7174

СООБЩЕНИЯ
ОБЪЕДИНЕННОГО
ИНСТИТУТА
ЯДЕРНЫХ
ИССЛЕДОВАНИЙ

ДУБНА

Р5 - 7174

А.Пазман , И.Н.Силин

МОДИФИКАЦИЯ τ -КРИТЕРИЯ И r^2 -КРИТЕРИЙ СРАВНЕНИЯ ГИПОТЕЗ

1973

ЛАБОРАТОРИЯ ВЫЧИСЛИТЕЛЬНОЙ ТЕХНИКИ И АВТОМАТИЗАЦИИ

P5 - 7174

А.Пазман*, И.Н.Силин

МОДИФИКАЦИЯ τ -КРИТЕРИЯ И r^2 -КРИТЕРИЙ СРАВНЕНИЯ ГИПОТЕЗ

* Институт теории измерения САН, Братислава, ЧССР.

Введение

В /I/ одним из авторов был предложен критерий для статистической проверки линейной гипотезы против линейной альтернативы. Этот критерий применялся для дискриминации решений при неоднозначном фазовом анализе /2-6/.

Опыт использования этого критерия и исследование его свойств методом Монте-Карло другим автором показали необходимость некоторого видоизменения τ -критерия (точнее его использования). Этому посвящен первый параграф настоящей статьи. Во втором параграфе излагается сделанное А.Пазманом изменение τ -критерия для учета ограничений на область изменения искомых параметров, если таковые имеются.

В третьем параграфе излагается предложенный И.Силиным дополнительный критерий (τ^2 -критерий) для оценки мощности τ -критерия и для использования в тех случаях, когда τ -критерий отказывает.

§ I. Использование τ -критерия

Существенное предположение, что одна из двух сравнимых гипотез должна быть верной, не всегда пригодно на практике. Однако это предположение не используется при оценке τ -распределения, а лишь при проверке худшей из гипотез. В действительности вполне может оказаться, что обе гипотезы неверны или недостаточно точны.

К тому же анализ τ-критерия показал, что он мощнее, чем предполагалось, а τ может оказаться большим даже при малой разности χ^2 гипотез. Поэтому нужно проверять по τ-критерию обе гипотезы (меняя местами гипотезы в выражении для τ), а проверку вести не по величине τ, а по величине τ^2 (знак τ в общем случае нельзя однозначно предсказать). Проверка по τ^2 делается аналогично проверке по χ^2-распределению с одной степенью свободы (τ^2 распределено как χ^2_1). Может оказаться, что τ-критерий отвергнет обе гипотезы. В частности, при разности χ^2 проверяемых гипотез порядка единицы, большими могут стать только оба τ одновременно, если число параметров в обеих гипотезах равно.

При дискриминации гипотез полезно учитывать, что отбрасывание одной из гипотез еще не означает, что вторая много лучше. В дальнейшем может оказаться, что придется уточнять обе гипотезы (например, добавлять параметры) после чего лучшая гипотеза станет худшей. Возможно также, что истина одинаково далека от обеих гипотез и расположена между ними. Экспериментальные же данные вследствие статистических ошибок уклонились от истинного значения в сторону одной из гипотез и, соответственно, удалились от другой гипотезы. В результате τ одной гипотезы превзошло критическое значение и она отвергнута, а τ другой гипотезы не превзошло критического значения.

Однако можно утверждать, что преимущество одной гипотезы перед другой не случайно, если $\left|\frac{\tau_0-\tau_1}{2}\right| \gg 1$, т.е. вероятность $P\left(\chi^2_1 \geqslant \left(\frac{\tau_0-\tau_1}{2}\right)^2\right)$ мала. Не вдаваясь в подробности (геометрический смысл τ-критерия более подробно рассмотрен в § 3), можно сказать, что при одинаковых знаках τ эксперимент расположен между гипотезами, а при разных знаках τ – в стороне от обеих гипотез. В последнем случае гипотеза, ближе расположенная к эксперименту, имеет меньшее по модулю τ и с большой вероятностью будет иметь также меньшее τ после повторных экспериментов. Но если ее τ велико, то обе гипотезы плохие.

Есть ситуации, в которых τ-критерий теряет мощность и τ оказывается малым для очень плохих гипотез. В связи с этим одновременно с τ полезно вычислять z^2 (см. § 3). $z^2 \gg \tau^2$ может сигнализировать о такой ситуации. В принципе, можно использовать и δ-критерий /7-10/, но он слишком трудоемок.

Если число гипотез велико, то появляется возможность большого ($n!$) числа парных проверок. Не следует этим увлекаться, так как пропорционально числу проверок возрастает вероятность случайной большой флюктуации τ в одной из проверок, что обязательно следует учитывать.

§ 2. τ-критерий с учетом ограничений на параметры

Пусть $y_1, \dots, y_n$ – данные из эксперимента, которые распределены нормально с дисперсиями $\sigma_1^2, \dots, \sigma_n^2$ ж).

Проверяемая гипотеза H_0 утверждает, что среднее значение вектора

$$\vec{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

ж) Предполагается, что $y_1, \dots, y_n$ – независимые случайные величины, хотя все рассуждения вполне сохраняются и для зависимых величин; нужно только другое определение скалярного произведения (2.4) и матрицы К.

равно

$$\mathcal{E}(\vec{y}) = \vec{\eta}(\vec{\theta}) \quad ,$$

а гипотеза $\mathcal{H}_1$ утверждает, что

$$\mathcal{E}(\vec{y}) = \vec{\nu}(\vec{\Phi}) .$$

Здесь

$$\vec{\theta} = \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_q \end{pmatrix} , \quad \vec{\Phi} = \begin{pmatrix} \Phi_1 \\ \vdots \\ \Phi_m \end{pmatrix}; \quad m < n , \quad q < n$$

векторы неизвестных параметров и $\eta_i(\vec{\theta}), \nu_i(\vec{\Phi}); i=1,...,n-$ заданные функции от $\vec{\theta}$ и от $\vec{\Phi}$, которые определены на области допустимых значений параметров линейными формами:

$$\vec{\eta}(\vec{\theta}) = F\vec{\theta} + \vec{f} \quad ; \quad \vec{\theta} \in \Omega_0 \qquad 2.1$$

и

$$\vec{\nu}(\vec{\Phi}) = G\vec{\Phi} + \vec{g} \quad ; \quad \vec{\Phi} \in \Omega_1 . \qquad 2.2$$

В этих формулах F и G - заданные матрицы ранга q и m соответственно, а $\vec{f}$ и $\vec{g}$ - заданные векторы. Области допустимых значений параметров Ω_0 и Ω_1 определены неравенствами

$$\alpha_i^{(0)} \le \theta_i \le \beta_i^{(0)} \quad ; \quad i = 1, \dots, q$$

$$\alpha_i^{(1)} \le \Phi_i \le \beta_i^{(1)} \quad ; \quad i = 1, \dots, m . \qquad 2.3$$

В дальнейшем полезно воспользоваться следующим скалярным произведением в n -мерном пространстве возможных векторов $\vec{y}$:

$$\langle \vec{a} , \vec{b} \rangle = \sum_{i=1}^{n} a_i b_i / \sigma_i^2 \qquad 2.4$$

для любых n -мерных векторов $\vec{a}$ и $\vec{b}$. Соответственно норма $\|\vec{a}\|$ равна

$$\|\vec{a}\| = \sqrt{\langle \vec{a}, \vec{a} \rangle} = \sqrt{\sum_{i=1}^{n} a_i^2 / \sigma_i^2} .$$

В таких обозначениях плотность вероятности нормального случайного вектора $\vec{y}$, в предположении, что $\vec{\theta}_0$ -истинное значение параметров, равна

$$\frac{1}{(2\pi)^{n/2} \prod_{i=1}^{n} \sigma_i} e^{-\frac{1}{2}\|\vec{y} - \vec{\eta}(\vec{\theta}_0)\|} .$$

Поэтому условие максимума правдоподобия сводится к условию минимума нормы, которое имеет явный геометрический смысл (минимум "расстояния", т.е. оценивание по методу максимума правдоподобия сводится к проектированию точки $\vec{y}$ на множество возможных значений векторов $\vec{\eta}(\vec{\theta})$).

