

ПОСТРОЕНИЕ ДОВЕРИТЕЛЬНЫХ ИНТЕРВАЛОВ С УЧЕТОМ АПРИОРНОЙ ИНФОРМАЦИИ

А. В. Лохов *, *Ф. В. Ткачев*

Институт ядерных исследований РАН, Москва

ВВЕДЕНИЕ	628
НЕЙМАНОВСКОЕ ПОСТРОЕНИЕ	631
Непрерывный случай	631
Дискретный случай	633
ПОСТРОЕНИЕ ССГВ	636
ПОСТРОЕНИЕ ФЕЛЬДМАНА И КАЗИНСА	637
КОНСТРУКЦИЯ МАНДЕЛЬКЕРНА И ШУЛЬЦА	638
МЕТОД ПРЕДЕЛА ЧУВСТВИТЕЛЬНОСТИ ДЛЯ ПАРАМЕТРА НЕПРЕРЫВНОГО РАСПРЕДЕЛЕНИЯ	639
Горизонтальные деформации	641
Доверительные интервалы в методе предела чувствительности	643
МЕТОД ПРЕДЕЛА ЧУВСТВИТЕЛЬНОСТИ В СЛУЧАЕ ДИСКРЕТНЫХ РАСПРЕДЕЛЕНИЙ	644
НАИЛУЧШИЙ ВЕРХНИЙ ПРЕДЕЛ (НЕПРЕРЫВНЫЙ И ДИСКРЕТНЫЙ СЛУЧАИ) В МЕТОДЕ ПРЕДЕЛА ЧУВСТВИТЕЛЬНОСТИ	653
Непрерывный случай	653
Дискретный случай. Метод предела чувствительности для параметра пуассоновского распределения с фоном: случай несимметричной системы доверительных интервалов	656
ОЦЕНКИ МАССЫ ЭЛЕКТРОННОГО АНТИНЕЙТРИНО В ЭКСПЕРИМЕНТАХ В МАЙНЦЕ И ТРОИЦКЕ	658

*E-mail: lokhov.alex@gmail.com

ВЫВОДЫ	661
СПИСОК ЛИТЕРАТУРЫ	663

ПОСТРОЕНИЕ ДОВЕРИТЕЛЬНЫХ ИНТЕРВАЛОВ С УЧЕТОМ АПРИОРНОЙ ИНФОРМАЦИИ

А. В. Лохов*, Ф. В. Ткачев

Институт ядерных исследований РАН, Москва

Дается обзор методов решения фундаментальной задачи учета априорной информации в виде одностороннего ограничения на оцениваемый параметр при построении доверительных интервалов. Подробно рассматривается так называемый метод предела чувствительности, дающий, как показывается, физически корректное решение задачи. Строятся решения для ситуаций с непрерывным распределением в случае неотрицательного оцениваемого параметра, а также для дискретных распределений, в частности для пуассоновского процесса с фоном. Для этих же двух случаев построен наилучший верхний предел, учитывающий наличие априорной информации. Для важного в физических приложениях случая пуассоновского распределения с неизвестным параметром распределения и известным фоном приведена таблица доверительных интервалов для параметра, а также представлена полная программа расчетов для произвольных доверительных уровней и значений фона (программа находится в открытом доступе).

We review the methods of constructing confidence intervals that account for *a priori* information about one-sided constraints on the parameter being estimated. We show that the so-called method of sensitivity limit yields a correct solution of the problem. Derived are the solutions for the cases of a continuous distribution with non-negative estimated parameter and a discrete distribution, specifically a Poisson process with background. For both cases, the best upper limit is constructed that accounts for the *a priori* information. A table is provided with the confidence intervals for the parameter of Poisson distribution that correctly accounts for the information on the known value of the background along with the software for calculating the confidence intervals for any confidence levels and magnitudes of the background (the software is freely available for download *via* Internet).

PACS: 29.85.Fj

ВВЕДЕНИЕ

Построение неймановских доверительных интервалов [1] для экспериментально измеряемых параметров — одна из важнейших задач обработки данных. При этом часто встречается случай, когда имеется априорная инфор-

*E-mail: lokhov.alex@gmail.com

мация об оцениваемых параметрах, и важно корректно и полностью учесть эту информацию при построении интервалов.

