Статистические критерии для поиска аномалий типа ступеньки в интегральных спектрах

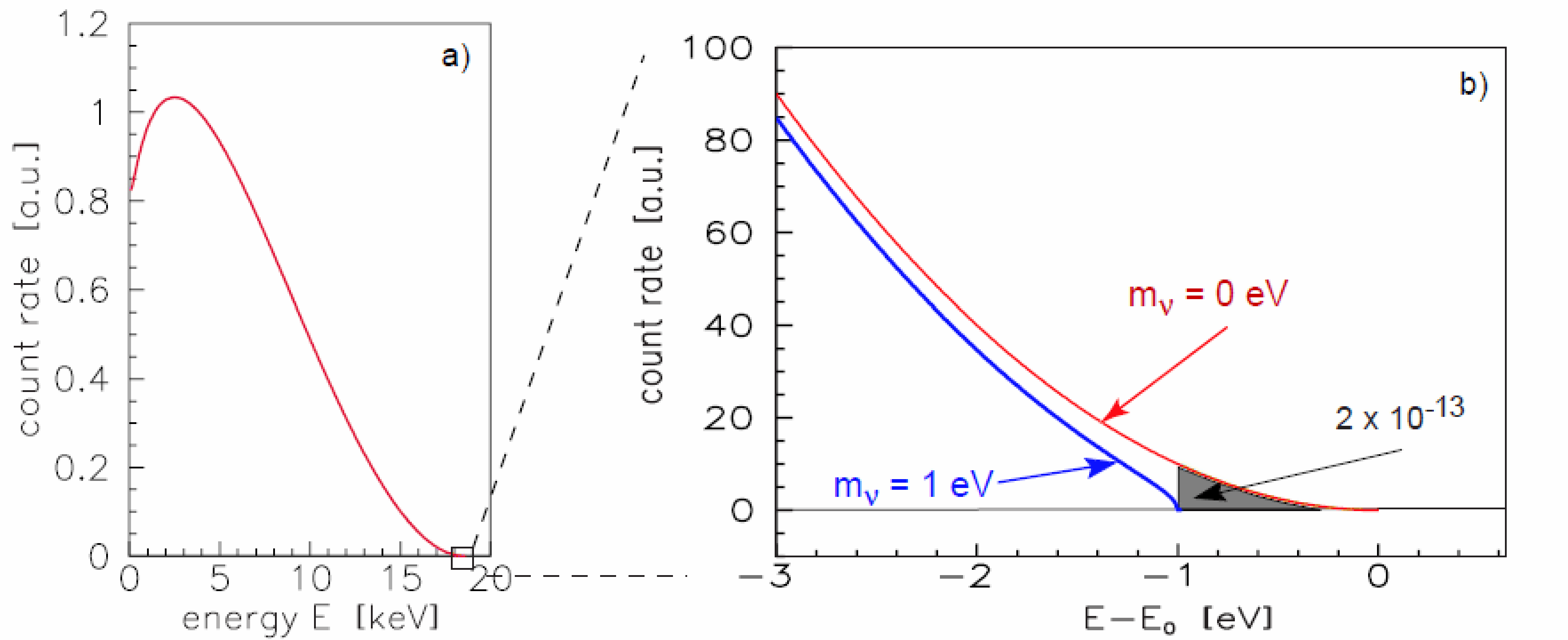
А.В. Лохов, Ф.В. Ткачёв, П.С. Труханов

Троицк – 24 января 2012

План доклада

1. Спектр электронов в β-распаде. Проблема «ступеньки» в спектре эксперимента Троицк-ν-масс
2. Статистические критерии
   - стандартные критерии
   - локально наиболее мощный критерий (метод квазиоптимальных моментов)
   - «квазиоптимальный» критерий
   - «критерий парных корреляций соседей»
   - сравнение критериев
3. Результаты использования данных 11 сеансов
4. Исследование значений критериев для всех сеансов
   - на равномерность
   - на симметричность
5. Вывод