Обозначим через $\tilde{\vec{\theta}}$ оценку вектора параметров в случае, когда гипотеза $\mathcal{H}_0$ верна. Она является решением уравнения

$$\| \vec{y} - \vec{\eta}(\tilde{\vec{\theta}}) \| = \min_{\vec{\theta} \in \Omega_0} \| \vec{y} - \vec{\eta}(\vec{\theta}) \| . \qquad 2.5$$

Через $\hat{\vec{\theta}}$ обозначим решение уравнения

$$\| \vec{y} - \vec{\eta}(\hat{\vec{\theta}}) \| = \min_{\vec{\theta} \in E^q} \| \vec{y} - \vec{\eta}(\vec{\theta}) \| , \qquad 2.6$$

где считается, что функция $\vec{\eta}(\vec{\theta}) = F\vec{\theta} + \vec{f}$ определена на целом q -мерном линейном пространстве E_q (оценка без учета ограничений. Известно/II/, что $\hat{\vec{\theta}}$ является решением нормального уравнения

$$[KF]'KF\,\hat{\vec{\theta}} = [KF]'K(\vec{y}-\vec{f}) \qquad 2.7$$

(штрих обозначает транспонирование матрицы), где

$$K = \begin{pmatrix} 1/\sigma_1 & & 0 \\ & \ddots & \\ 0 & & 1/\sigma_n \end{pmatrix}.$$

Подобным образом, как в (I) воспользуемся вспомогательной оценкой $\tilde{\vec{\Phi}}_{\tilde{\theta}}$, которая равна решению уравнения

$$\|\vec{\nu}(\tilde{\vec{\Phi}}_{\tilde{\theta}}) - \vec{\eta}(\tilde{\vec{\theta}})\| = \min_{\vec{\Phi}\in\Omega_1} \|\vec{\nu}(\vec{\Phi}) - \vec{\eta}(\vec{\theta})\|. \qquad 2.8$$

Кроме матриц

$$C = F[(KF)'KF]^{-1}F'K^2$$

и

$$R = I - C\,,$$

приведенных также в (I), воспользуемся еще матрицами $C^{\tilde{\theta}}$ и $R^{\tilde{\theta}}$, зависящими от оценки $\tilde{\vec{\theta}}$. Пусть τ компонент вектора $\tilde{\vec{\theta}}$ равно граничным значениям (т.е. $\alpha_i^{(0)}$ или $\beta_i^{(0)}$). Тогда $F^{\tilde{\theta}}$ будет подматрица типа $n\times(q-\tau)$ матрицы F :

$$F_{ij}^{\tilde{\theta}} = F_{ij}\,,$$

где $i=1,\dots,n$ и j пробегает те значения, для которых $\tilde{\theta}_j$ не равны граничным значениям.

Обозначим

$$C^{\tilde{\theta}} = F^{\tilde{\theta}}(F^{\tilde{\theta}'}K^2F^{\tilde{\theta}})^{-1}F^{\tilde{\theta}'}K^2$$

и

$$R^{\tilde{\theta}} = I - C^{\tilde{\theta}}.$$

Непосредственным расчетом можно проверить, что

$$CC = C\,,\quad RR = R\,,\quad C^{\tilde{\theta}}C^{\tilde{\theta}} = C^{\tilde{\theta}},\quad R^{\tilde{\theta}}R^{\tilde{\theta}} = R^{\tilde{\theta}}$$

$$R^{\tilde{\theta}}C = CR^{\tilde{\theta}} \qquad C^{\tilde{\theta}}C = CC^{\tilde{\theta}} = C^{\tilde{\theta}} \qquad 2.9$$

и что для любых векторов $\vec{a}, \vec{b}$

$$\langle \vec{a}, A\vec{b} \rangle = \langle A\vec{a}, \vec{b} \rangle\,, \qquad 2.10$$

где A равно C или R или $C^{\tilde{\theta}}$ или $R^{\tilde{\theta}}$.

Критерий проверки гипотезы H_0 будет основан на случайной величине

$$\tau' = \frac{\langle \vec{y} - \vec{\eta}(\tilde{\vec{\theta}}),\ \nu(\tilde{\vec{\Phi}}_{\tilde{\theta}}) - \vec{\eta}(\tilde{\vec{\theta}}) \rangle}{\|R^{\tilde{\theta}}[\vec{\nu}(\tilde{\vec{\Phi}}_{\tilde{\theta}}) - \vec{\eta}(\tilde{\vec{\theta}})]\|} = \frac{\sum_{i=1}^{n}[y_i - \eta_i(\tilde{\vec{\theta}})]\chi_i/\sigma_i^2}{\sqrt{\sum_{i=1}^{n}\chi_i^2/\sigma_i^2 - \sum_{i,j=1}^{n}\sum_{k,l=1}^{q-\tau}\chi_i F_{ik} D_{kl}^{\tilde{\theta}} F_{jl}\chi_j/\sigma_i^2\sigma_j^2}}\,, \qquad 2.11$$

где $\chi_i = \nu_i(\tilde{\vec{\Phi}}_{\tilde{\theta}}) - \eta_i(\tilde{\vec{\theta}})$ и где $D^{\tilde{\theta}} = ([KF]^{\tilde{\theta}'}[KF]^{\tilde{\theta}})^{-1}$ – условная матрица ошибок параметров $\tilde{\theta}_1,\dots,\tilde{\theta}_{q-\tau}$ при заданных $\tilde{\theta}_{q-\tau+1},\dots,\tilde{\theta}_q$ (после выражения в (2.II) написано в предположении, что именно последние τ параметров достигают граничных значений).

Выражение (2.II) для вычисления τ' похоже на выражение для τ в /I/, но вместо оценки $\hat{\vec{\theta}}$ употребляется оценка $\tilde{\vec{\theta}}$, учитывающая ограничения, и вместо матрицы R употребляется случайная матрица $R^{\tilde{\theta}}$.

