Асимптотические формулы для оценки статистической значимости в коллайдерных экспериментах

Д. Э. Горин,* О. С. Василевский,† Л. В. Дудко,‡ Э. Э. Абасов§
Московский государственный университет имени М. В. Ломоносова,
физический факультет, кафедра физики атомного ядра и квантовой теории столкновений
Россия, 119991, Москва, Ленинские горы, д. 1, стр. 2
(Поступила в редакцию 26.12.2023; подписана в печать 22.04.2024)

В настоящей работе представлены уточнения и развитие ранее предложенных асимптотических методов статистического анализа в физике высоких энергий, основанных на использовании функции правдоподобия. С учетом результатов Уилкса и Вальда представлен явный вывод ранее полученных асимптотических формул для оценки статистической значимости открытия новых явлений и постановки верхних ограничений на параметры, характеризующие модели «новой физики». Полученные формулы в совокупности с применением так называемого «азимовского набора данных» представляют большой интерес в силу своей простоты и возможности учета систематических неопределенностей. Также предложена методология применения этих формул и проведена проверка их точности и границ применимости с помощью Монте–Карло моделирования.

PACS: 02.70.Rr, 02.50.Cw, 02.50.Ng, 02.30.Mv, 02.70.Uu УДК: 519.2, 519.246, 51-73
Ключевые слова: асимптотические формулы, статистический анализ, статистическая значимость, чувствительность эксперимента, функция правдоподобия, азимовский набор данных, Монте–Карло моделирование.

ВВЕДЕНИЕ

В рамках исследований, проводимых на Большом адронном коллайдере (LHC), особое внимание уделяется поиску процессов, предсказываемых в рамках моделей «новой физики». Определение статистической значимости обнаруженных сигналов осуществляется с помощью p-значения или эквивалентной гауссовой значимости. Традиционно такой анализ предполагает использование методов Монте–Карло, требующих значительных вычислительных ресурсов и времени.

В настоящей работе проведено уточнение и дальнейшее развитие асимптотическоого подхода, предложенного Коуэном и др. [1], основанного на результатах Уилкса [2] и Вальда [3], позволяющего оценивать статистическую значимость без обращения к методам Монте–Карло. Используемый подход включает оценку ожидаемой значимости с помощью замены экспериментальных наборов данных одним искусственным набором данных с подавленными статистическими флуктуациями, именуемым «азимовским набором данных».

Настоящая работа является по своей сути детализацией работы [1], и ее основным результатом является явный вывод формул для оценки статистической значимости открытия и исключения сигнала, учитывающих систематическую неопределенность фона и позволяющих совершать быстрые и асимптотически точные расчеты. Помимо этого в работе подробно описано применение этих формул и проведена численная проверка асимптотических значений путем проведения Монте–Карло моделирования псеведоэкспериментов. Упомянутые формулы получаются непосредственным применением методики, описанной Коуэном и др. [1], и потому для облегчения понимания в настоящей работе сохранены ранее используемые обозначения.

В разделе I дается обзор представленных в работе Коуэна и др. теоретических основ — описываются математический формализм статистического анализа эксперимента, функция правдоподобия, аппроксимация распределения используемой статистики и принципы применения азимовского набора данных. Далее, в разделах II и III вводятся статистики q_0 и q_μ для открытия и исключения сигнала соответственно и с учетом описанных в предыдущем разделе приближений проводится непосредственный вывод формул для оценки соответствующих статистических значимостей. В разделе IV предлагается обобщение этих формул для более общего случая статистического сравнения форм распределений, представленных в виде гистограмм. В заключительном разделе V показано применение упомянутых формул на примере простейших счетных экспериментов, и сравнение аналитического результата с Монте–Карло моделированием, исследуется их точность и границы применимости.

1. СТАТИСТИЧЕСКИЙ АНАЛИЗ ЭКСПЕРИМЕНТА

Статистический анализ в физике высоких энергий обычно проводится в контексте проверки статистических гипотез. Для оценки статистической значимости открытия нового сигнального процесса задается нулевая гипотеза H_0, предполагающая наличие только фоновых процессов, которая подвергается проверке против альтернативной гипотезы H_1, включающей как фон, так и предполагаемый новый сигнал. Если речь

* E-mail: gorin.de20@physics.msu.ru
† E-mail: vasilevskii.os20@physics.msu.ru
‡ E-mail: lev.dudko@cern.ch
§ E-mail: emil@abasov.ru

идет о постановке ограничений на параметры изучаемых моделей, то комбинация «фон + сигнал» рассматривается как нулевая гипотеза H_0, тогда как чисто фоновое предположение выступает в роли альтернативы.

Для количественной оценки степени соответствия между наблюдаемыми данными и заданной гипотезой H используется p-значение. Оно представляет собой вероятность получить данные, которые в равной или большей степени расходятся с предсказаниями гипотезы H, при условии, что эта гипотеза верна. Если вычисленное p-значение оказывается меньше заранее установленного порога, то гипотеза H отклоняется.

p-значение можно пересчитать в соответствующую ему статистическую значимость Z. Эта значимость определяется по правилу, согласно которому случайная величина, подчиняющаяся нормальному распределению Гаусса и отклоняющаяся на величину Z стандартных отклонений вверх от своего среднего, имеет ту же вероятность p в правой части своего распределения. Другими словами, Z рассчитывается по формуле:

$$Z = \Phi^{-1}(1-p), \tag{1}$$

где Φ^{-1} — функция, обратная к функции стандартного гауссова распределения.

В последующих разделах будет также идти речь об ожидаемой значимости. Стоит заметить, что из-за нелинейности связи между Z и p, ожидаемая значимость, полученная для некоторой гипотезы, не будет совпадать со значимостью Z, рассчитанной по формуле (1) из ожидаемого p-значения. Однако же медианные значения Z и p будут удовлетворять равенству (1) вследствие его монотонности. По этой причине в дальнейшем под ожидаемой значимостью будет всегда пониматься медианная значимость.

1.1. Функция правдоподобия

Разбив сигнальную область на N бинов, можно в каждом из них провести измерение числа событий n_j , $j = 1, ..., N$. Ожидаемое (среднее) же число событий в каждом бине можно вычислить по формуле $E[n_j] = \mu s_j + b_j$, где s_j и b_j представляют собой средние числа сигнальных и фоновых событий соотвествено, а μ является силой сигнала — параметром, характеризующим возможный вклад сигнала.

Помимо измерений в сигнальной области можно также провести измерения в дополнительной, так называемой, контрольной области, где отсутствует сигнал, чтобы уточнить значения случайных «мешающих» параметров $\boldsymbol{\theta}$, от которых зависят величины s_j и b_j. Проводя такие измерения, можно найти число событий m_k , $k = 1, ..., M$ в каждом из M бинов контрольной области, а их ожидаемое число может быть найдено как $E[m_k] = u_k(\boldsymbol{\theta})$.