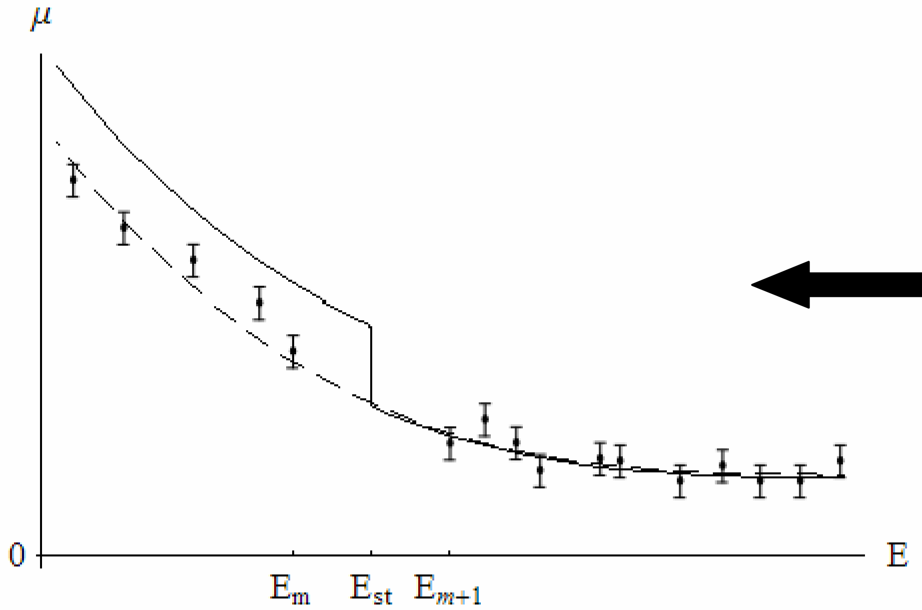
Спектр электронов в β-распаде

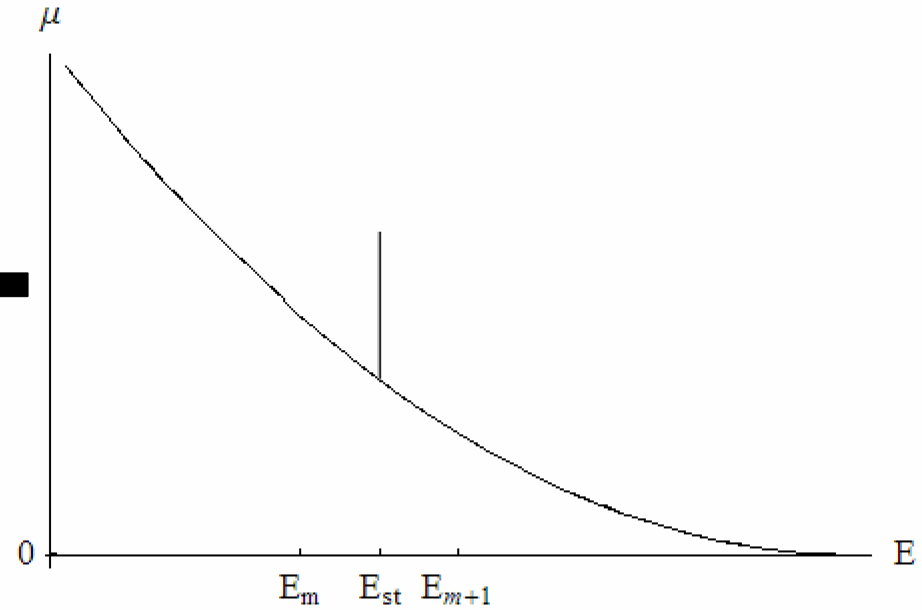


$$\boxed{\frac{dN}{dE} = C \times F(Z,E)\, p\left(E + m_e c^2\right)\left(E_0 - E\right)\left[\left(E_0 - E\right)^2 - m_\nu^2\right]^{1/2} \Theta\left(E_0 - E - m_\nu\right)}$$

При $(E_0 - E)^2 \gg m_\nu$ $\quad \frac{dN}{dE} \sim (E_0 - E)^2$ $\quad C = \frac{G_F^2}{2\pi^3}\cos^2\theta_C \left|M\right|^2$

Спектр электронов в эксперименте Троицк-ν-масс





Характерный вид интегрального спектра вблизи граничной энергии в отсутствии аномальных вкладов (штрихованный) и при наличии ступеньки (сплошной)

Измеряется спектр $\mu(E)$ для набора значений E_i $(i=1,..,M)$. Для каждого E_i измеряется число событий N_i за некоторое фиксированное время.