Чтобы проанализировать статистические свойства величины τ', выразим числитель формулы (2.II) в виде суммы двух компонент $\rho + \varkappa$, где ж)

$$\rho = \langle \vec{y} - \vec{\eta}(\hat{\theta}), \vec{\nu}(\tilde{\vec{\Phi}}_{\tilde{\theta}}) - \vec{\eta}(\tilde{\vec{\theta}}) \rangle \qquad 2.12$$

$$\varkappa = \begin{cases} \langle \vec{\eta}(\hat{\vec{\theta}}) - \vec{\eta}(\tilde{\vec{\theta}}), \vec{\nu}(\tilde{\vec{\Phi}}_{\theta}) - \vec{\eta}(\tilde{\vec{\theta}}) \rangle \text{ при } \hat{\vec{\theta}} \neq \tilde{\vec{\theta}} \\ 0 \text{ при } \hat{\vec{\theta}} = \tilde{\vec{\theta}}. \end{cases} \qquad 2.13$$

Воспользуемся еще вспомогательной случайной величиной

$$\psi = \langle cR^{\tilde{\theta}}[\vec{\eta}(\hat{\vec{\theta}}) - \vec{\eta}(\vec{\theta}_o)], \vec{\nu}(\tilde{\vec{\Phi}}_{\tilde{\theta}}) - \vec{\eta}(\tilde{\vec{\theta}}) \rangle, \qquad 2.14$$

где $\vec{\theta}_o$ – истинное значение вектора параметров в предположении, что гипотеза $\mathcal{H}_o$ правдива.

Следующая теорема описывает статистические свойства величин $\rho, \psi, \varkappa$ при условии, что задано значение оценки $\tilde{\vec{\theta}}$.

ж) Обозначим через $L_i(\vec{\theta}|\vec{y})$ функцию правдоподобия при условии, что гипотеза $\mathcal{H}_i$ верна. Подобным образом, как в /I/, можно показать, что

$$\rho + \varkappa = \ln \frac{L_1(\tilde{\vec{\Phi}}_{\tilde{\theta}}|\vec{y})}{L_o(\tilde{\vec{\theta}}|\vec{y})} - \mathcal{E}_{\tilde{\theta}} \ln \frac{L_1(\tilde{\vec{\Phi}}_{\tilde{\theta}}|\vec{y})}{L_o(\tilde{\vec{\theta}}|\vec{y})}.$$

Теорема I. Предположим, что гипотеза $\mathcal{H}_o$ является правдивой. Тогда при заданном $\tilde{\vec{\theta}}$:

I. Условное распределение величины ρ независимо от условного распределения величин $\varkappa$ и ψ.

2. Условные распределения величин ρ и ψ являются нормальными распределениями с нулевыми средними и с дисперсиями

$$\sigma_1 = \sigma_{\tilde{\theta}}(\rho) = \| R[\vec{\nu}(\tilde{\vec{\Phi}}_{\tilde{\theta}}) - \vec{\eta}(\tilde{\vec{\theta}})\|$$

и

$$\sigma_2 = \sigma_{\tilde{\theta}}(\psi) = \| R^{\tilde{\theta}} c[\vec{\nu}(\tilde{\vec{\Phi}}_{\tilde{\theta}}) - \vec{\eta}(\tilde{\vec{\theta}})]\|.$$

3. В случае, когда вектор $cR^{\tilde{\theta}}[\vec{\nu}(\tilde{\vec{\Phi}}_{\tilde{\theta}}) - \vec{\eta}(\tilde{\vec{\theta}})]$ лежит внутри области

$$\mathcal{L} = \{\vec{z} : \vec{z} = \vec{\eta}(\vec{\theta}), \vec{\theta} \in \Omega_o\}, \qquad 2.15$$

условная плотность вероятности величины $\varkappa$ равна

$$f_{\varkappa}(\chi) = \begin{cases} \frac{1}{\sqrt{2\pi}\,\sigma_2} \exp\{-\chi^2/2\sigma_2^2\}, & \text{если } \chi \le d \\ \frac{\delta(\chi)}{\sqrt{2\pi}\,\sigma_2} \int_d^{\infty} \exp\{-t^2/2\sigma_2^2\}\,dt, & \text{если } \chi > d \end{cases}, \qquad 2.16$$

где $d = \langle \vec{\eta}(\vec{\theta}) - \vec{\eta}(\vec{\theta}_o), cR^{\tilde{\theta}}[\vec{\nu}(\tilde{\vec{\Phi}}_{\tilde{\theta}}) - \vec{\eta}(\tilde{\vec{\theta}})] \rangle$
и, где $\delta(\chi)$ – функция Дирака.

Если $cR^{\tilde{\theta}}[\vec{\nu}(\tilde{\vec{\Phi}}_{\tilde{\theta}}) - \vec{\eta}(\tilde{\vec{\theta}})]$ лежит вне $\mathcal{L}$, то в (2.16) нужно заменить d на $-d$ и интегрировать от $-\infty$ до d.

Доказательство. Обозначим через

$$\vec{h} = [\vec{\nu}(\tilde{\Phi}_{\tilde{\theta}}) - \vec{\eta}(\vec{\tilde{\theta}})]$$

вектор, который будет постоянным при заданном $\tilde{\vec{\theta}}$.

Из (2.7) вытекает, что

$$[\vec{\eta}(\hat{\vec{\theta}}) - \vec{\eta}(\vec{\theta}_0)] = F[\hat{\vec{\theta}} - \vec{\theta}_0] = C[\vec{y} - \vec{\eta}(\vec{\theta}_0)].$$

Это позволяет выразить ρ и ψ в виде

$$\rho = \langle R[\vec{y} - \vec{\eta}(\vec{\theta}_0)], \vec{h} \rangle$$
$$\psi = \langle CR^{\tilde{\theta}}[\vec{y} - \vec{\eta}(\vec{\theta}_0)], \vec{h} \rangle \qquad 2.17$$

значения вектора $\vec{\eta}(\tilde{\vec{\theta}})$, которое определено в (2.5), т.е. просто проекция $\vec{y}_1$ на множество $\mathcal{L}$. С другой стороны, из (2,6) вытекает, что $\vec{\eta}(\vec{\theta})$ - результат проектирования $\vec{y}$ на линейное многообразие

$$\mathcal{F} = \{\vec{z} : \vec{z} = \vec{\eta}(\vec{\theta}),\ \vec{\theta} \in E^q\}. \qquad 2.18$$

Так как $\mathcal{L} \subset \mathcal{F}$, то $\vec{\eta}(\tilde{\vec{\theta}})$ является также результатом проектирования $\vec{\eta}(\hat{\vec{\theta}})$ на множество $\mathcal{L}$.

Поэтому при $\hat{\vec{\theta}} \neq \tilde{\vec{\theta}}$

$$\begin{aligned} x &= \langle \vec{\eta}(\hat{\vec{\theta}}) - \vec{\eta}(\tilde{\vec{\theta}}), \vec{h} \rangle = \langle CR^{\tilde{\theta}}[\vec{\eta}(\hat{\vec{\theta}}) - \vec{\eta}(\tilde{\vec{\theta}})], \vec{h} \rangle = \\ &= \langle CR^{\tilde{\theta}}[\vec{y} - \vec{\eta}(\vec{\theta}_0) + (\eta(\vec{\theta}_0) - \vec{\eta}(\tilde{\vec{\theta}}))], \vec{h} \rangle = \psi - d, \end{aligned} \qquad 2.19$$

где мы воспользовались правилами (2.9) и (2.10).