Пример априорной информации — ограниченность области значений параметра, и проблема возникает в случае, когда экспериментальная оценка — являясь случайной величиной — выпадает из этой области. Например, в эксперименте «Троицк-ню-масс» по измерению массы нейтрино в β -распаде трития [2] параметр m_ν^2 заведомо больше либо равен нулю, однако полученное формальным фитированием значение m_ν^2 оказалось меньше нуля.

Другой важный момент, когда нужно учитывать априорную информацию, — построение доверительных интервалов в случае пуассоновского процесса при наличии пуассоновского же фона. Это обычная ситуация при изучении редких событий (в экспериментах по поиску двойного безнейтринного β -распада [3], по осцилляциям нейтрино, например T2K, MINOS [4] и др.).

Было предложено несколько вариантов решения этой задачи. Решения можно разбить на две группы по способу использования свободы, заложенной в неймановской процедуре построения доверительных интервалов.

В первой группе решений априорная информация используется на стадии построения доверительных областей. К этому типу решений относится построение Фельдмана и Казинса [6], а также способ ограничения интервалов с помощью функции мощности, предложенный Кауэном и др. (так называемый метод CCGV [7]). Однако интервалы, построенные, например, по рецепту Фельдмана и Казинса (способ, имеющий высокую цитируемость), не позволяют произвести объективное сравнение результатов разных экспериментов (рецепт может дать лучший интервал для эксперимента с заведомо худшей чувствительностью и систематикой).

Другая группа решений предполагает использование априорной информации о параметрах при выборе самой оценки (estimator). К этой группе, по сути, принадлежит конструкция доверительных интервалов, указанная в работе [8], где, однако, рассматривается только случай оценивания по методу максимального правдоподобия, а рассуждения не вполне прозрачны (и местами искусственны), и даже не приводится явная формула для оценки. По этим причинам не видно, что окончательный рецепт может быть использован с любыми методами оценивания. Надо полагать, что именно поэтому конструкция доверительных интервалов [8] — физически, как выясняется, корректная — оставалась, судя по истории цитирования, практически неизвестной.

Полное решение, не зависящее от способа оценивания (будь то метод моментов, метод наименьших квадратов или метод максимального правдоподобия), с явной формулой для оценки было найдено в работе [9], где, в отличие от работы [8], дана ясная графическая и аналитическая интерпретация для всей конструкции.

Оказывается, что эта конструкция для нефизических, т. е. выпадающих за априорную границу значений оценки совпадает с эмпирическим рецептом так называемого предела чувствительности (использованным, например, среди прочих методов для представления результатов эксперимента по измерению массы нейтрино в Майнце [10]), поэтому мы будем называть рассматриваемую конструкцию методом предела чувствительности. Чтобы избежать путаницы, необходимо подчеркнуть, что предел чувствительности — это величина, характеризующая данный эксперимент и соответствующие неопределенности, она может быть вычислена до начала измерений. Метод предела чувствительности, в свою очередь, дает систему доверительных интервалов (доверительный пояс), построенную в соответствии с неймановской процедурой [1].

Возможно, главным достоинством метода предела чувствительности является тот факт, что он позволяет напрямую объективно сравнивать результаты различных экспериментов, не требуя повторной обработки. Именно так, например, новые результаты эксперимента «Троицк-ню-масс» [2] можно сравнивать (см. ниже разд. 8) как с упомянутым старым измерением в Майнце, так и с результатом старой обработки данных «Троицк-ню-масс» [11].

В этом отношении метод качественно превосходит построение Фельдмана и Казинса [6], где в нефизической области доверительный интервал зависит от экспериментального значения оценки параметра и даже естественно уменьшается по мере удаления от физической границы в нефизическую область. Фактически рецепт Фельдмана и Казинса не дает результатов, которые можно было бы сравнивать.

Интересно, что применение второго подхода (т. е. учет априорной информации еще при построении оценки) позволяет также указывать корректные и оптимальные односторонние (сверху или снизу) ограничения на параметры [12].