Функция правдоподобия в таком случае записывается как произведение пуассоновских вероятностей для всех N бинов сигнальной и M бинов контрольной областей:

$$L(\mu, \boldsymbol{\theta}) = \prod_{j=1}^{N} \frac{(\mu s_j + b_j)^{n_j}}{n_j!} e^{-(\mu s_j + b_j)} \prod_{k=1}^{M} \frac{u_k^{m_k}}{m_k!} e^{-u_k} . \tag{2}$$

Согласно лемме Неймана-Пирсона в задаче проверки простой статистической гипотезы против простой альтернативы критерий, основанный на отношении функций правдоподобия, является наиболее мощным, т.е. при заданной ошибке первого рода ошибка второго рода минимальна. Поэтому в качестве статистики имеет смысл взять следующее отношение:

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\boldsymbol{\theta}}})}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})} , \tag{3}$$

где $\hat{\mu}$ и $\hat{\boldsymbol{\theta}}$ — оценки максимального правдоподобия парамтеров μ и $\boldsymbol{\theta}$, то есть такие их значения, что функция правдоподобия L максимальна, а $\hat{\hat{\boldsymbol{\theta}}}(\mu)$ — значение $\boldsymbol{\theta}$, условно максимизирующее L для заданного μ. В дальнейшем под ожидаемым значением силы сигнала μ будет пониматься это же значение $\hat{\mu}$.

1.2. Аппроксимация распределения $\boldsymbol{\lambda(\mu)}$

Как будет показано далее, для расчета p-значения гипотезы необходимо знать плотность вероятности $f(\lambda(\mu)|\mu)$ используемой статистики. Здесь аргумент λ относится к проверяемому в гипотезе значению силы сигнала μ, а второй аргумент плотности вероятности — к предполагаемому в данных.

Кроме того, для определения ожидаемой значимости в случае, когда данные соответствуют силе сигнала, отличной от проверяемой в гипотезе, требуется знать функцию плотности вероятности $f(\lambda(\mu)|\mu')$, где $\mu' \neq \mu$.

Пусть проверяется гипотеза о силе сигнала μ, и пусть данные соответствуют силе сигнала μ'. Искомая плотность вероятности $f(\lambda(\mu)|\mu')$ может быть найдена с помощью результатов Вальда [3], который показал, что в случае одного параметра интереса:

$$-2\ln\lambda(\mu) = \frac{(\mu - \hat{\mu})^2}{\sigma^2} + \mathcal{O}(1/\sqrt{N}) , \tag{4}$$

где $\hat{\mu}$ распределена по Гауссу со средним μ' и стандартным отклонением σ, а N — объем выборки данных.

Пренебрегая асимптотическим членом $\mathcal{O}(1/\sqrt{N})$, можно показать, что случайная величина $t_\mu = -2\ln\lambda(\mu)$ подчиняется нецентральному распределению хи-квадрат с одной степенью свободы [1]:

$$f(t_\mu, \Lambda) = \frac{1}{2\sqrt{t_\mu}} \frac{1}{\sqrt{2\pi}} \left[\exp\left(-\frac{1}{2} \left(\sqrt{t_\mu} + \sqrt{\Lambda} \right)^2 \right) + \right.$$
$$\left. + \exp\left(-\frac{1}{2} \left(\sqrt{t_\mu} - \sqrt{\Lambda} \right)^2 \right) \right] , \tag{5}$$

где $\Lambda = \frac{(\mu\ ---\mu')^2}{\sigma^2}$ — параметр нецентральности.

В частном случае, когда $\mu' = \mu$, параметр нецентральности $\Lambda = 0$ и $-2\ln\lambda(\mu)$ подчиняется в пределе распределению хи-квадрат с одной степенью свободы, что было показано Уилксом [2].

1.3. Азимовский набор данных

Для расчета ожидаемой значимости очень удобным является особый искусственный набор данных, называемый «азимовским набором данных» [1], в котором устранены все статистические флуктуации. То есть для азимовского набора данных, соответствующего силе сигнала μ', выполняется:

$$n_{i,A} = E[n_i] = \mu' s_i(\boldsymbol{\theta}) + b_i(\boldsymbol{\theta}) \tag{6}$$

$$m_{i,A} = E[m_i] = u_i(\boldsymbol{\theta})\,. \tag{7}$$

Таким образом, при использовании азимовского набора данных оценки максимального правдоподобия $\hat{\mu}$ и $\hat{\boldsymbol{\theta}}$ параметров μ и $\boldsymbol{\theta}$ равны их ожидаемым значениям:

$$\hat{\mu} = \mu' \tag{8}$$

$$\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}\,. \tag{9}$$

С помощью азимовского набора данных можно также построить «азимовскую функцию правдоподобия» L_A и с помощью нее найти:

$$\lambda_A(\mu) = \frac{L_A(\mu, \hat{\hat{\boldsymbol{\theta}}})}{L_A(\hat{\mu}, \hat{\boldsymbol{\theta}})} = \frac{L_A(\mu, \hat{\hat{\boldsymbol{\theta}}})}{L_A(\mu', \boldsymbol{\theta})}\,, \tag{10}$$

где в последнем равенстве использованы соотношения (8) и (9).

2. ОЦЕНКА СТАТИСТИЧЕСКОЙ ЗНАЧИМОСТИ ОТКРЫТИЯ СИГНАЛА

2.1. Статистика q_0

Для открытия новых процессов нужна статистика, при использовании которой фоновая гипотеза $\mu = 0$ будет проверяться против набора альтернативных гипотез с $\mu \geq 0$. Отвержение контрольной гипотезы в таком случае и будет означать открытие нового сигнала.

Подходящей для этих целей является статистика q_0 [1]:

$$q_0 = \begin{cases} -2\ln\lambda(0)\ , \ \hat{\mu} \geq 0 \\ 0\ ,\ \hat{\mu} < 0 \end{cases}\,, \tag{11}$$

где $\lambda(0)$ задается выражением (3) с $\mu = 0$.

Здесь предполагается, что фоновая гипотеза не подтверждается только в том случае, когда $\hat{\mu} > 0$. Значение $\hat{\mu} < 0$ в типовых задачах открытия «новой физики» не рассматривается, поскольку в таком случае данные свидетельствуют скорее не о наличии сигнала, а о некой систематической ошибке, за исключением некоторых специальных задач.

Таким образом, с ростом числа наблюдаемых событий, то есть с ростом $\hat{\mu}$, увеличивается и значение q_0, что означает увеличение несоотвествия между наблюдаемыми данными и контрольной гипотезой $\mu = 0$.

Чтобы количественно оценить величину этого несоответствия, можно рассчитать p-значение, используя наблюдаемое значение q_0:

$$p_0 = \int_{q_{0,\mathrm{obs}}}^{\infty} f(q_0|0)dq_0\,, \tag{12}$$

где $f(q_0|0)$ — плотность вероятности статистики q_0 в предположении контрольной гипотезы $\mu = 0$.

2.2. Аппроксимация распределения статистики q_0

В рамках приближения (4) статистика q_0 в соответствии с определением (11) записывается следующим образом:

$$q_0 = \begin{cases} \hat{\mu}^2/\sigma^2\ ,\ \hat{\mu} \geq 0 \\ 0\ ,\ \hat{\mu} < 0 \end{cases}\,, \tag{13}$$

где $\hat{\mu}$ распределена по Гауссу со средним μ' и стандартным отклонением σ.