N_i распределены по Пуассону со средними $\mu_i = \mu(E_i)$.

Фитируются параметры:

N_0 – нормировка спектра

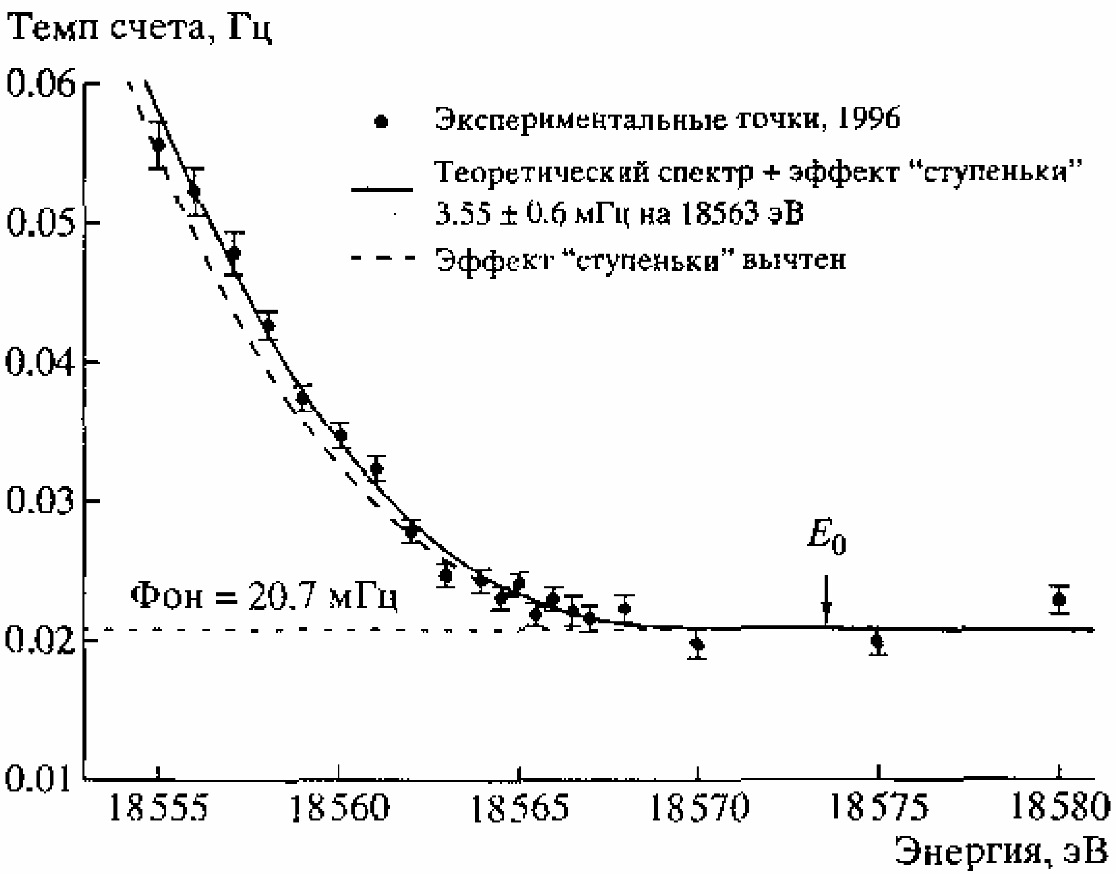
B – фон (константа)

E_0 – граничная энергия

$m_\nu{}^2$ – квадрат массы электронного антинейтрино

Данные – $\left\{E_i, N_i, \mu_i\right\}_q$

Проблема «ступеньки»



Часть β-спектра вблизи граничной энергии E_0

$$\mu'(E) = \mu(E) + \Delta \cdot \theta(E_{st} - E)$$

$\mu(E)$ — спектр без аномалии

Δ — высота ступеньки

E_{st} — положение аномалии

$\theta(E_{st} - E)$ — функция Хевисайда

Без ступеньки:

$$\boxed{m_\nu^2 \sim -(10 \div 20)\ \ eV^2}$$

После вычитания эффекта ступеньки:

$$\boxed{m_\nu^2 = -2.3 \pm 2.5_{fit} \pm 2.0_{syst}\, eV^2}$$

$$m_\nu < 2.05\ \ eV$$

В.М. Лобашев, Вестник Российской Академии Наук, 2003, том 70, №1

Проблема «ступеньки»