Утверждения I. и 2. теоремы являются следствиями того, что по формуле (2.17) величины ρ, ψ являются линейными функциями от ортогональных компонент нормального вектора. Чтобы получить дисперсию ρ, выразим явно скалярное произведение (2.17)

$$\rho = [\vec{y} - \vec{\eta}(\vec{\theta}_0)]' KKR\vec{h},$$

и дисперсия ρ равна /2/

$$\sigma^2_{\tilde{\theta}}(\rho) = \vec{h}' RK^2(K^{-1})^2 K^2 R\vec{h} = \| R\vec{h} \|^2.$$

Так же получим и $\sigma^2_{\tilde{\theta}}(\psi)$.

Утверждение 3. теоремы непосредственно вытекает из (2.19) и нормальной функции распределения величины ψ.

Теорема доказана.

Теорема 2. Если гипотеза H_0 является правдивой, то вероятность (полную) того, что τ'^2 превышает заданное число c^2, можно оценить следующими неравенствами:

$$P\{\tau'^2 \geq c^2\} \leq \frac{2,4}{\sqrt{2\pi}} \int_c^\infty e^{-\frac{t^2}{2}} dt \qquad 2.20$$

для всех $c > 1$, и

$$P\{\tau'^2 \geq c^2\} \leq \frac{2}{\sqrt{2\pi}} \int_c^\infty e^{-\frac{t^2}{2}} dt \qquad 2.21$$

для всех

$$c^2 > 1 + (\ln 4) \max_{\tilde{\theta} \in \Omega_0} \left[\frac{\sigma^2_{\tilde{\theta}}(\rho)}{\sigma^2_{\tilde{\theta}}(\psi)} \right].$$

Доказательство. Из теоремы I вытекает, что условное распределение суммы $\rho + \psi$ при заданном $\tilde{\tilde{\theta}}$ определено сверткой распределения для ρ с распределением для ψ :

$$f_{\rho+\psi}(x/\tilde{\tilde{\theta}}) = s\int_{-\infty}^{\infty} e^{-\frac{y^2}{2\sigma_2^2}}\, e^{-\frac{(x-y)^2}{2\sigma_1^2}}\, dy \quad ,$$

где $s = 1/2\pi\sigma_1\sigma_2$. Подобным образом из (2.16) и независимости ρ и $\varkappa$ получим условное распределение для $\rho+\varkappa$ в виде

$$f_{\rho+\varkappa}(x/\tilde{\tilde{\theta}}) = s\left\{\int_{-\infty}^{d} e^{-\frac{y^2}{2\sigma_2^2}}\, e^{-\frac{(x-y)^2}{2\sigma_1^2}}\, dy + e^{-\frac{x^2}{2\sigma_1^2}}\int_{d}^{\infty} e^{-\frac{y^2}{2\sigma_2^2}}\, dy\right\}.$$

Поэтому разность $\Delta(x/\tilde{\tilde{\theta}}) \equiv f_{\rho+\psi}(x/\tilde{\tilde{\theta}}) - f_{\rho+\varkappa}(x/\tilde{\tilde{\theta}})$ равна

$$\Delta(x/\tilde{\tilde{\theta}}) = s\, e^{-\frac{x^2}{2\sigma_1^2}}\left\{\int_{d}^{\infty} e^{-\frac{y^2}{2}\left[\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}\right]+\frac{xy}{\sigma_1^2}}\, dy - \int_{d}^{\infty} e^{-\frac{y^2}{2\sigma_2^2}}\, dy\right\}.$$

Воспользуемся сокращенным обозначением

$$a = \sqrt{\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}} \qquad 2.22$$

и напишем $\Delta(x/\tilde{\tilde{\theta}})$ в виде

$$\Delta(x/\tilde{\tilde{\theta}}) = s\, e^{-\frac{x^2}{2\sigma_1^2}}\left\{\int_{d}^{\infty} e^{-\frac{a^2}{2}\left[y-\frac{x}{a^2\sigma_1^2}\right]+\frac{x^2}{2\sigma_1^4 a^2}}\, dy - \int_{d}^{\infty} e^{-\frac{y^2}{2\sigma_2^2}}\, dy\right\} =$$

$$= s\, e^{-\frac{x^2}{2\sigma_1^2}}\left\{\frac{1}{a}\, e^{\frac{x^2}{2a^2\sigma_1^4}}\, \Phi(ad - x/a\sigma_1^2) - \sigma_2\, \Phi(d/\sigma_2)\right\},$$

где $\Phi(u) = (1/\sqrt{2\pi})\int_{u}^{\infty} \exp\{-t^2/2\}\, dt$.

Нам нужна сумма

$$g(x/\tilde{\tilde{\theta}}) \equiv \Delta(x/\tilde{\tilde{\theta}}) + \Delta(-x/\tilde{\tilde{\theta}}) = s\, e^{-\frac{x^2}{2\sigma_1^2}}\left\{\frac{1}{a}\, e^{\frac{x^2}{2a^2\sigma_1^4}}\left[\Phi(ad - x/a\sigma_1^2) + \Phi(ad + x/a\sigma_1^2)\right] - 2\sigma_2\Phi(d/\sigma_2)\right\}. \qquad 2.23$$

Из того, что $ad \le 0$, $d/\sigma_2 \le 0$, вытекают неравенства

$$2 \geqslant \Phi(ad + x/a\sigma_1^2) + \Phi(ad - x/a\sigma_1^2) \geqslant 1$$

$$1 \geqslant \Phi(d/\sigma_2) \geqslant 1/2. \qquad 2.24$$

Оценим $g(x/\tilde{\tilde{\theta}})$ снизу для значений x , выполняющих неравенство

$$\exp\{x^2/2a^2\sigma_1^4\} \geqslant a\sigma_2. \qquad 2.25$$

Тогда из (2.23), (2.24) и (2.25) получим

$$g(x/\tilde{\tilde{\theta}}) \geqslant -\frac{1}{\sqrt{2\pi}}\,\frac{1}{\sqrt{2\pi}\,\sigma_1}\, e^{-\frac{x^2}{2\sigma_1^2}}. \qquad 2.26$$

Выясним смысл неравенства (2.25). Оно сводится к неравенству

$$\frac{x^2}{\sigma_1^2+\sigma_2^2} \geqslant \frac{\sigma_1^2}{\sigma_2^2}\, \ln\left[1+\frac{\sigma_2^2}{\sigma_1^2}\right]. \qquad 2.27$$

Функция $[\ln(1+z)]/z$ - монотонно убывающая функция от z для $z \geqslant 0$. Это можно показать, если вычислить производную

$$\frac{d}{dz}\left(\frac{\ln(1+z)}{z}\right) = -\frac{1}{z^2(1+z)}\left[(1+z)\ln(1+z) - z\right]. \qquad 2.28$$

Выражение в квадратных скобках в (2.28) равно нулю при $z=0$ и является возрастающей функцией при $z \geqslant 0$. Действительно,

$$\frac{d}{dz}\left[(1+z)\ln(1+z) - z\right] = \ln(1+z) \geqslant 0.$$

Поэтому $\frac{d}{dz}\left(\frac{\ln(1+z)}{z}\right) \leq 0$ для всех z и

$$\sup_{z\geqslant 0} \frac{\ln(1+z)}{z} = \lim_{z\to 0} \frac{\ln(1+z)}{z} = 1. \qquad 2.29$$

Поэтому для всех X, выполняющих неравенство

$$\frac{X^2}{\sigma_1^2+\sigma_2^2} \geqslant 1, \qquad 2.30$$

выполнено (2.27) и (2.26).