Можно показать, что плотность вероятности q_0 в таком случае можно записать как [1]:

$$f(q_0|\mu') = \left(1 - \Phi\left(\frac{\mu'}{\sigma}\right)\right)\delta(q_0) + \\ + \frac{1}{2}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{q_0}}\exp\left[-\frac{1}{2}\left(\sqrt{q_0} - \frac{\mu'}{\sigma}\right)^2\right]\,, \tag{14}$$

откуда соотвествующая функция распределения:

$$F(q_0|\mu') = \Phi\left(\sqrt{q_0} - \frac{\mu'}{\sigma}\right)\,. \tag{15}$$

Тогда для частного случая $\mu' = 0$:

$$F(q_0|0) = \Phi\left(\sqrt{q_0}\right)\,. \tag{16}$$

В таком случае p-значение гипотезы $\mu = 0$ в соотвествии с (12):

$$p_0 = 1 - F(q_0|0)\,, \tag{17}$$

и следовательно из определения (1) получается выражение для значимости:

$$Z_0 = \Phi^{-1}(1 - p_0) = \sqrt{q_0}\,. \tag{18}$$

2.3. Вывод формулы для ожидаемой значимости Z_{disc} открытия сигнала

Для вычисления чувствительности эксперимента важна не значимость, рассчитанная на одном наборе данных, а ожидаемая (точнее, как упоминалось ранее, медианная) значимость, с которой можно было бы отвергнуть различные значения μ. Непосредственно для открытия нового сигнала необходимо знать медианную значимость при условии, что гипотеза $\mu = 1$ верна, с которой можно было бы отвергнуть фоновую гипотезу $\mu = 0$.

Пусть в сигнальной и контрольной областях проводится по одному измерению количества событий n и m соответственно. Будем считать, что имеется только один «мешающий» параметр $\theta = b$, и предположим также, что m подчиняется распределению Пуассона со средним τb, где τ — сила фона. Параметр τ можно использовать для введения в статистическую модель систематической неопределенности предсказания фона.

В таком случае, функция правдоподобия записывается следующим образом:

$$L(\mu, b) = \frac{(\mu s + b)^n}{n!} e^{-(\mu s+b)} \frac{(\tau b)^m}{m!} e^{-\tau b}. \qquad (19)$$

Для расчета статистик открытия и исключения понадобятся выражения для параметров $\hat{\mu}$ и $\hat{b}$, совместно максимизирующих функцию правдоподобия (19), а также параметра $\hat{\hat{b}}$, максимизирующего функцию правдоподобия (19) для заданного μ [1]:

$$\hat{\mu} = \frac{n - m/\tau}{s}, \qquad (20)$$

$$\hat{b} = \frac{m}{\tau}, \qquad (21)$$

$$\hat{\hat{b}} = \frac{n + m - (1-\tau)\mu s}{2(1+\tau)} + \left[\frac{(n + m - (1+\tau)\mu s)^2 + 4(1+\tau)m\mu s}{4(1+\tau)^2}\right]^{\frac{1}{2}}. \qquad (22)$$

Для расчета статистической значимости Z_{disc} необходимо пользоваться статистикой q_0, определенной в (11):

$$q_0 = \begin{cases} -2\ln\lambda(0)\,, & \hat{\mu} \geq 0, \\ 0\,,\ \hat{\mu} < 0, \end{cases} \quad \text{где } \lambda(\mu) = \frac{L(\mu, \hat{\hat{b}}(\mu))}{L(\hat{\mu}, \hat{b})}.$$

Подставляя выражения (20), (21) и (22) при $\mu = 0$, можно получить:

$$\begin{aligned} -2\ln\lambda(0) &= -2\ln\frac{L(0,\hat{\hat{b}}(0))}{L(\hat{\mu},\hat{b})} = -2\ln\frac{\left(\frac{n+m}{1+\tau}\right)^n \frac{e^{-\frac{n+m}{1+\tau}}}{n!}\left(\tau\frac{n+m}{1+\tau}\right)^m \frac{e^{-\tau\frac{n+m}{1+\tau}}}{m!}}{\left(\frac{n-m/\tau}{s}s + \frac{m}{\tau}\right)^n \frac{e^{-\left(\frac{n-m/\tau}{s}s+\frac{m}{\tau}\right)}}{n!}\left(\tau\frac{m}{\tau}\right)^m\frac{e^{-\left(\tau\frac{m}{\tau}\right)}}{m!}} = \\ &= -2\ln\left[\frac{\left(\frac{n+m}{1+\tau}\right)^n}{n^n}\frac{e^{-\frac{n+m}{1+\tau}}}{e^{-n}}\frac{\left(\tau\frac{n+m}{1+\tau}\right)^m}{m^m}\frac{e^{-\tau\frac{n+m}{1+\tau}}}{e^{-m}}\right] = \\ &= -2\left[n\ln\left(\frac{n+m}{(1+\tau)n}\right) - \frac{n+m}{1+\tau} + n + m\ln\left(\frac{\tau(n+m)}{(1+\tau)m}\right) - \tau\frac{n+m}{1+\tau} + m\right] = \\ &= -2\left[n\ln\left(\frac{n+m}{(1+\tau)n}\right) + m\ln\left(\frac{\tau(n+m)}{(1+\tau)m}\right) + \overbrace{\frac{n(1+\tau) - n - m + m(1+\tau) - \tau(n+m)}{1+\tau}}^{=0}\right]. \end{aligned}$$

Таким образом, выражение для статистики q_0 можно записать в следующем виде:

$$q_0 = \begin{cases} -2\left[n\ln\left(\frac{n+m}{(1+\tau)n}\right) + m\ln\left(\frac{\tau(n+m)}{(1+\tau)m}\right)\right]\,,\ n \geq \frac{m}{\tau} \\ 0\,,\ n < \frac{m}{\tau} \end{cases}.$$

В силу того, что q_0 монотонна по $\hat{\mu}$, ее медианное значение можно записать через соответствующую статистику $q_{0,A}$, построенную на азимовском наборе данных согласно определениям (10) и (11):

$$\text{med}[q_0(\hat{\mu})] = q_0(\text{med}[\hat{\mu}]) = q_0(\mu') = q_{0,A}\,. \qquad (23)$$

В свою очередь, в силу монотонности значимости Z_0 по q_0 (что видно из выражения (18)), медианная значимость является значимостью медианного значения статистики. Таким образом, для медианной значимости открытия получается следующее выражение:

$$\text{med}[Z_0|1] = \sqrt{\text{med}[q_0]} = \sqrt{q_{0,A}}\,, \qquad (24)$$

где статистика $q_{0,A}$ построена на азимовском наборе данных с силой сигнала $\mu' = 1$.

Таким образом, для расчета ожидаемой значимости открытия необходимо воспользоваться выражением (24) и азимовским набором данных с силой сигнала $\mu' = 1$, то есть сделать замену $n = s + b$ и $m = \tau b$:

$$Z_{disc} = \mathrm{med}[Z_0|1] = \sqrt{-2\left[(s+b)\ln\left(\frac{s+(1+\tau)b}{(1+\tau)(s+b)}\right) + \tau b \ln\left(1+\frac{s}{(1+\tau)b}\right)\right]}\,. \tag{25}$$

Полезно переписать выражение (25), используя систематическую ошибку δ фона.