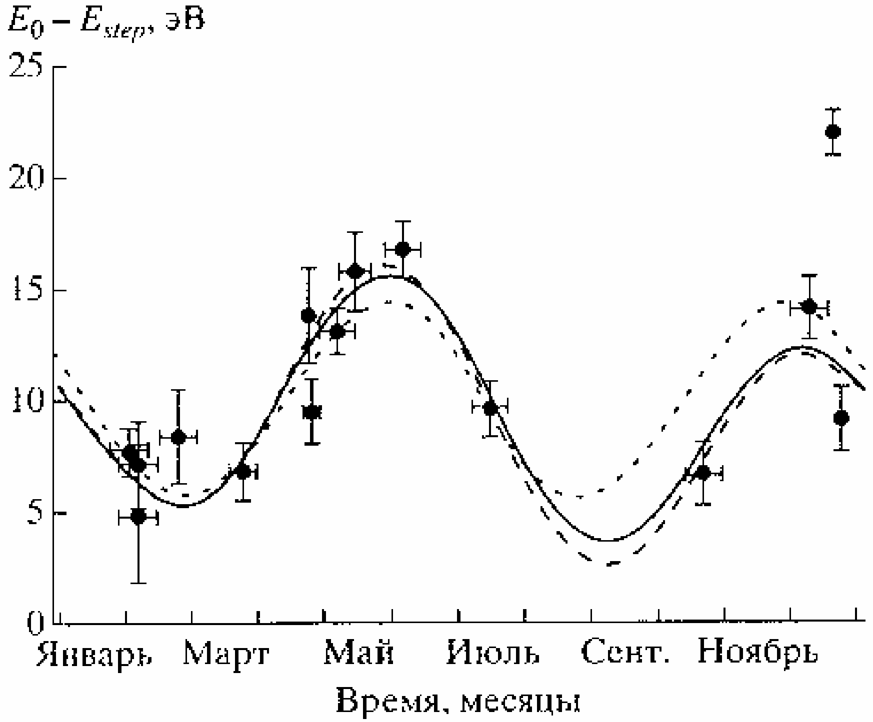


Рис. 2. Зависимость положения ступеньки от времени года. E_0 – граничная энергия, E_{st} – оценка положения ступеньки.(период изменения составляет полгода с примесью годичного)

Необходимы специальные статистические критерии, чувствительные к аномальным вкладам типа ступеньки

- в рамках первой обработки данных возникла необходимость вводить дополнительные параметры
- вторая обработка с помощью метода квазиоптимальных весов, специально оптимизированного для данной задачи + более точная теоретическая модель установки: физически разумный результат без введения ступеньки
- во второй обработке введение дополнительных параметров не потребовалось
- вопрос о наличии ступеньки остаётся открытым, так как фит недостаточно чувствителен к наличию в спектре ступеньки

Статистические критерии

Мера отклонения выборочной функции распределения от теоретической

$S\left(\left\{N_i\right\}\right)$ - критерий = функция данных

Выбираем критическое значение S_0

$S\left(\left\{N_i\right\}\right) \leq S_0 \Longrightarrow$ Данные согласуются с гипотезой

$S\left(\left\{N_i\right\}\right) > S_0 \Longrightarrow$ Гипотеза отвергнута опытом

Стандартные критерии

Статистика критерия Хи-квадрат:

$$\boxed{\chi^2 = \frac{1}{M-1}\sum_i \xi_i^2, \qquad \xi_i = \frac{N_i - \mu_i}{\sqrt{\mu_i}}}$$

Модификация критерия Колмогорова-Смирнова
(критерий *типа* Колмогорова-Смирнова)

Переходим к нормализованным остаткам $N_i \to \xi_i = \dfrac{N_i - \mu_i}{\sqrt{\mu_i}}$

Статистика критерия *типа* Колмогорова-Смирнова:

$$\boxed{D = \max_i \left| F_0(\xi_i) - S_i \right|}$$

$F_0(\xi)$ – теоретическая функция распределения величины ξ

S_i – выборочная функция распределения величины ξ

➡ Стандартная процедура проверки гипотез:

H0: высота ступеньки Δ=0

H1: высота ступеньки Δ>0

Локально наиболее мощный (ЛНМ) критерий

ЛНМ критерий строится на основе метода квазиоптимальных моментов. Пусть данные отфитированы в предположении нулевой гипотезы, получены оценки средних μ_i. Положение ступеньки в спектре будем полагать известным.

$$f_i(N) = \frac{\mu_i'^N e^{-\mu_i'}}{N!}$$ – распределение числа событий (с учётом ступеньки)

$$\mu_i' \to \begin{cases} \mu_i + \Delta_-, & i > m \\ \mu_i + \Delta_+, & i \le m \end{cases}$$

Точка с номером m близка к предполагаемому положению ступеньки

$$E_m \le E_{st} < E_{m+1}$$

Для оценки параметров Δ выпишем оптимальные веса:

$$\omega_i^+(N) = \frac{\partial \ln f_i}{\partial \Delta_+} = \begin{cases} 0, & i > m \\ \dfrac{N}{(\mu_i + \Delta_+)} - 1, & i \le m \end{cases}$$

$$\omega_i^-(N) = \frac{\partial \ln f_i}{\partial \Delta_-} = \begin{cases} 0, & i \le m \\ \dfrac{N}{(\mu_i + \Delta_-)} - 1, & i > m \end{cases}$$

Локально наиболее мощный (ЛНМ) критерий

Приравниваем экспериментальные средние весов теоретическим средним (теоретические средние равны нулю в для таких весов):

$$h_{-}^{\exp} = \frac{1}{M}\sum_{i=1}^{M}\omega_i^{-} = \frac{1}{M}\sum_{i=m+1}^{M}\left(\frac{N_i}{\mu_i + \Delta_-} - 1\right) = 0$$

$$h_{+}^{\exp} = \frac{1}{M}\sum_{i=1}^{M}\omega_i^{+} = \frac{1}{M}\sum_{i=1}^{m}\left(\frac{N_i}{\mu_i + \Delta_+} - 1\right) = 0$$

$\Rightarrow$ Оценки Δ_+ и Δ_-

Статистика ЛНМ критерия:

$$\boxed{\overline{\Delta} = \Delta_+ - \Delta_-}$$

Предполагая малость Δ_+ и Δ_- по сравнению с μ_i, можно получить оценки в явном виде:

$$\begin{cases} \displaystyle\sum_{i=m+1}^{M}\left(\frac{N_i}{\mu_i} - \frac{N_i}{\mu_i^2}\Delta_- - 1\right) = 0 \\ \displaystyle\sum_{i=1}^{m}\left(\frac{N_i}{\mu_i} - \frac{N_i}{\mu_i^2}\Delta_+ - 1\right) = 0 \end{cases} \rightarrow \begin{cases} \Delta_- = \left(\displaystyle\sum_{i=m+1}^{M}\frac{N_i}{\mu_i} - m\right)\Big/\displaystyle\sum_{i=m+1}^{M}\left(\frac{N_i}{\mu_i^2}\right) \\ \Delta_+ = \left(\displaystyle\sum_{i=1}^{m}\frac{N_i}{\mu_i} - m\right)\Big/\displaystyle\sum_{i=1}^{m}\left(\frac{N_i}{\mu_i^2}\right) \end{cases}$$

«Квазиоптимальный» критерий

Строим критерий менее чувствительный к положению ступеньки в спектре, но сохраняющий эффективность, близкую к ЛНМ критерию.

Статистика такого «квазиоптимального» критерия:

$$\boxed{S_{q-opt} = \sum_{i=1}^{M} w_i \cdot \xi_i}$$

$$\xi_i = \frac{N_i - \mu_i}{\sqrt{\mu_i}}$$

$$w_i = \begin{cases} \dfrac{(m-i)}{(m)}, & i \le m, \\ \dfrac{(m-i)}{(M-m)}, & i > m, \end{cases}$$

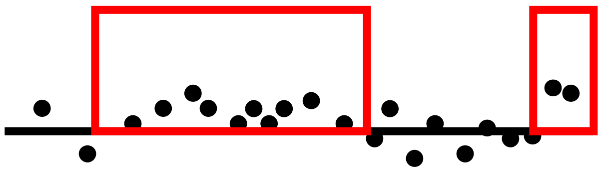
$$E_m \le E_{st} < E_{m+1}$$

«Критерий попарных корреляций соседей»