Из того, что $\rho+\psi$ условно распределено нормально с нулевым средним и дисперсией $\sigma_1^2+\sigma_2^2$, получим при помощи (2.26)

$$f_{\rho+\varkappa}(x|\tilde{\bar{\theta}}) + f_{\rho+\varkappa}(-x|\tilde{\bar{\theta}}) \leq \frac{2}{\sqrt{2\pi(\sigma_1^2+\sigma_2^2)}}\, e^{-\frac{x^2}{2(\sigma_1^2+\sigma_2^2)}} + \frac{0{,}4}{\sqrt{2\pi}\cdot\sigma_1}\, e^{-\frac{x^2}{2\sigma_1^2}}$$

и вероятность

$$P\left\{\frac{X^2}{\sigma_1^2+\sigma_2^2} \geqslant c^2 \,\middle|\, \tilde{\bar{\theta}}\right\} \leq \frac{2{,}4}{\sqrt{2\pi}}\int_C^\infty e^{-\frac{t^2}{2}}\,dt \qquad 2.31$$

для всех $C \geqslant 1$. Для полного доказательства неравенства (2.20) нужно еще переписать сумму $\sigma_1^2 + \sigma_2^2$

$$\sigma_1^2 + \sigma_2^2 = \sigma_{\tilde\theta}^2(\rho) + \sigma_{\tilde\theta}^2(\psi) =$$
$$= \|R\vec{h}\|^2 + \|CR^{\tilde\theta}\vec{h}\|^2 =$$
$$= \|(R+CR^{\tilde\theta})\vec{h}\|^2 = \|(I-C^{\tilde\theta})\vec{h}\|^2 = \|R^{\tilde\theta}\vec{h}\|^2,$$

где мы воспользовались правилами (2.9). Из (2.31) тогда получим

$$P\{\tau'^2 \geqslant c^2\} = P\left\{\frac{X^2}{\sigma_1^2+\sigma_2^2} \geqslant c^2\right\} = \int_{\Omega_0} P\left\{\frac{X^2}{\sigma_1^2+\sigma_2^2} \geqslant c^2 \,\middle|\, \tilde{\bar\theta}\right\} df(\tilde{\bar\theta}) \geqslant \frac{2{,}4}{\sqrt{2\pi}}\int_C^\infty e^{-\frac{t^2}{2}}\,dt.$$

Подобным образом доказывается оценка (2.21). При помощи (2.29) получим, что из неравенства

$$\frac{X^2}{\sigma_1^2+\sigma_2^2} \geqslant 1 + (\ln 4)\max_{\tilde{\bar\theta}\in\Omega_0}\left[\frac{\sigma_1^2}{\sigma_2^2}\right]$$

вытекает, что

$$X^2 \geqslant \frac{\sigma_1^2+\sigma_2^2}{\sigma_1^2\,\sigma_2^2}\,\sigma_1^4 \ln\left[4\sigma_2^2\,\frac{\sigma_1^2+\sigma_2^2}{\sigma_1^2\,\sigma_2^2}\right] =$$
$$= 2a^2\sigma_1^4 \ln[2\sigma_2 a]$$

и что

$$\frac{\exp\{X^2/2a^2\sigma_1^4\}}{a} \geqslant 2\sigma_2.$$

Используя это неравенство вместе с (2.24), получим из (2.23), что

$$g(x|\tilde{\bar\theta}) \geqslant 0$$

т.е.

$$f_{\rho+\varkappa}(x|\tilde{\bar\theta}) + f_{\rho+\varkappa}(-x|\tilde{\bar\theta}) \leq \frac{2}{\sqrt{2\pi(\sigma_1^2+\sigma_2^2)}}\, e^{-\frac{x^2}{2(\sigma_1^2+\sigma_2^2)}}.$$

Поэтому

$$P\{\tau'^2 > c^2 | \tilde{\bar\theta}\} \leq \frac{2}{\sqrt{2\pi}}\int_C^\infty e^{-\frac{t^2}{2}}\,dt$$

для всех $\tilde{\bar{\theta}}$ и для всех $c^2 \geq 1 + (\ln 4)\max[\sigma_1^2/\sigma_2^2]$. Усредняя последнее неравенство по распределению величины $\tilde{\bar{\theta}}$, получим (2.21).

Теорема доказана.

Замечание: В случае ограничений на параметры применение τ -критерия аналогично случаю, когда нет ограничений на параметры, только вместо величины τ вычисляется величина τ' по формуле (2.II). Если $\tau'_{\text{экс}}$ результат таких вычислений, (и $|\tau'_{\text{экс}}| \geq 1$), то, следуя формуле (2.20), вычислим интеграл

$$\frac{2,4}{\sqrt{2\pi}} \int\limits_{|\tau'_{\text{экс}}|}^{\infty} e^{-\frac{t^2}{2}} dt \qquad 2.32$$

при помощи таблиц нормального распределения. Тем и получим оценку сверху вероятности того, что гипотеза $\mathcal{H}_0$ верна (и если эта вероятность мала, гипотезу следует отвергнуть).

Меняя местами обе гипотезы в (2.II), вычислим новое значение для $\tau'_{\text{экс}}$ и интеграл (2.32) оценивает вероятность ошибки при отбрасывании гипотезы $\mathcal{H}_1$ (т.е. что $\vec{\nu}(\vec{\Phi})$ правильна).

Если функции $\vec{\eta}(\vec{\theta})$, $\vec{\nu}(\vec{\Phi})$ нелинейны, то придется их линеаризовать, т.е. вместо матриц F, G воспользоваться матрицами производных $\left\| \frac{\partial \eta_i(\tilde{\bar{\theta}})}{\partial \theta_j} \right\|$ и $\left\| \frac{\partial \nu_i(\tilde{\vec{\Phi}})}{\partial \Phi_i} \right\|$.

§ 3. τ^2 -критерий

Здесь мы будем придерживаться по возможности тех же обозначений, что и в предыдущем параграфе. Но основная часть данного параграфа будет посвящена случаю линейных гипотез с неограниченными параметрами.

Для уменьшения обозначений мы будем называть гипотезу $\mathcal{H}_0$ гипотезой $\vec{\eta}(\vec{\theta})$, а гипотезу $\mathcal{H}_1$ - гипотезой $\vec{\nu}(\vec{\Phi})$. Выражение: вектор ортогонален гипотезе $\vec{\eta}(\vec{\theta})$, мы будем понимать в смысле ортогональности геометрическому семейству $\mathcal{F}$ точек $\vec{\eta}(\vec{\theta})$ при всевозможных значениях $\vec{\theta}$ (аналогично для $\vec{\nu}(\vec{\Phi})$). Ортогональность гипотез мы будем понимать в смысле ортогональности соответствующих им геометрических семейств точек.