Поскольку $\hat{b} = \frac{m}{\tau}$, а дисперсия m равняется ее среднему, τb, то дисперсия $\hat{b}$:

$$D\hat{b} \equiv \sigma_b^2 = \frac{1}{\tau^2} Dm = \frac{b}{\tau}\,.$$

Таким образом, систематическая ошибка δ:

$$\delta \equiv \frac{\sigma_b}{b} = \frac{1}{\sqrt{\tau b}}\,. \tag{26}$$

Подставляя (26) в (25):

$$\begin{aligned}
Z_{disc} &= \sqrt{-2\left[(s+b)\ln\left(\frac{s+(1+\tau)b}{(1+\tau)(s+b)}\right) + \tau b \ln\left(1+\frac{s}{(1+\tau)b}\right)\right]} = \\
&= \sqrt{-2\left[(s+b)\ln\left(\frac{s+b+\tau b}{s+\tau s+\tau b+b}\right) + \tau b \ln\left(1+\frac{s}{b+\tau b}\right)\right]} = \\
&= \sqrt{-2\left[(s+b)\ln\left(\frac{s+b+1/\delta^2}{s+\tau s+1/\delta^2+b}\right) + \frac{1}{\delta^2} \ln\left(1+\frac{s}{b+1/\delta^2}\right)\right]} = \\
&= \sqrt{-2\left[(s+b)\ln\left(\frac{1+\delta^2(s+b)}{1+\delta^2(s+b)+s/b}\right) + \frac{1}{\delta^2} \ln\left(1+\delta^2\frac{s}{1+\delta^2 b}\right)\right]} = \\
&= \sqrt{-2\left[(s+b)\ln\left(\frac{b+\delta^2 b(s+b)}{b+\delta^2 b(s+b)+s}\right) + \frac{1}{\delta^2} \ln\left(1+\delta^2\frac{s}{1+\delta^2 b}\right)\right]} \quad ,
\end{aligned}$$

можно получить окончательное выражение для ожидаемой значимости открытия с учетом вклада систематической ошибки предсказания фона:

$$Z_{disc} = \sqrt{2\left[(s+b)\ln\left(\frac{(s+b)(1+\delta^2 b)}{b+\delta^2 b(s+b)}\right) - \frac{1}{\delta^2} \ln\left(1+\delta^2\frac{s}{1+\delta^2 b}\right)\right]}\,. \tag{27}$$

В случае, если фон известен точно, $\delta \to 0$ и (27) переходит в:

$$\begin{aligned}
Z_{disc} &= \sqrt{2\left[(s+b)\ln\left(1+\frac{s}{b}\right) - \lim_{\delta\to 0}\left(\frac{1}{\delta^2}\ln\left(1+\delta^2\frac{s}{1+\delta^2 b}\right)\right)\right]} = \\
&= \sqrt{2\left[(s+b)\ln\left(1+\frac{s}{b}\right) - \lim_{\delta\to 0}\left(\frac{\frac{\partial}{\partial\delta}\left(\ln\left(1+\delta^2\frac{s}{1+\delta^2 b}\right)\right)}{\frac{\partial}{\partial\delta}\delta^2}\right)\right]} = \\
&= \sqrt{2\left[(s+b)\ln\left(1+\frac{s}{b}\right) - \lim_{\delta\to 0}\left(\frac{2\delta s}{2\delta(1+2b\delta^2+b^2\delta^4+s\delta^2+sb\delta^4)}\right)\right]} = \\
&= \sqrt{2\left[(s+b)\ln\left(1+\frac{s}{b}\right) - s\right]}\,.
\end{aligned} \tag{28}$$

3. ОЦЕНКА СТАТИСТИЧЕСКОЙ ЗНАЧИМОСТИ ПОСТАНОВКИ ВЕРХНИХ ОГРАНИЧЕНИЙ

3.1. Статистика q_μ

Для постановки верхних ограничений на силу сигнала необходимо проверить нулевую гипотезу о наличии сигнала против альтернативной контрольной гипотезы и найти такое значение μ, для которого p-значение нулевой гипотезы будет соответствовать заданному уровню доверия.

В таком случае удобно пользоваться статистикой q_μ [1]:

$$q_\mu = \begin{cases} -2\ln\lambda(\mu)\ , \ \hat{\mu} \le \mu \\ 0\ ,\ \hat{\mu} > \mu \end{cases} \quad , \tag{29}$$

где $\lambda(\mu)$ задается выражением (3).

Причина зануления q_μ в случае $\hat{\mu} > \mu$ заключается в том, что данные с такой силой сигнала не рассматриваются как более сильное отклонение от гипотезы μ, чем фактически наблюдаемые результаты. По этой причине они не включаются в критическую область гипотезы μ.

Так же как и в случае со статистикой q_0, чем больше значение q_μ, тем больше несоответствие между наблюдаемыми данными и гипотетическим значением μ. Аналогичным образом можно количественно оценить величину этого несоответствия, рассчитав p-значение, используя наблюдаемое значение q_μ:

$$p_\mu = \int_{q_{\mu,\mathrm{obs}}}^{\infty} f(q_\mu|\mu) dq_\mu\,, \tag{30}$$

где $f(q_\mu|\mu)$ — плотность вероятности статистики q_μ в предположении гипотезы μ.

3.2. Аппроксимация распределения статистики q_μ

В рамках приближения (4) статистика q_μ в соответствии с определением (29) записывается следующим образом:

$$q_\mu = \begin{cases} \frac{(\mu-\hat{\mu})^2}{\sigma^2}\ , \ \hat{\mu} \le \mu \\ 0\ ,\ \hat{\mu} > \mu \end{cases} \quad , \tag{31}$$

где $\hat{\mu}$ распределена по Гауссу со средним μ' и стандартным отклонением σ.

Плотность вероятности q_μ можно записать как [1]:

$$f(q_0|\mu') = \Phi\left(\frac{\mu'-\mu}{\sigma}\right)\delta(q_\mu) + \\ + \frac{1}{2}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{q_\mu}}\exp\left[-\frac{1}{2}\left(\sqrt{q_\mu} - \frac{\mu-\mu'}{\sigma}\right)^2\right] \quad , \tag{32}$$

откуда соотвествующая функция распределения:

$$F(q_\mu|\mu') = \Phi\left(\sqrt{q_\mu} - \frac{\mu-\mu'}{\sigma}\right)\,, \tag{33}$$

и для частного случая $\mu' = \mu$ получается:

$$F(q_\mu|\mu) = \Phi\left(\sqrt{q_\mu}\right)\,. \tag{34}$$

Тогда в соответсвии с (30) p-значение гипотезы μ :

$$p_\mu = 1 - F(q_\mu|\mu) = 1 - \Phi\left(\sqrt{q_\mu}\right)\,, \tag{35}$$

и следовательно выражение для значимости записывается как:

$$Z_\mu = \Phi^{-1}(1-p_\mu) = \sqrt{q_\mu}\,. \tag{36}$$

3.3. Вывод формулы для ожидаемой значимости Z_{excl} исключения сигнала

При постановке верхних ограничений чувствительность эксперимента характеризуется медианной значимостью, в предположении гипотезы $\mu = 0$, с которой отвергается некоторое ненулевое значение μ. Обычно наибольший интерес представляет случай отвержения значения $\mu = 1$, называемый «исключением» (exclusion).

Для расчета статистической значимости Z_{excl} необходимо пользоваться статистикой q_μ, определенной в (29):

$$q_\mu = \begin{cases} -2\ln\lambda(\mu)\ , \ \hat{\mu} \le \mu \\ 0\ ,\ \hat{\mu} > \mu, \end{cases}$$

где в случае исключения следует положить $\mu = 1$.