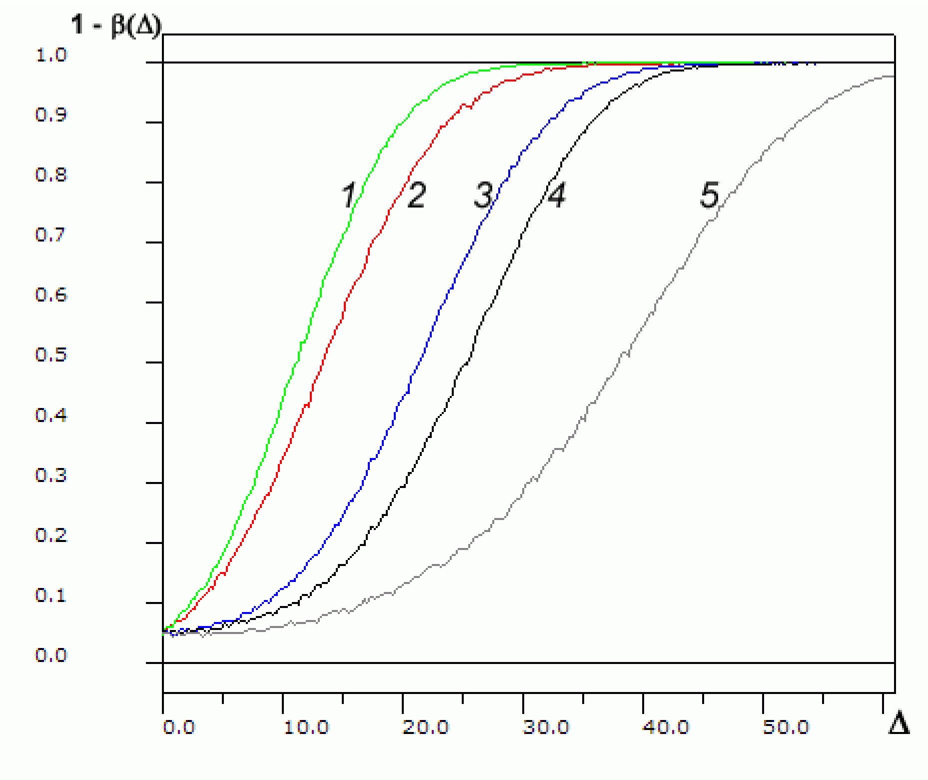
Альтернативный критерий (чувствителен к аномалиям более общего вида)

Статистика критерия попарных корреляций соседей:

$$\boxed{S_{pair} = \sum_i \xi_i \cdot \xi_{i+1}} \qquad \xi_i = \left(N_i - \mu_i\right) \Big/ \sqrt{\mu_i}$$