Большим достоинством τ -критерия является простота вычислений, связанная с нормальностью распределения величины

$$\tau = \frac{\langle \vec{y} - \vec{\eta}(\hat{\vec{\theta}}), \vec{\nu}(\hat{\vec{\Phi}}_\theta - \vec{\eta}(\hat{\vec{\theta}}) \rangle}{\| R[\vec{\nu}(\hat{\vec{\Phi}}_\theta) - \vec{\eta}(\hat{\vec{\theta}})] \|} = \frac{\langle \vec{y} - \vec{\eta}(\hat{\vec{\theta}}), R(\vec{\nu}(\hat{\vec{\Phi}}_\theta) - \vec{\eta}(\hat{\vec{\theta}})) \rangle}{\| R[\vec{\nu}(\hat{\vec{\Phi}}_\theta) - \vec{\eta}(\hat{\vec{\theta}})] \|}, \qquad 3.I$$

где матрица R (см. § 2) есть оператор проектирования, переводящий вектор, на который он действует, в перпендикуляр к гипотезе (гиперплоскости) $\vec{\eta}(\vec{\theta})$, что фактически сводится к решению (2.6). Второе равенство определяется тем, что $\vec{y} - \vec{\eta}(\hat{\vec{\theta}}) = R(\vec{y} - \vec{\eta}(\hat{\vec{\theta}}))$, так как $\vec{y} - \vec{\eta}(\hat{\vec{\theta}})$ уже ортогонально $\vec{\eta}(\vec{\theta})$, и $RR = R$ (2.9). Следовательно, τ есть скалярное произведение вектора "невязки" $\vec{y} - \vec{\eta}(\hat{\vec{\theta}})$ и единичного вектора $\frac{\vec{h}}{\|\vec{h}\|}$, ортогонального гипотезе $\vec{\eta}(\hat{\vec{\theta}})$, где

$$\vec{h} = R(\vec{\nu}(\hat{\vec{\Phi}}_\theta) - \vec{\eta}(\hat{\vec{\theta}})). \qquad 3.2$$

Нормальность распределения τ связана с тем, что вектор $\vec{h}$ зависит лишь от составляющих $\vec{y}$, параллельных гиперплоскости $\vec{\eta}(\vec{\theta})$ и ортогональных к $\vec{h}$.

Обычно вектор $\vec{h}$ примерно направлен в сторону наибольшего различия гипотез $\vec{\nu}$ и $\vec{\eta}$. τ-критерий показывает, насколько случайно отклонение эксперимента от гипотезы в этом направлении (в отличие от χ^2-критерия, который суммирует все составляющие ошибок, затемняя наиболее важную для сравнения гипотез информацию).

Однако, например, в случае ортогональных гипотез $\vec{\nu}(\vec{\Phi})$ и $\vec{\eta}(\vec{\theta})$ вектор $\vec{h}$ ортогонален обеим гипотезам. Действительно, он ортогонален гипотезе $\vec{\eta}(\vec{\theta})$ по построению и ортогонален $\vec{\nu}(\vec{\Phi})$, так как является линейной комбинацией векторов $\vec{\nu}(\hat{\Phi}_\theta)-\vec{\eta}(\hat{\theta})$ и $(I-R)(\vec{\nu}(\hat{\vec{\Phi}}_\theta)-\vec{\eta}(\hat{\vec{\theta}}))$, ортогональных $\vec{\nu}(\vec{\Phi})$. Естественно, что такой вектор никакого отношения к различию гипотез не имеет. Если расстояние между непересекающимися гипотезами мало, а $\vec{y}$ лежит вблизи одной из гиперплоскостей и вдали от области скрещивания, τ будет мало для обеих гипотез даже когда χ^2 резко различны.

По этой причине желательно иметь возможность контроля мощности τ-критерия, а также критерий, работающий там, где τ-критерий оказывает (δ-критерий слишком трудоемок и трудно сопоставим с τ-критерием).

Вполне естественной является попытка заменить в формуле (3.2) вектор $\vec{\nu}(\hat{\vec{\Phi}}_\theta)-\vec{\eta}(\hat{\vec{\theta}})$ вектором $\vec{\nu}(\hat{\vec{\Phi}})-\vec{\eta}(\hat{\vec{\theta}})$, соединяющим оценки экспериментальных величин по гипотезам H_0 и H_1. Этот вектор всегда, по построению, направлен в сторону наибольшего различия гипотез.

Обозначим
$$\vec{z}=R(\vec{\nu}(\hat{\vec{\Phi}})-\vec{\eta}(\hat{\vec{\theta}}))$$

и
$$r=\frac{\langle \vec{y}-\vec{\eta}(\hat{\vec{\theta}}),\vec{z}\rangle}{\|\vec{z}\|}. \quad 3.3$$

К сожалению, распределение величины r не является таким простым, как у τ, и требует специального исследования.

Вектор $\vec{z}=R[\vec{\nu}(\hat{\vec{\Phi}})-\vec{\eta}(\hat{\vec{\theta}})]$ является перпендикуляром, опущенным из точки, лежащей на гипотезе $\vec{\nu}(\vec{\Phi})$ на гипотезу $\vec{\eta}(\vec{\theta})$. Таким образом, он зависит (притом линейно) лишь от m степеней свободы гипотезы $\vec{\nu}(\vec{\Phi})$. То-есть можно написать

$$\vec{z}=\alpha_0\vec{e}_0+\sum_{i=1}^{m}\alpha_i(\vec{y})\,\vec{e}_i\,, \quad 3.4$$

где $\vec{e}_i$ – единичные векторы, нормированные в согласии с (2.4). Выбором системы координат мы можем добиться, чтобы $\vec{e}_i$ были ортогональны друг другу и остальным компонентам невязки $\vec{y}-\eta(\vec{\theta}_0)=\sum_{i=0}^{n-1}x_i\vec{e}_i$, где x_i – нормально распределенные величины с единичной дисперсией и средним, равным нулю. В r дадут вклад лишь компоненты $\vec{y}-\vec{\eta}(\vec{\theta}_0)$, не ортогональные к $\vec{z}$ и, следовательно, $r=\alpha_0x_0+\sum_{i=1}^{m}\alpha_i(y)x_i$. Вершина вектора $\frac{\vec{z}}{\|\vec{z}\|}$ перемещается в подпространстве размерности $m+1$. Поэтому вклад в $\vec{z}$ могут дать не более, чем $m+1$ независимая компонента случайного шума $\vec{y}-\vec{\eta}(\vec{\theta}_0)$, т.е.

$$r=\frac{\langle \sum_{k=0}^{m}x_k\vec{e}_k,\vec{z}\rangle}{\|\vec{z}\|}. \quad 3.5$$

Модуль проекции вектора не может превышать модуль самого вектора. Поэтому

$$r^2\le\sum_{k=0}^{m}x_k^2. \quad 3.6$$

Таким образом, τ^2 распределено не хуже, чем χ^2_{m+1}. При этом наихудшее распределение получается при сильной корреляции $\vec{z}$ и $x_0,\dots,x_m$. Тем не менее, полная корреляция невозможна, так как у вектора $\vec{z}$ лишь m степеней свободы. При α_0 в (3.4), равном нулю, распределение τ^2 не хуже, чем χ^2_m. χ^2_m -распределение достижимо. Если $\vec{z}$ не коррелировано с $x_0,\dots,x_m$, то τ^2 распределено как χ^2_1.

(3.6) уже позволяет как-то использовать τ^2-критерий. Вычисление τ даже проще, чем τ, и может проводиться одновременно с вычислением τ.

Легко видеть, что направление векторов $\vec{z}$ в (3.3) и $\vec{h}$ в (3.2) обычно весьма сходно за исключением почти ортогональных гипотез. Поэтому, если величина τ^2 заметно больше, чем τ^2, то это может служить указанием на трудности для τ-критерия. В этом случае можно использовать (3,6), но, конечно, интересно оценить распределение точнее.