Подставляя выражения (20) и (21), можно получить:

$$-2\ln\lambda(1) = -2\ln\frac{L(1,\hat{\hat{b}}(1))}{L(\hat{\mu},\hat{b})} == -2\ln\frac{\frac{(s+\hat{\hat{b}}(1))^n}{n!}e^{-(s+\hat{\hat{b}}(1))}\frac{(\tau\hat{\hat{b}}(1))^m}{m!}e^{-\tau\hat{\hat{b}}(1)}}{\left(\frac{n-m/\tau}{s}s+\frac{m}{\tau}\right)^n\frac{e^{-\left(\frac{n-m/\tau}{s}s+\frac{m}{\tau}\right)}}{n!}\left(\tau\frac{m}{\tau}\right)^m\frac{e^{-\left(\tau\frac{m}{\tau}\right)}}{m!}} = \\ = -2\ln\left[\frac{(s+\hat{\hat{b}}(1))^n}{n^n}\frac{e^{-(s+\hat{\hat{b}}(1))}}{e^{-n}}\frac{(\tau\hat{\hat{b}}(1))^m}{m^m}\frac{e^{-\tau\hat{\hat{b}}(1)}}{e^{-m}}\right] = \\ = -2\left[n\ln\left(\frac{s+\hat{\hat{b}}(1)}{n}\right) + m\ln\left(\frac{\tau\hat{\hat{b}}(1)}{m}\right) - (s+\hat{\hat{b}}(1)) + n - \tau\hat{\hat{b}}(1) + m\right]\,.$$

Таким образом, выражение для статистики q_1 можно записать в следующем виде:

$$q_1 = \begin{cases} -2\left[n\ln\left(\frac{s+\hat{\hat{b}}(1)}{n}\right) + m\ln\left(\frac{\tau\hat{\hat{b}}(1)}{m}\right) - (s+\hat{\hat{b}}(1)) + n - \tau\hat{\hat{b}}(1) + m\right]\ ,\ n - m/\tau \le s \\ 0\ ,\ n - m/\tau > s \end{cases} .$$

Используя азимовский набор данных с силой сигнала $\mu' = 0$ для медианного значения статистики q_μ и медианной значимости исключения сигнала можно получить аналогичные (23) и (24) выражения:

$$\mathrm{med}[q_\mu(\hat{\mu})] = q_\mu(\mathrm{med}[\hat{\mu}]) = q_\mu(\mu') = q_{\mu,A} \tag{37}$$

$$\mathrm{med}[Z_1|0] = \sqrt{\mathrm{med}[q_1]} = \sqrt{q_{1,A}}\,. \tag{38}$$

Таким образом, для расчета ожидаемой значимости исключения сигнала необходимо воспользоваться выражением (38) и азимовским набором данных с силой сигнала $\mu' = 0$, то есть сделать замену $n = b$ и $m = \tau b$.

Стоит отдельно обратиться к выражению (22) при $\mu = 1$ и переписать его, используя систематическую ошибку $\delta = \frac{1}{\sqrt{\tau b}}$ фона:

$$\begin{aligned}
\hat{\hat{b}}(1) &= \frac{n+m-(1-\tau)s}{2(1+\tau)} + \left[\frac{(n+m-(1+\tau)s)^2 + 4(1+\tau)ms}{4(1+\tau)^2}\right]^{\frac{1}{2}} = \\
&= \frac{n+m-(1-\tau)s + \sqrt{(n+m-(1+\tau)s)^2 + 4(1+\tau)ms}}{2(1+\tau)} = \\
&= \frac{b+\tau b-(1-\tau)s + \sqrt{(b+\tau b-(1+\tau)s)^2 + 4(1+\tau)\tau bs}}{2(1+\tau)} = \\
&= \frac{b-s}{2} + \frac{1}{2}\sqrt{\left(\frac{b+\tau b-(1+\tau)s}{1+\tau}\right)^2 + \frac{4\tau bs}{1+\tau}} = \frac{b-s}{2} + \frac{1}{2}\sqrt{(b-s)^2 + \frac{4\tau bs}{1+\tau}} = \\
&= \frac{b-s}{2} + \frac{1}{2}\sqrt{(b-s)^2 + \frac{4s/\delta^2}{1+\tau}} = \frac{b-s}{2} + \frac{1}{2}\sqrt{(b-s)^2 + \frac{4s}{\delta^2+1/b}} = \frac{b-s}{2} + \frac{1}{2}\sqrt{(b-s)^2 + \frac{4sb}{\delta^2 b+1}} = \\
&= \frac{b-s}{2} + \frac{1}{2}\sqrt{(b+s)^2 - 4sb + \frac{4sb}{\delta^2 b+1}} = \frac{b-s}{2} + \frac{1}{2}\sqrt{(b+s)^2 - \frac{4\delta^2 b^2 s}{\delta^2 b+1}} = \frac{b-s+x}{2}, \\
&\text{где}\quad x = \sqrt{(b+s)^2 - \frac{4\delta^2 b^2 s}{\delta^2 b+1}}\,.
\end{aligned}$$

Таким образом, выражение для ожидаемой значимости Z_{excl} может быть записано в следующем виде:

$$\begin{aligned}
Z_{excl} &= \sqrt{-2\left[b\ln\left(\frac{s+\frac{b-s+x}{2}}{b}\right) + \tau b\ln\left(\frac{b-s+x}{2b}\right) + b(1+\tau) - s - (1+\tau)\frac{b-s+x}{2}\right]} = \\
&= \sqrt{-2\left[b\ln\left(\frac{b+s+x}{2b}\right) + \tau b\ln\left(\frac{b-s+x}{2b}\right) - s + (1+\tau)(b - \frac{b-s+x}{2})\right]} = \\
&= \sqrt{-2\left[b\ln\left(\frac{b+s+x}{2b}\right) + \tau b\ln\left(\frac{b-s+x}{2b}\right) - s + (1+\tau)\frac{b+s-x}{2}\right]} = \\
&= \sqrt{2\left[s - b\ln\left(\frac{b+s+x}{2b}\right) - \tau b\ln\left(\frac{b-s+x}{2b}\right)\right] - (b+s-x)(1+\tau)}\,.
\end{aligned}$$

В очередной раз переписывая полученное, используя систематическую ошибку δ фона, можно получить окончательное выражение для ожидаемой значимости исключения:

$$Z_{excl} = \sqrt{2\left[s\ \ ---b\ln\left(\frac{b+s+x}{2b}\right) - \frac{1}{\delta^2}\ln\left(\frac{b-s+x}{2b}\right)\right] - (b+s-x)\left(1+\frac{1}{\delta^2 b}\right)}\quad , \tag{39}$$

где x задается выражением $x = \sqrt{(b+s)^2 - \frac{4\delta^2 b^2 s}{\delta^2 b+1}}$.

В случае, если фон известен точно, $\delta \to 0$ и $\lim_{\delta\to 0} x = b + s$. Тогда (39) переходит в:

$$Z_{excl} = \sqrt{2\left[s - b\ln\left(1+\frac{s}{b}\right)\right]}\,. \tag{40}$$

4. ОБОБЩЕННЫЕ ФОРМУЛЫ ДЛЯ ОЖИДАЕМОЙ ЗНАЧИМОСТИ ОТКРЫТИЯ И ИСКЛЮЧЕНИЯ СИГНАЛА

Приведенные в предыдущем разделе выводы можно обобщить на случай N сигнальных и M контрольных бинов [4]. Предполагается, что количествам сигнальных событий $s_1, s_2, \ldots, s_N$ в каждом из N сигнальных бинов соотвествует одна и та же сила сигнала μ, а также что контрольная область не пересекается с сигнальной и независима от нее. Для $i = 1, \ldots, N$ сигнальных бинов существуют матрицы $\boldsymbol{\tau}^i$; $i = 1, \ldots, N$, где $\boldsymbol{\tau}^i$ — матрица, сопоставляющая фоны из i-го сигнального бина с M контрольными бинами (например, τ^i_{jk} сопоставляет k-ый фон из i-го сигнального бина с j-ым контрольным бином). В таком случае функция правдоподобия:

$$L(\mu, \boldsymbol{B}) = \prod_{i=1}^{N} P\left(n_i \,\middle|\, \mu s_i + \sum_{j=1}^{M} b_{ji}\right) \times \\ \times \prod_{j=1}^{M} P\left(m_j \,\middle|\, \sum_{j'=1}^{M} \tau^{i'}_{jj'} b_{j'i'}\right), \quad (41)$$

где $P(n\,|\,\lambda)$ — вероятность пуассоновского распределения, $\boldsymbol{B}$ — матрица фонов в сигнальных бинах ($b_{ji} \equiv [\boldsymbol{B}]_{ji}$ соответствует j-му фону в i-ом сигнальном бине), а i' — любое целое число от 1 до N. Для определенности можно положить $i' = 1$.

