оценочные параметры которых определены однозначно.

Пространство p^0 можно условно назвать пространством реализации ситуационно-структурного управления (реализации S - принципа) и обобщенно представить в виде класса множеств $M^0 = \left\{M_1^0, M_2^0, ... M_j^0, ... M_n^0\right\}$, где $M_j^0 = \left\{M_{j1}^0, ... M_{jk}^0\right\}$ - множество частных моделей $M_{jk}^0 =< E_{jk}^0, V_{jk}^0 >$ параметрически идентифированных структурных состояний p_{jk}^0, описывающих некоторую детерминированную структурную ситуацию p_j^0.

Под структурной идентификацией будем понимать отображение $\varphi : A_s^p \to p^0$, где φ - способ (правило) формирования отображения; $\varphi\left(A_s^p\right)$ - образ идентифицируемого объекта A_s^p, в общем случае принадлежащего некоторому пространству объектов А, удовлетворяющий условию $\varphi\left(A_s^p\right) \in p^0$. Структура A_s^p называется идентифицируемой если существует $\varphi\left(A_s^p\right) \in p^0$.

В зависимости от способа отображения $\varphi \in \Phi$, где Φ - пространство отображений, и от целей оценивания идентифицируемой структуры, могут решаться различные задачи структурной идентификации, вся совокупность которых составляет совокупность задач идентификационно-структурного анализа (IS – анализа).