Сравнение критериев



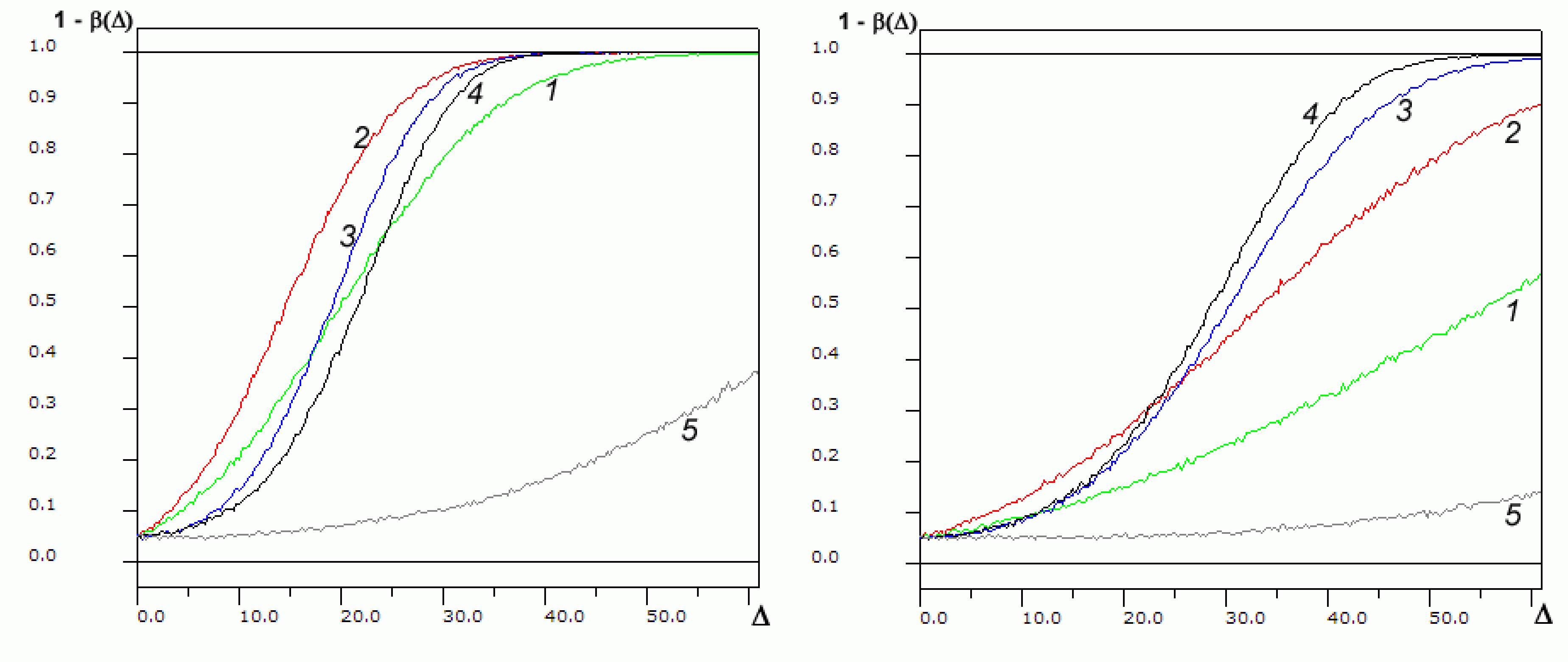
1. ЛНМ критерий
2. «Квазиоптимальный» критерий
3. «Критерий попарных корреляций соседей»
4. Критерий Хи-квадрат
5. Критерий *типа* Колмогорова-Смирнова.

Функция мощности (1 - $\beta(\Delta)$) в точке $\Delta>0$ равна вероятности отклонить нулевую гипотезу ($\Delta=0$), когда верна гипотеза H1 (высота ступеньки равна $\Delta>0$)

•Монте-Карло моделирование

•на основе фита данных 22 сеанса Троицк-ν-масс

Чувствительность к положению ступеньки



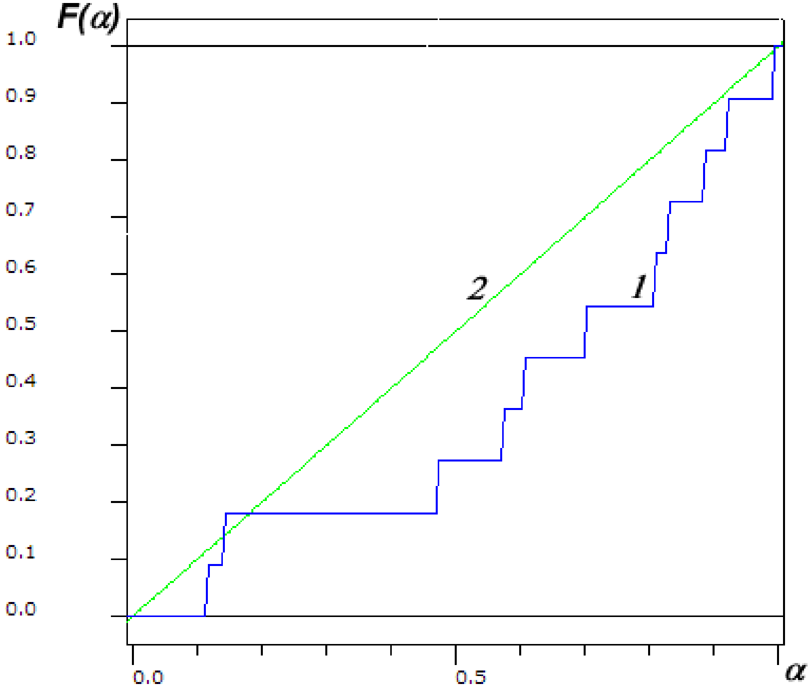
Функции мощности критериев для сдвига фактического положения ступеньки на 12 eV (слева) и 25 eV (справа)