Тогда, рассчитав оценки максимального правдоподобия $\hat{\mu}$ и $\hat{\boldsymbol{B}}$ и воспользовавшись выражением (24) и азимовским набором данных, можно получить выражение для ожидаемой значимости открытия сигнала:

$$Z_{disc} = \sqrt{-2\left(\sum_{i=1}^{N}\left\{n_i \ln\left(\frac{\sum_{j=1}^{M}\hat{\hat{b}}_{ji}}{n_i}\right) + n_i - \sum_{j=1}^{M}\hat{\hat{b}}_{ji}\right\} + \sum_{j=1}^{M}\left\{m_j \ln\left(\frac{[\boldsymbol{\tau}^1\cdot\hat{\hat{\boldsymbol{B}}}]_{j1}}{[\boldsymbol{\tau}^1\cdot\boldsymbol{B}]_{j1}}\right) + [\boldsymbol{\tau}^1\cdot(\boldsymbol{B}-\hat{\hat{\boldsymbol{B}}})]_{j1}\right\}\right)}, \quad (42)$$

где $\hat{\hat{\boldsymbol{B}}}$ задается выражением:

$$\sum_{i=1}^{N}\frac{\tau^1_{1\ell}}{\tau^i_{1\ell}}\cdot\left(\frac{n_i}{\sum_{j=1}^{M}\hat{\hat{b}}_{ji}} - 1\right) + \sum_{j=1}^{M}\tau^1_{j\ell}\cdot\left(\frac{m_j}{[\boldsymbol{\tau}^1\cdot\hat{\hat{\boldsymbol{B}}}]_{j1}} - 1\right) = 0\,; \quad \ell = 1, \ldots, M\,. \quad (43)$$

Аналогичным образом можно получить выражение для ожидаемой значимости исключения сигнала, воспользовавшись формулой (38):

$$Z_{excl} = \sqrt{-2\left(\sum_{i=1}^{N}\left\{n_i \ln\left(\frac{s_i + \sum_{j=1}^{M}\hat{\hat{b}}_{ji}}{n_i}\right) + n_i - s_i - \sum_{j=1}^{M}\hat{\hat{b}}_{ji}\right\}\right) + \overline{\sum_{j=1}^{M}\left\{m_j \ln\left(\frac{[\boldsymbol{\tau}^1\cdot\hat{\hat{\boldsymbol{B}}}]_{j1}}{[\boldsymbol{\tau}^1\cdot\boldsymbol{B}]_{j1}}\right) + [\boldsymbol{\tau}^1\cdot(\boldsymbol{B}-\hat{\hat{\boldsymbol{B}}})]_{j1}\right\}}}, \quad (44)$$

где $\hat{\hat{\boldsymbol{B}}}$ задается выражением:

$$\sum_{i=1}^{N}\frac{\tau^1_{1\ell}}{\tau^i_{1\ell}}\cdot\left(\frac{n_i}{s_i + \sum_{j=1}^{M}\hat{\hat{b}}_{ji}} - 1\right) + \sum_{j=1}^{M}\tau^1_{j\ell}\cdot\left(\frac{m_j}{[\boldsymbol{\tau}^1\cdot\hat{\hat{\boldsymbol{B}}}]_{j1}} - 1\right) = 0\,; \quad \ell = 1, \ldots, M\,. \quad (45)$$

5. МЕТОДОЛОГИЯ ПРИМЕНЕНИЯ ФОРМУЛ ДЛЯ АНАЛИЗА ЭКСПЕРИМЕНТА И ЧИСЛЕННАЯ ПРОВЕРКА РЕЗУЛЬТАТОВ

В данном разделе показана методология применения аналитических формул (28) и (40) для оценки статистической значимости открытия сигнала в эксперименте и постановки верхних ограничений на параметр модели «новой физики». Рассматриваются два простейших счетных эксперимента, каждый из которых содержит одну сигнальную область. В первом случае в сигнальной области имеется $n = 135$ событий, из которых подразумевается $b = 90$ фоновых и $s = 45$ сигнальных. Тогда, так как количество сигнальных событий значительно, возможно говорить об открытии сигнала на некотором уровне статистической значимости, которую необходимо оценить. Во втором случае имеется $n = 110$ событий в сигнальной области, $s = 10$ из которых — сигнальные и $b = 100$ — фоновые. В этом случае количество сигнальных событий относительно невелико, и следовательно можно достоверно установить только верхние ограничения на параметр силы сигнала μ.

5.1. Аналитический результат оценки статистической значимости и постановки верхних ограничений

В анализе первого эксперимента, пользуясь выведенной в предыдущих разделах формулой (28), можно получить значение статистической значимости открытия сигнала. Помимо самой формулы (28) полезно рассмотреть ее предельный случай, когда $b \gg s$:

$$Z_{disc} = \sqrt{2\left[(s+b)\ln\left(1+\frac{s}{b}\right)-s\right]} \approx \frac{s}{\sqrt{b}}. \quad (46)$$

Подставляя s = 50, b = 100 в (28) и (46) имеем:

$Z_{disc} = 4.41$ — значение, полученное из (28),
$Z_{approx} = 4.74$ — значение, полученное из (46).

Таким образом, получены оценки статистической значимости, с которой отвергается нулевая гипотеза ($s = 0$), то есть принимается альтернативная ($s = 45$) об открытии сигнала.

В случае второго эксперимента, как было сказано выше, можно лишь поставить верхние ограничения на параметр μ. В анализе такого типа необходимо выбрать уровень значимости. Общепринятым как в физике высоких энергий, так и в математической статистике в целом, считается уровень достоверности $\alpha = 95\%$, что соответсвует статистической значимости односторонних верхних ограничений $Z = 1.65$. Для постановки ограничений можно воспользоваться вариантом формулы (40), которая была получена в предыдущих разделах при условии, что предполагаемое значение параметра силы сигнала $\mu = 1$. Аналогично выводу, приведенному выше, полагая значение параметра μ неизвестным и требущим оценки, можно получить формулу:

$$Z_{excl} = med[Z_\mu|\mu] = \sqrt{2\left(\mu s - b\ln\left(1+\frac{\mu s}{b}\right)\right)}. \quad (47)$$

Подставляя в (47) значения $Z = 1.65$, $s = 10$ и $b = 100$ и решая уравнение относительно μ, имеем верхнее ограничение:

$\mu \leq 1.7$ (на уровне значимости 95%).

Следует отметить, что аналитические формулы (28), (46) и (47) дают лишь оценку ожидаемых значений. Тем не менее, эти формулы весьма удобны в применении. Задача стоит в том, чтобы проверить точность этих оценок и убедиться в их достоверности.