Для распространения решения на случай дискретных распределений также проще всего (и понятней) использовать процедуры работ [9, 12].

Сначала в разд. 1 мы напоминаем неймановскую конструкцию, формулируя ее в виде, удобном для вывода метода предела чувствительности. При этом особо рассматривается дискретный случай, где — в силу самой дискретности — необходимо отказаться от равенств, описывающих вероятностное содержание доверительных интервалов, в пользу неравенств (как это было и в рецепте Фельдмана и Казинса). Обсуждается построение интервалов симметричных, несимметричных, а также способ построения Стерна, Кроу и Гарднера [5].

В разд. 2–4 дается обзор конструкций Кауэна и др., Фельдмана и Казинса и Манделькерна и Шульца соответственно.

В разд. 5, согласно [9], проводится построение метода предела чувствительности для непрерывного случая, когда априорная информация об измеряемом параметре имеет вид неравенства $\theta \geq 0$. В разд. 6 рассматривается

дискретный случай пуассоновского процесса с неизвестным μ , но известным значением пуассоновского фона b . Наконец, в разд. 7 по методу предела чувствительности строятся наилучшие верхние пределы для непрерывного и дискретного случая. В разд. 8 на основе результатов экспериментов по измерению массы нейтрино в β -распаде трития проиллюстрирована сравнимость доверительных интервалов, построенных по методу предела чувствительности. Выводы суммированы в разд. 9.

Подчеркнем, что в данной работе рассматривается только неймановская процедура построения доверительных интервалов. Такие интервалы имеют ясную интерпретацию в рамках частотного подхода к интервальному оцениванию. Другой используемый подход — байесовский — хотя и может быть с натяжками проинтерпретирован в терминах статистических ансамблей [13], но применяемая экспериментаторами процедура построения байесовских интервалов использует неизвестную априорную функцию плотности вероятности для параметра, что нарушает условия применимости теоремы Байеса в узком смысле и затрудняет интерпретацию результатов. Мы не рассматриваем в настоящей работе и построение доверительных интервалов методом CLs [14], поскольку он не имеет ясной интерпретации в рамках ни частотного, ни байесовского подхода, хотя и широко используется, например, в представлении результатов поиска бозона Хиггса [15].

1. НЕЙМАНОВСКОЕ ПОСТРОЕНИЕ

1.1. Непрерывный случай. Чтобы зафиксировать обозначения, которые понадобятся в дальнейшем, кратко опишем стандартную процедуру построения доверительных интервалов по Нейману [1].

Пусть нужно оценить неизвестный параметр θ по набору экспериментальных данных X . Предположим, что для этого применяется какой-нибудь обычный метод (например, фундаментальный метод моментов [16, 17]), дающий некоторую оценку $\hat{\theta} = \hat{\theta}(X)$ без учета априорного ограничения. Оценка $\hat{\theta}$ является случайной величиной, плотность вероятности которой $d_{\theta}(\hat{\theta})$ зависит от θ как от параметра. Предполагается, что плотность вероятности — это известная несингулярная функция, как того требует стандартная неймановская процедура построения доверительных интервалов [1]. Плотность $d_{\theta}(\hat{\theta})$ содержит в себе всю информацию об измерении параметра θ в данном эксперименте, включая метод оценивания.

Пусть α, α' — малые неотрицательные величины. Определим $L_{\alpha}(\theta)$ и $U_{\alpha'}(\theta)$ как

$$P(-\infty < \hat{\theta} < L_{\alpha}(\theta)) = \alpha, \quad P(U_{\alpha'}(\theta) < \hat{\theta} < +\infty) = \alpha'. \quad (1)$$

Вероятность получить величину оценки меньше $L_{\alpha}(\theta)$ равна α , а больше $U_{\alpha'}(\theta)$ — α' .