Результаты использования данных

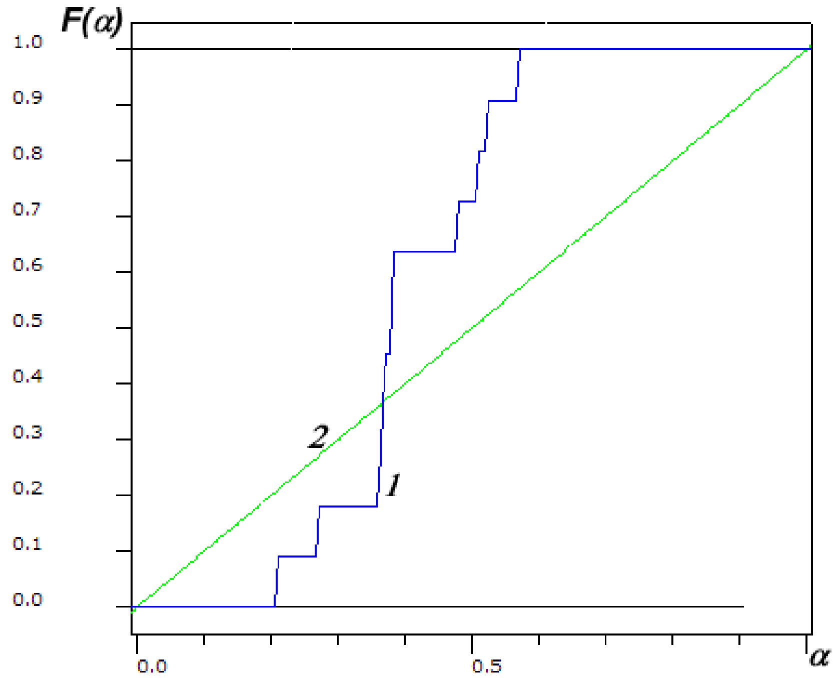
№ сеанса	$\alpha_{\text{q-opt}}$	α_{pair}
22	0.382	0.883
23	0.359	0.571
24_150	0.381	0.471
24_160_mod	0.522	0.604
25_mod	0.478	0.920
28	0.266	0.829
29	0.510	0.141
30	0.371	***0.994***
31	0.365	0.113
33	0.570	0.702
36	0.207	0.810

$$\alpha^{q}_{q-opt} = F_1\left(S^{q}_{q-opt}\right) \qquad \alpha^{q}_{pair} = F_2\left(S^{q}_{pair}\right)$$

Исследование полученных наборов значений на равномерность



$$\{\alpha_{pair}\}$$



$$\{\alpha_{q-opt}\}$$

Исследование полученных наборов значений на равномерность и симметричность

Критерий	$\alpha_{q\text{-}opt}$	α_{pair}	Критическая область 95%	Критическая область 99%
$\chi 2$	2.273	2.273	<3.841	<6.635
Колмогоров	**0.430**	0.299	<0.391	<0.468
Шерман	**0.492**	0.334	<0.469	<0.521
Моран	**13.013**	5.243	[2.704, 11.375]	[1.81, 14.21]
Ченг-Спиринг	1.56	1.25	[0.92, 1.73]	[0.83, 1.99]
Фроцини	**0.611**	0.464	<0.581	<0.749
Янг	**0.241**	0.435	[0.3, 0.6]	[0.2, 0.6]
Знаковый	3	3	>1	>0
Вилкоксон	**58**	49	>51	>58

Обсуждение результатов

- Построены три специальных статистических критерия:
  - ❖ЛНМ критерий, сильно зависящий от положения аномалии в спектре
  - ❖квазиоптимальный критерий: удобный, эффективный, слабочувствительный к положению ступеньки
  - ❖«критерий попарных корреляций соседей», эффективен для поиска аномалий более общего вида

- Получены экспериментальные значения критериев для данных 11 сеансов Троицк-ν-масс. Лишь одно из значений отклоняет гипотезу об отсутствии ступеньки с большой доверительной вероятностью (99%)

- Проведено статистическое исследование наборов значений критериев на предмет наличия аномальных характеристик, присутствующих во всех сеансах. Отклонений на уровне 99% нет, а доверительного уровня в 95% недостаточно для указания на аномальные вклады