5.2. Численное моделирование экспериментов с помощью метода Монте–Карло

Можно получить значения статистической значимости отрытия сигнала и верхних ограничений с помощью численного Монте–Карло (МС) моделирования экспериментов. Точные значения, полученые в результате такого моделирования, затем можно сравнить с аналитическими оценками. Для проведения МС моделирования используется пакет ROOT [5] и его класс RooStats [6].

Согласно нулевой гипотезе H_0, распределение данных в сигнальной области будет соответствовать распределению Пуассона с параметром $\mu = 0$. Тогда альтернативная H_1 гипотеза — $\mu = 1$. Основываясь на определении статистики q_0 (11), которая являетя функцией отношения максимумов правдоподобия (3), необходимо рассчитать значения параметров, максимизирующих функцию правдоподобия (19). Рассчитанные с помощью формул (20) — (22) значения параметров:

$$\hat{\mu} = 1, \quad \hat{b} = 90, \quad \hat{\hat{b}} = 90.$$

Распределение статистики q_0 моделируется для обеих гипотез путем проведения случайных псевдоэкспериментов методом Монте–Карло. Выбирается количество этих псевдоэкспериментов равное двумстам тысячам с целью уменьшения ошибки. После того, как распределение статистики смоделировано, проводится статистический тест, описанный выше. С помощью численного интегрирования находится p-значение, которое затем переводится в термины статистической значимости. В результате такого МС моделирования полученное точное значение статистической значимости открытия сигнала в случае первого эксперимента:

$$Z_0 = 4.26 \pm 0.03.$$

На рис.1 представлен график полученных распределений.

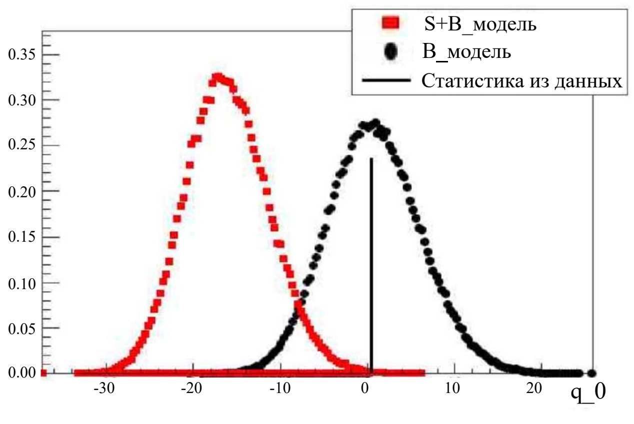

Рис. 1. Монте–Карло смоделированные распределения статистики q_0 для нулевой (черные кружки) и альтернативной (красные квадраты) гипотез, черная линия — значение статистики q_0, полученное из данных

В случае второго эксперимента метод проведения моделирования распределения статистики q_μ аналогичен. Имеется $s = 10$ сигнальных и $b = 100$ фоновых событий в сигнальной области. Теперь в качестве нулевой гипотезы берется сигнальная ($\mu = 1$), а в качестве альтернативной — фоновая ($\mu = 0$). Для постановки верхних ограничений, как описано выше, выбирается уровень значимости $\alpha = 95\%$. Значения максимизирующих параметров функций правдоподобия:

$$\hat{\mu} = 1, \quad \hat{b} = 100, \quad \hat{\hat{b}} = 125.$$

Достаточно смоделировать распределение статистики q_μ только для сигнальной гипотезы. Далее необходимо провести численное интегрирование распределения до точки, в которой интеграл равен 0.95 (найти квантиль уровня 95%). Значение статистики в этой точке подставляется в (29), откуда затем находится значение μ. Именно это значение и будет являтся верхним ограничением на параметр силы сигнала на уровне значимости $\alpha = 95\%$. Результатом такого MC моделирования является точное значение верхнего предела:

$$\mu_{mc} \leq 1.66 \pm 0.03 \text{ (на уровне значимости 95\%)}.$$

На рис. 2 представлен график распределения статистики q_μ для сигнальной гипотезы.

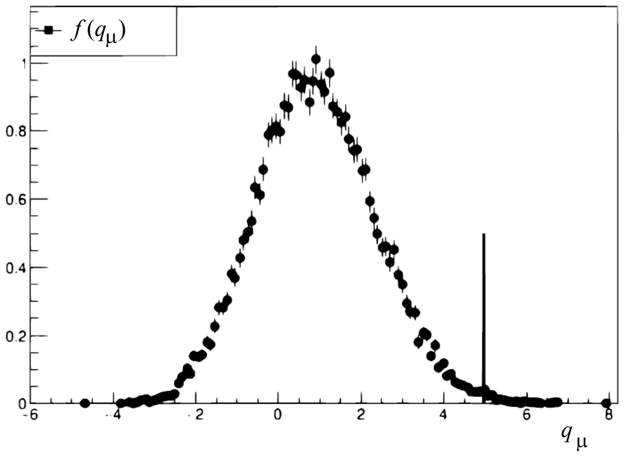

Рис. 2. Монте–Карло смоделированное распределения статистики q_μ для сигнальной гипотезы (черные точки), черная линия - 95% квантиль распределения

5.3. Сравнение оценок, полученных с помощью аналитических формул, с результатами численного моделирования

Для сравнения значений, полученных с помощью аналитических формул (28), (46) и (47) необходимо повторить процедуру, описанную в предыдущих секциях несколько раз для разлчиных значений количества событий. Такой подход позволит определить качество оценок, полученных на основе этих формул, и границы их применимости.

5.3.1. Сравнение для статистическая значимости открытия сигнала

Для определения границ применимости формул (28) и (46) проводится тринадцать счетных экспериментов для различного числа фоновых событий b в диапазоне от b=2 до b = 90. В каждом эксперименте отношение числа сигнальных событий к числу фоновых равно $\frac{s}{b} = \frac{1}{2}$. Результатом такой процедуры является сравнение значений статистической значимости открытия сигнала, полученной с помощью аналитических формул и с помощью точного Монте–Карло моделирования (рис. 3)

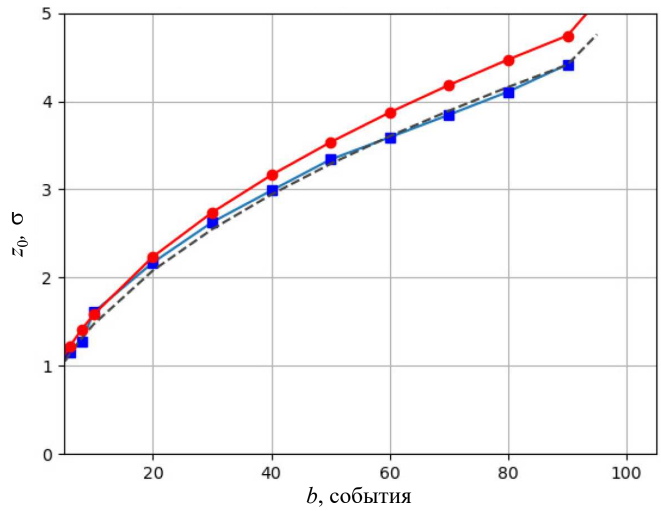

Рис. 3. Сравнение значений статистической значимости, полученной с помощью формул (28) (серый пунктир) и (46) (красная сплошная линия и кружки) и MC моделирования (сплошная синия линия и квадраты) см. текст

На графике видно хорошее согласие между аналитическими значениями статистической значимости, полученными с помощью формулы (28) и значениями, полученными с помощью MC моделирования экспериментов. Можно сделать вывод о достаточно широкой области применимости данной формулы, начиная со значений количества событий порядка единицы. С ростом количества событий наблюдаются некоторые отклонения значений статистической значимости открытия сигнала, полученных с помощью асимптотической формулы (46), от точных значений MC моделирования. Такое различие может быть объяснено тем, что условие s < < b не выполняется при $\frac{s}{b} = \frac{1}{2}$, ошибка становится больше при увеличении количества событий.