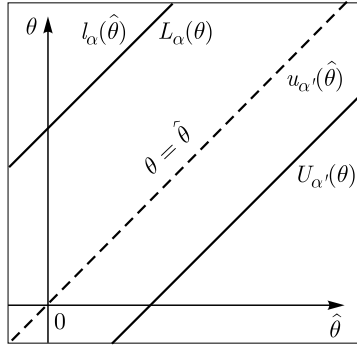


Рис. 1. Показаны функции $\theta = l_\alpha(\hat{\theta})$ и $\theta = u_{\alpha'}(\hat{\theta})$ (или, с другой точки зрения, $\hat{\theta} = L_\alpha(\theta)$ и $\hat{\theta} = U_{\alpha'}(\theta)$). В общем случае диагональ $\theta = \hat{\theta}$ не лежит между двумя кривыми (на последующих иллюстрациях диагональ отображаться не будет). Две сплошные линии будут использоваться в дальнейшем для случая, когда они составляют стандартную симметричную систему доверительных интервалов с уровнем доверия $\beta = 1 - 2\alpha$. Чем меньше величина β , тем уже доверительные интервалы

Таким образом, $L_\alpha(\theta)$ соответствует определенной в разд. 9.1.1 [16] величине Z_α .

Если еще предположить, что $L_\alpha(\theta)$ и $U_{\alpha'}(\theta)$ суть обратимые функции θ , то (1) можно переписать в виде

$$P(l_\alpha(\hat{\theta}) < \theta) = \alpha, \quad P(\theta < u_{\alpha'}(\hat{\theta})) = \alpha', \quad (2)$$

где $u_\alpha = U_\alpha^{-1}$, $l_\alpha = L_\alpha^{-1}$. Из (2) следует, что вероятность получить случайную величину $l_\alpha(\hat{\theta})$ ($u_{\alpha'}(\hat{\theta})$), меньшую (большую), чем неизвестное истинное значение θ , равна α (α') (см. рис. 1).

Можно переписать (1) в виде

$$P(L_\alpha(\theta) < \hat{\theta} < U_{\alpha'}(\theta)) = 1 - \alpha - \alpha' \equiv \beta. \quad (3)$$

Тогда из эквивалентного выражения

$$P(u_{\alpha'}(\hat{\theta}) < \theta < l_\alpha(\hat{\theta})) = \beta \quad (4)$$

следует, что с вероятностью β (доверительный уровень, например, 90%-й) случайный интервал $[u_{\alpha'}(\hat{\theta}), l_\alpha(\hat{\theta})]$ накрывает неизвестное значение θ .

Выбирая $\alpha = \alpha' = (1 - \beta)/2$, получим стандартный симметричный доверительный пояс (систему доверительных интервалов).

Описанное построение содержит следующую свободу. Зафиксируем уровень доверия β (например, $\beta = 90\%$, далее везде полагаем β фиксированным). Выберем две функции L, U , удовлетворяющие условию

$$P(L(\theta) < \hat{\theta} < U(\theta)) = \beta. \quad (5)$$

Если эти функции также монотонны, то существуют обратные им $u = U^{-1}$, $l = L^{-1}$. Тогда из эквивалентного выражения

$$P(u(\hat{\theta}) < \theta < l(\hat{\theta})) = \beta \quad (6)$$

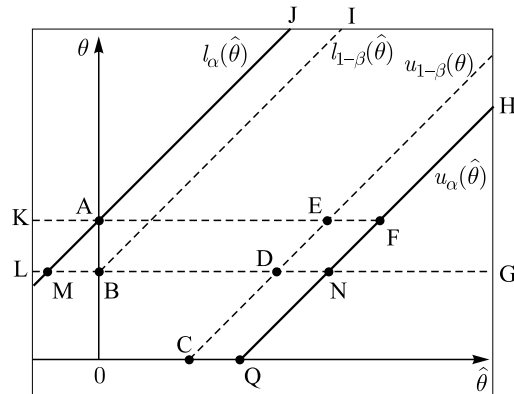


Рис. 2. Пары сплошных и штриховых наклонных линий, ограничивающие симметричные доверительные пояса для доверительных уровней $\beta = 1 - 2\alpha$ и $\tilde{\beta} = 1 - 2(1 - \beta) = 1 - 4\alpha$; см. рис. 1. На рисунке указаны функции, соответствующие линиям. Точка А — пересечение вертикальной оси с линией $\theta = l_\alpha(\hat{\theta})$. Точка А задает горизонтальную линию KF с другими точками пересечений. Точки С и Q — пересечения линий $\theta = u_{1-\beta}(\hat{\theta})$ и $\theta = u_\alpha(\hat{\theta})$ с горизонтальными осями

снова следует, что случайный интервал $[u(\hat{\theta}), l(\hat{\theta})]$ накрывает неизвестное значение θ с вероятностью β .