Построим систему векторов $\vec{e}_0,\vec{e}_1,\dots,\vec{e}_m,\dots,\vec{e}_{m+q}$, ортонормированных так, чтобы $\vec{e}_0$ был направлен вдоль линии кратчайшим образом соединяющей гипотезы, $\vec{e}_{m+1},\dots,\vec{e}_{m+q}$ лежали в гиперплоскости $\vec{\eta}(\vec{\theta})$, а $\vec{e}_1,\dots,\vec{e}_m$ были образованы ортогонализацией к $\vec{e}_0,\vec{e}_{m+1},\dots,\vec{e}_{m+q}$ координатной системы гиперплоскости $\vec{\nu}(\vec{\Phi})$.

В качестве начала координат выберем точку гипотезы $\vec{\eta}(\vec{\theta})$, ближайшую к $\vec{\nu}(\vec{\Phi})$. Тогда

$$\vec{z} = \alpha_0\vec{e}_0 + \sum_{k=1}^{m+q} x_k R_\eta C_\nu \vec{e}_k \quad ж) \qquad 3.7$$

ж) Здесь $R_\eta = R$. $C_\nu = I - R_\nu$ – оператор проектирования вектора на гипотезу $\vec{\nu}(\vec{\Phi})$.

Действительно, при построении $\vec{z}$ компонента $x_k\vec{e}_k$ проектируется на гипотезу $\vec{\nu}(\vec{\Phi})$, т.е. строится $C_\nu x_k\vec{e}_k$. Компонента $C_\nu x_k\vec{e}_k$, ортогональная $\vec{\eta}(\vec{\theta})$, т.е. $x_k R_\eta C_\nu \vec{e}_k$, дает приращение вектора $\vec{z}$. По построению

$$\vec{z} = \vec{z}_0 + R_\eta C_\nu (\vec{y} - \vec{\eta}(\vec{\theta}_0)).$$

Представляя $\vec{y} - \vec{\eta}(\vec{\theta}_0) = \sum_{k=0}^{n-1} x_k\vec{e}_k$, мы увидим, что лишь $C_\nu R_\eta \vec{e}_k$ для $k=0,\dots,m+q$ могут оказаться ненулевыми (другие $\vec{e}_k$ ортогональны к гипотезам). Отметим также, что $R_\eta C_\nu \vec{e}_k$ остается линейной комбинацией лишь $\vec{e}_0,\vec{e}_1,\dots,\vec{e}_m$, так как $R_\eta C_\nu \vec{e}_k$ ортогонально к $\vec{\eta}(\vec{\theta})$ и, следовательно, к $\vec{e}_k$, параллельным $\vec{\eta}(\vec{\theta})$.

Если гипотезы имеют общие точки, т.е. пересекаются, $\vec{e}_0$ не требуется. Для единообразия его можно вводить с нулевым α_0.

(3.7) можно еще упростить. Вращением векторов $\vec{e}_{m+1},\dots,\vec{e}_{m+q}$ добьемся, чтобы $C_\nu\vec{e}_{m+1},\dots,C_\nu\vec{e}_{m+q}$ стали ортогональны (некоторые из них могут стать нулевыми). Для этого нам следует унитарными преобразованиями d добиться, чтобы

$$\sum_{i,\ell=m+1}^{m+q} \langle C_\nu d_{ji}\vec{e}_i, C_\nu d_{k\ell}\vec{e}_\ell\rangle = \sum_{i,\ell=m+1}^{m+q} d_{ji} d_{k\ell} c_{i\ell} = \delta_{jk}\lambda_k,$$

где

$$c_{i\ell} = \langle C_\nu\vec{e}_i, C_\nu\vec{e}_\ell\rangle,$$

то-есть нам нужно найти собственные векторы симметричной матрицы $c_{i\ell}$, что всегда возможно. Одновременно для новых $\vec{e}_{m+1},\dots\vec{e}_{m+q}$ будет $C_\nu\vec{e}_k\perp\vec{e}_j$ при $j\ne k$ ввиду $C_\nu C_\nu = C_\nu$ согласно (2.9). $\vec{e}_k$ для $k=1,\dots,m$ выберем, ортогонализуя ненулевые $C_\nu\vec{e}_{k+m}$ к $\vec{e}_{k+m}$, так что

$$C_\nu\vec{e}_{k+m} = \gamma_k\vec{e}_k + \gamma_{k+m}\vec{e}_{k+m}. \qquad 3.8$$

Они окажутся также ортогональны между собой в силу $C_\nu\vec{e}_{k+m}\perp C_\nu\vec{e}_{j+m}$, $\vec{e}_{k+m}\perp\vec{e}_{j+m}$ и $C_\nu\vec{e}_{k+m}\perp\vec{e}_{j+m}$ при $k\ne j$.

Чтобы не возникало проблем с нумерацией векторов при $q > m$, следует расставить векторы $\vec{e}_{m+1}, \dots, \vec{e}_{m+q}$ так, чтобы векторы, для которых $C_\nu \vec{e}_k = 0$, имели максимальные номера.

Если окажется, что число ненулевых $C_\nu \vec{e}_k$ меньше m (как, например, в случае $m > q$), то $C_\nu \vec{e}_k$ следует дополнить ортогональными к ним и между собой единичными векторами до полной системы векторов в гиперплоскости $\vec{\nu}(\vec{\Phi})$. Они же дополнят ортонормированную систему векторов $\vec{e}_k$ до $m+q+1$ (с γ_k, равными нулю).

Векторы $\vec{e}_{m+1}, \dots, \vec{e}_{m+q}$, будучи спроектированы на $\vec{\nu}(\vec{\Phi})$ и обратно на $\vec{\eta}(\vec{\theta})$, не меняют направления. В нашей системе координат построение вектора $\vec{z}$ в (3.7) можно вести покомпонентно, пользуясь двумерной геометрией. К примеру, из (3.8) следует $R_\eta C_\nu \vec{e}_{k+m} = \gamma_k \vec{e}_k$. Вспомним также, что мы выбрали нумерацию $\vec{e}_k$ так, что при

$q > m$ и $m < k \le q$ – $C_\nu \vec{e}_k = 0$.

Таким образом,

$$\vec{z} = \alpha_0 \vec{e}_0 + \sum_{k=1}^{m} (\alpha_k X_k + \alpha_{k+m} X_{k+m}) \vec{e}_{k+m}. \qquad 3.9$$

В действительности между коэффициентами α_k и α_{k+m} существует связь. Пусть угол между векторами $\vec{e}_k$ и $C_\nu \vec{e}_k$ есть θ_k. Пусть также смещение $\vec{y} - \vec{\eta}(\theta)$ вдоль $\vec{e}_k$ есть X_k, а вдоль $\vec{e}_{k+m}$ есть X_{k+m}. К каким изменениям в $\vec{z}$ приведут эти смещения? Смещение X_{k+m} приведет к смещению оценки $\vec{\nu}(\vec{\Phi})$ на $X_k \cos\theta_k$, что, в свою очередь, приведет к смещению $\vec{z}$ на величину $X_k \cos\theta_k \sin\theta_k \vec{e}_{k+m}$. Смещение на X_k в направлении $\vec{e}_k$ приведет к смещению $\vec{z}$ на величину $X_k \sin^2\theta_k \vec{e}_k$. Таким образом,

$$\alpha_k = \sin^2\theta_k \; ; \quad \alpha_{k+m} = \cos\theta_k \cdot \sin\theta_k \, , \qquad 3.10$$

$$\vec{y} - \vec{\eta}(\theta_0) = \sum_{k=0}^{n-1} X_k \vec{e}_k \, , \qquad 3.11$$

$$\tau = \frac{\alpha_0 X_0 + \sum_{k=1}^{m} X_k (\alpha_k X_k + \alpha_{k+m}(X_{k+m} + \beta_k))}{\sqrt{\alpha_0^2 + \sum_{k=1}^{m} (\alpha_k X_k + \alpha_{k+m}(X_{k+m} + \beta_k))^2}} \, , \qquad 3.12$$

где X_k – нормально распределенные независимые переменные с единичной дисперсией и средним, равным нулю. β_k связаны с положением точного решения $\vec{\eta}(\vec{\theta}_0)$ и неизвестны.