5.3.2. Сравнение для постановки значений верхних ограничений

Аналогичная процедура проведения MC моделирования нескольких экспериментов с целью постановки верхних ограничений на параметр силы сигнала μ и дальнейшее сравнение полученных значений со значениями, полученными с помощью формулы (47), позволит исследовать ее границы применимости. Выбираеются эксперименты с количеством фоновых событий в диапазоне от b=10 до b=130, отношение числа сигнальных событий к числу фоновых событий всюду равно $\frac{s}{b} = \frac{1}{10}$. В результате проведения такой процедуры получено сравнение аналитических значений верхних ограничений и значений из MC моделирования (рис. 4).

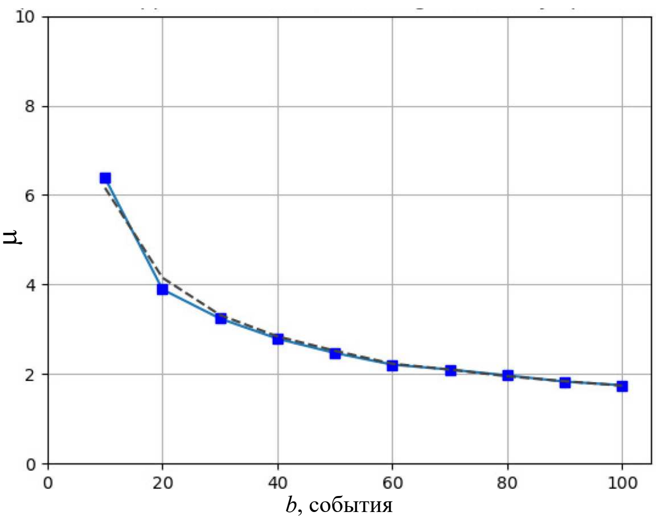

Рис. 4. Сравнение значений верхних ограничений на параметр силы сигнала μ, полученных с помощью формулы (47) (пунктирная линия) и с помощью численного МК моделирования экспериментов (сплошная синяя линия и квадраты) см. текст

На представленном графике наблюдается очень хорошее согласование значений верхних пределов, полученных с использованием аналитической формулы (47) и с помощью MC моделирования. Можно сделать вывод, что данная формула также имеет достаточно широкую границу применимости и дает весьма неплохие оценки верхних пределов на параметр силы сигнала.

ЗАКЛЮЧЕНИЕ

В представленной работе проведено уточнение, дальнейшее развитие и численная проверка асимптотических методов статистического анализа коллайдерных экспериментов, предложенных ранее в статье [1]. Авторами представляемой работы выбраны два наиболее актуальных сценария статистического анализа. Первый сценарий — это оценка статистической значимости измерения сигнала нового процесса на коллайдере, описываемого в рамках некоторой модели «новой физики». Второй сценарий предполагает малую статистическую значимость выделения такого сигнала «новой физики» и предоставляет возможность установить верхние ограничения на параметры, характеризующие возможный вклад такой «новой физики». В работе были выведены необходимые асимптотические формулы, некоторые из них не были представлены ранее в опубликованных статьях. В статистических моделях введена возможность учета систематических неопределенностей предсказания фоновых процессов. Проведено обобщение статистической модели для случая анализа формы распределений событий, представленных в виде гистограмм, что позволяет напрямую использовать полученные формулы для анализа дискриминантов многомерных статистических методов, таких как нейронные сети. В заключении проведено численное моделирование методом Монте–Карло и сравнение результатов асимптотических формул и точного моделирования. Продемонстрировано хорошее согласие оценки статистической значимости, полученной в точном моделировании и в асимптотических формулах.

В качестве основного результата работы представлен явный вывод асимптотических формул (27) и (39) для оценки статистической значимости открытия и исключения сигнала соответственно. Кроме того, предложено обобщение этих формул на случай большего числа бинов в сигнальной и контрольной областях: представленная в работе [4] формула (42) для оценки статистической значимости открытия и не опубликованная ранее аналогичная формула (44) для оценки статистической значимости исключения сигнала. Также показана методология применения формул (28), (46) и (47) на примерах простейших счетных экспериментов и проведена их численная проверка с помощью Монте–Карло моделирования.

Благодарности

Исследование выполнено в рамках научной программы Национального центра физики и математики, направление № 5.2 «Физика частиц и космология»

[1] *Cowan G., Cranmer K., Gross E., Vitells O.* // Eur. Phys. J. C. **71**. (2011). [arXiv:1007.1727].
[2] *Wilks S.S.* // Annals Math. Statist. **9**, N 1. 60 (1938).
[3] *Wald A.* // Transactions of the American Mathematical Society. **54**, N 3. 426 (1943).
[4] *Basso M.J.* // J. Phys. G. **49**. (2022). [arXiv:2102.04275]
[5] *Brun R., Rademakers F.* // ROOT: An object oriented data analysis framework, Nucl. Instrum. Meth. A389, 81 (1997).
[6] *Moneta L., Belasco K., Cranmer K.* et. al. // The RooStats Project, arXiv:1009.1003

Asymptotic formulae for estimating statistical significance of experiments on LHC

D. E. Gorin[a], **O. S. Vasilevskii**[b], **L. V. Dudko**[c], **E. E. Abasov**[d]

Department of Nuclear Physics, Faculty of Physics, Lomonosov Moscow State University. Moscow 119991, Russia
E-mail: [a]gorin.de20@physics.msu.ru, [b]vasilevskii.os20@physics.msu.ru, [c]lev.dudko@cern.ch, [d]emil@abasov.ru

This paper elaborates and expands on earlier presented likelihood-based asymptotic methods of statistical analysis in high-energy physics. An explicit derivation of earlier presented asymptotic formulae for estimating the statistical significance of the discovery of new phenomena and setting upper limits on parameters of «new physics» models is given using the results of Wilks and Wald. The formulae obtained in combination with the use of the so-called «Asimov dataset» are of great interest due to their simplicity and the possibility of accounting for systematic uncertainties. The methodology of using these formulae is also shown in this paper and their accuracy is being verified by Monte Carlo modeling.

PACS: 02.70.Rr, 02.50.Cw, 02.50.Ng, 02.30.Mv, 02.70.Uu

Keywords: asymptotic formulae, statistical analysis, statistical significance, experimental sensitivity, likelihood, Asimov data set, Monte Carlo modeling.

Received 26 December 2023.

Сведения об авторах

1. Горин Даниил Эдуардович — студент; e-mail: gorin.de20@physics.msu.ru.
2. Василевский Олег Сергеевич — студент; e-mail: vasilevskii.os20@physics.msu.ru.
3. Дудко Лев Владимирович — доктор физ.-мат. наук, зав. лаб.; e-mail: lev.dudko@cern.ch.
4. Абасов Эмиль Эльданизович — студент; e-mail: emil@abasov.ru.