Заметим, что линия $\theta = u(\hat{\theta})$ расположена ниже $\theta = u_{1-\beta}(\hat{\theta})$ (т. е. $u(\hat{\theta}) \leq u_{1-\beta}(\hat{\theta})$). Аналогично, линия $\theta = l(\hat{\theta})$ ограничена снизу. Каждая пара таких линий задает допустимый доверительный пояс для доверительного уровня β .

На рис. 2, в дополнение к симметричному доверительному поясу с уровнем доверия β , введен также более узкий пояс для меньшего уровня доверия $\tilde{\beta} = 1 - 2(1 - \beta) = 1 - 4\alpha < \beta$ (наклонные штриховые линии). Буквенные обозначения различных точек пересечения и горизонтальных линий введены для удобства: одинаково обозначенные на разных рисунках точки совпадают с точками на данном рисунке.

Обозначим ординату точки А (пересечение вертикальной оси и прямой KF) θ_A :

$$\theta_A = l_\alpha(0). \tag{7}$$

Числа $\theta_C < \theta_E < \theta_F$ суть ординаты точек С, Е, F:

$$\theta_C = U_{1-\beta}(0), \quad \theta_E = U_{1-\beta}(\theta_A), \quad \theta_F = U_\alpha(\theta_A). \tag{8}$$

1.2. Дискретный случай. Будем рассматривать ситуацию, когда в эксперименте детектируется число событий n . Число зарегистрированных событий

может быть распределено, например, по Пуассону:

$$P_{\mu}(n) = \frac{\mu^n}{n!} e^{-\mu}. \quad (9)$$

Здесь μ — параметр пуассоновского распределения, т. е. среднее число событий. (Дальнейшие рассуждения не зависят от конкретного вида распределения.)

Дискретность распределения приводит к некоторым изменениям в построении доверительных интервалов по сравнению с непрерывным случаем.

Для построения доверительных интервалов сначала выберем, как обычно, доверительный уровень α (например, 95 %).

Для каждого значения μ можно указать значения n , удовлетворяющие условию

$$P_{\mu}(n \in [n_1(\mu), n_2(\mu)]) \geq \alpha. \quad (10)$$

В дискретном случае невозможно обеспечить здесь равенство (как в формулах (5) и (6) для непрерывного случая), что и составляет специфику дискретного случая. (Необходимость перехода в дискретном случае к неравенствам обсуждалась в [6].) Поэтому для оцениваемого параметра μ можно только искать интервалы, которые будут содержать неизвестное истинное значение μ_0 в доле экспериментов, не меньшей α :

$$P(\mu_0 \in [\mu_1, \mu_2]) \geq \alpha. \quad (11)$$

Рассмотрим сначала односторонний интервал:

$$P(n \leq n_{\alpha}(\mu)) \geq \alpha. \quad (12)$$

Для определения доверительного интервала для μ можно преобразовать выражение в скобках:

$$P(\mu \geq \mu^*(n)) \geq \alpha. \quad (13)$$

Неравенство (13) означает, что в доле $\geq \alpha$ экспериментов (измерений числа n) будет получено значение n такое, что неизвестное истинное значение параметра удовлетворяет условию $\mu \geq \mu^*(n)$.

В случае дискретного распределения для величины n функция $\mu^*(n)$ определена, на первый взгляд, неоднозначно (каждому значению n соответствует набор значений μ^* , при которых выполнено неравенство (13)). Ниже мы покажем, что условие консервативности интервала [6] (требование выполнения (13) для любых фиксированных μ) определяет единственное значение μ^* для каждого n . Это значение μ^* и будет границей доверительного интервала.