Выражение (3.12) уже сравнительно простое и может быть исследовано. Все коэффициенты (3.12), за исключением β_k, могут быть вычислены (β_k могут быть оценены).

Для сравнения можем выразить в тех же терминах вектор $\vec{h}$, используемый в τ.

$$\vec{h} = \alpha_0 \vec{e}_0 + \sum_{k=1}^{m} \cos\theta_k \cdot \sin\theta_k (\alpha_{k+m} + \beta_k) \vec{e}_k \, . \qquad 3.13$$

Один из возможных способов использования (3.12) – это оценка его методом Монте-Карло, что не слишком трудоемко ввиду простоты выражения.

Из (3.13) и (3.12) видно, что $\vec{h}$ и $\vec{z}$ весьма мало отличаются, если велик хотя бы один из параметров α_0 или β_k, и, соответственно, распределение величины τ^2 стремится к распределению τ, т.е. к χ_1^2. Таким образом, большие значения α_0 и β_k приводят к уменьшению дисперсии τ^2. Однако, как мы видели ранее, наличие, например, α_0, в принципе допускает ухудшение распределения τ^2 (χ^2_{m+1} вместо χ^2_m). Моделирование распределения (3.13) методом Монте-Карло показало, что действительно, при не слишком больших значениях α_0 распределение несколько ухудшается по сравнению с $\alpha_0 = 0$. С другой стороны, моделирование показало, что τ^2-распределение при нулевых α_0 и β_k хорошо оценивается χ^2-распределением (вообще говоря, с дробным числом степеней свободы), а появление ненулевых α_0 и β_k не увеличивает число степеней свободы оценивающего χ^2-распределения больше чем на единицу.

Используя оценки дисперсии τ^2 в асимптотике, подправленные на неасимптотические варианты, переходящие в χ^2-распределение, удалось получить следующие формулы:

$$
\begin{aligned}
S &= 2\left(\frac{U}{L}\right)^2 D_U + \frac{1}{2}\left(\frac{U}{L}\right)^4 D_L - 2\left(\frac{U}{L}\right)^3 D_{UL} \\
U &= \sqrt{\left(\sum_1^m \alpha_k\right)^2 + \frac{1}{2}\sum_1^m \alpha_{k+m}^2} \\
L &= \sum_1^m \left(\alpha_k^2 + \alpha_{k+m}^2\right)^2 \\
D_U &= \sum_1^m \left(2\alpha_k^2 + \alpha_{k+m}^2\right) \\
D_L &= 2\sum_1^m \left(\alpha_k^2 + \alpha_{k+m}^2\right)^2 \\
D_{UL} &= 2\sum_1^m \alpha_k \left(\alpha_k^2 + \alpha_{k+m}^2\right).
\end{aligned}
\qquad 3.14
$$

Здесь S – число степеней свободы оценивающего χ^2распределения. В принципе оценка (3.14) может быть сделана и без преобразования системы координат, так как является функцией некоторых моментов числителя и знаменателя (3.12), а следовательно, и (3.5).

В случае задачи с ограничениями на параметры выражение для τ не меняется, ($\vec{\eta}(\hat{\theta})$ и $\vec{\nu}(\hat{\Phi})$ естественно заменяются на $\vec{\eta}(\hat{\vec{\theta}})$ и $\vec{\nu}(\hat{\vec{\Phi}})$). Оценка (3.6) остается справедливой. Детального анализа τ^2-распределения, однако, в этом случае не проводилось.

Замечание. Как видно, точная оценка τ^2-распределения довольно затруднительна. Поэтому прибегать к ней имеет смысл лишь в

крайних случаях при $\tau^2 \gg \tau^2$. Грубая же оценка может привести к более слабым результатам, чем τ-критерий.

ЛИТЕРАТУРА

1. A. PAZMAN. A method for testing complex linear hypothesis and its use in the phase shift analysis. Препринт ОИЯИ Е-5-3775, 1968.

2. С.И.Биленькая, Л.Н.Глонти, Ю.М.Казаринов, В.С.Киселев. Измерение параметра Вольфенштейна А в p-n рассеянии при энергии 605 Мэв. ЖЭТФ 59, 1049, 1970.

3. S.I. BILENKAYA, G.GOZZIKA and P.LEHAR and Z.JANOUT. Phase-shift analysis of p-p and n-p elastic scattering at 735 Mev. Nuclear Physics. B13. 375-384 1969.

4. S.I. BILENKAYA, J. BYSTRICKY, Z. JANOUT, YU.M. KAZARINOV, L.LEHAR, A.PAZMAN. N-N scattering phase-shift analysis at energies near to 20 Mev. CZECHOSLOVAK Journal of Physics. Volume B19, 891, 1969.

5. Л.Н.Глонти, Ю.М.Казаринов, В.С.Киселев, И.Н.Силин. Матрица упругого нуклон-нуклонного рассеяния при энергии 630 Мэв. I. Фазовый анализ p-p рассеяния. Сообщение ОИЯИ P1-6339, 1972.

6. Л.Н.Глонти, Ю.М.Казаринов, В.С.Киселев, И.Н.Силин. Матрица упругого нуклон-нуклонного рассеяния при энергии 630 Мэв. II Фазовый анализ N-N рассеяния. Сообщение ОИЯИ P1-6387, 1972.

7. Ю.М.Казаринов, В.С.Киселев, А.М.Розанова, И.Н.Силин.

Дискриминация статистических гипотез при близких значениях χ^2. Припринт ОИЯИ PI-3268, Дубна, 1967.

8. И.Н.Силин.

Сравнение гипотез при близких значениях χ^2. Материалы Международной школы по физике высоких энергий. Попрадске Плесо, Чехословакия, октябрь, 1967.

9. Ю.М.Казаринов, В.С.Киселев, А.М.Розанова, И.Н.Силин.

Возможный метод дискриминации статистических гипотез при близких значениях χ^2 и использование его в фазовом анализе нуклон- нуклонного рассеяния. ЯФ 7, вып. 2, 340, 1968.

10. Л.Н.Глонти, Ю.М.Казаринов, А.М.Розанова, И.Н.Силин.

Новые результаты фазового анализа нуклон-нуклонного рассеяния при энергии 630 Мэв. ЯФ 7, 1060, 1968.

11. Ю.В.Линник.

Метод наименьших квадратов и основы математико-статистической теории обработки наблюдений. Физматгиз, 1958.

Рукопись поступила в издательский отдел
22 мая 1973 года.