Для нахождения $\mu^*(n)$ введем следующее обозначение: пусть μ_n — значение параметра, при котором выполняется точное равенство

$P(n \leq n_\alpha(\mu_n) - 1) = \alpha$ (при этом $P(n \leq n_\alpha(\mu_n + \varepsilon) - 1) < \alpha$ и $P(n \leq n_\alpha(\mu_n + \varepsilon)) > \alpha$), где ε — сколь угодно малое положительное число.

Покажем, что если выбрать в качестве $\mu^*(n)$ величину $\mu^*(n) = \mu_n + \varepsilon$, то условие консервативности не будет выполнено.

При таком выборе $\mu^*(n)$, если истинное значение μ лежит в интервале $(\mu_n, \mu_n + \varepsilon)$, условие (13) будет нарушаться. Действительно, в таком случае μ окажется больше либо равной случайной величине $\mu^*(n) = \mu_n + \varepsilon$ только в тех случаях, когда при измерении будут получены значения $n = n_\alpha(\mu) - 1$, $n = n_\alpha(\mu) - 2$ и т. д. Вероятность получить такие значения n равна $P(n \leq n_\alpha(\mu) - 1) < \alpha$ (согласно выбору величины μ_n). Следовательно, интервал $[\mu_n + \varepsilon, +\infty)$ по определению не является доверительным интервалом.

При выборе в качестве $\mu^*(n)$ случайной величины μ_n условие $\mu \geq \mu_n$ будет выполнено в доле измерений $\geq \alpha$, что удовлетворяет определению доверительного интервала для параметра μ .

Таким образом, односторонний доверительный интервал в случае дискретного распределения задается формулой $\mu \geq \mu_n$, где величина μ_n определяется условием $P(n \leq n_\alpha(\mu_n) - 1) = \alpha$.

Аналогично можно рассмотреть интервал $P(n \geq n_\alpha(\mu)) \geq \alpha$. Соответствующая верхняя граница доверительного интервала для μ будет определяться формулой $\mu \leq \mu'_n$. Здесь μ'_n задается условием $P(n \geq n_\alpha(\mu'_n) + 1) = \alpha$.

Дискретность распределения вносит дополнительную неопределенность в построение двусторонних доверительных интервалов. Так, например, можно задать двусторонний доверительный интервал, объединяя верхнюю и нижнюю границы односторонних интервалов:

$$P(n_1(\mu) \leq n \leq n_2(\mu)) \geq \alpha \Rightarrow P(\mu_1(n) \leq \mu \leq \mu_2(n)) \geq \alpha. \quad (14)$$

В качестве μ_1 и μ_2 выбираются соответствующие значения μ_n и μ'_n . Значения этих величин для каждого n , очевидно, зависят от доверительного уровня α .

Такой выбор приводит к приблизительно симметричному доверительному интервалу. Приблизительность состоит в том, что вероятностное содержание областей $n < n_1(\mu)$ и $n_2(\mu) < n$ может не быть одинаковым.

Для указания интервала можно также руководствоваться физическими соображениями. Так, если важно получить ограничение на параметр сверху, то можно потребовать выполнения условия $P(n < n_1(\mu)) < (1 - \alpha)/2$. В таком случае вероятностное содержание области выше верхней границы будет $< (1 - \alpha)/2$. Построим нижнюю границу доверительного интервала, используя значения $n_1(\mu)$ как левую границу в условии $P(n_1(\mu) \leq n \leq n_2(\mu)) \geq \alpha$. В этом случае может оказаться, что условие $P(n > n_2(\mu)) < (1 - \alpha)/2$ не выполнено. Нижняя граница доверительного интервала при таком построении может не совпадать с границей одностороннего интервала, полученной

из условия (12). В то же время такой интервал будет менее перекрывающим, т. е. вероятностное содержание соответствующей доверительной области будет ближе к заданному значению α (как справедливо указано в [6], это качество — перекрывание — является нежелательным для доверительных интервалов).

Аналогично можно указать интервал, если важно более строгое ограничение на измеряемый параметр снизу.

Еще один способ построения доверительных интервалов, использующий специфику дискретного случая, выражающуюся в необходимости перехода к неравенствам (что дает дополнительную степень свободы для всей конструкции), предложен Стерном, Кроу и Гарднером [5]. Идея способа заключена в конструировании доверительной области добавлением в нее точек в порядке убывания соответствующей вероятности (в отличие от построения симметричной системы интервалов). Впервые способ был применен к биномиальному (дискретному) распределению, хотя возможно его распространение и на непрерывные распределения. Такое построение приводит в общем случае к несимметричным интервалам, но соответствующие интервалы имеют меньшую длину. Как показал Кроу, такой способ дает доверительную полосу с наименьшей общей площадью.

Таким образом, для каждого измерения величины n можно указать доверительный интервал для неизвестного параметра μ в форме

$$\mu \in [\mu_{n\alpha}, \mu'_{n\alpha}]$$

несколькими способами. Выбор конкретного способа задания доверительного интервала, как обычно в задачах обработки данных, остается за экспериментатором.

2. ПОСТРОЕНИЕ CCGV

Перейдем теперь к основным, известным из литературы, способам учета априорной информации в процедуре построения доверительных интервалов. Рассмотрим вначале попытки учета априорной информации при построении доверительной области (решения первой группы).

Кауэн и др. предложили метод, получивший название ограничения интервалов с помощью функции мощности [7]. Этот метод состоит в следующем. Выбирается некий статистический критерий со статистикой $q_\mu = q_\mu(x)$ (для удобства выбирается статистика, увеличивающаяся при увеличении несогласия между экспериментальными данными x и величиной параметра μ). Для этого критерия строится функция мощности, которая и является ключевым элементом этого построения:

$$M_{\mu'}(\mu) = P(q_\mu(x) > q_{\mu, \text{crit}} | \mu').$$

Рассматриваются две гипотезы: $\mu' = 0$ (сигнал отсутствует) и $\mu > 0$ (существует ненулевой сигнал) и соответствующая функция мощности $M_0(\mu)$. Выбирается некоторое пороговое значение функции мощности M_{\min} , и область значений μ разбивается на две: если для некоторого значения μ значение $M_0(\mu)$ оказывается ниже порогового, то считается, что чувствительность к этому параметру слишком мала, и такие значения μ не могут быть проверены. Таким образом, μ не включается в доверительный интервал для некоторого конкретного набора экспериментальных данных, если 1) μ отвергается критерием q_μ при заданном доверительном уровне α , и 2) чувствительность к μ достаточна, т. е. $M_0(\mu) \geq M_{\min}$.

Все значения μ , не удовлетворяющие условиям 1) или 2), составляют искомый доверительный интервал.

Вероятность, с которой доверительный интервал покрывает данное значение μ , равна 100% для тех μ , для которых функция мощности меньше порогового значения, и α для тех μ , для которых мощность больше либо равна пороговой. Выбор порогового значения M_{\min} остается за экспериментатором.

Таким образом, построение Кауэна и др.:

- 1) содержит перекрытие (избыточное вероятностное содержание доверительной области) в самой процедуре;
- 2) не дает ясной интерпретации значения M_{\min} , которое и определяет конечный вид интервалов;
- 3) не решает проблему уменьшения доверительных интервалов в нефизической области.

3. ПОСТРОЕНИЕ ФЕЛЬДМАНА И КАЗИНСА

Построение доверительных интервалов, предложенное Фельдманом и Казинсом [6], так же как и способ Стерна, Кроу и Гарднера, основано на специальном порядке добавления точек в доверительную область. Порядок добавления точек определяется отношением правдоподобий.

Рассмотрим следующий пример. Пусть в эксперименте измеряется число событий n , распределенное по Пуассону, с параметром $(\mu + b)$. Здесь μ — неизвестный параметр, который необходимо оценить, b — известный фон. Если обозначить вероятность получить в эксперименте число событий n при $\mu = \mu_1$ через $P(n|\mu_1)$, то вышеупомянутое отношение правдоподобий есть

$$R(n) = \frac{P(n|\mu_1)}{P(n|\mu_{\text{best}})}, \quad (15)$$

где величина $\mu_{\text{best}} = \max(0, n - b)$ 1) максимизирует $P(n|\mu)$ при данном n ; 2) неотрицательна. Далее, для каждого значения μ точки n добавляются в

доверительную область в порядке уменьшения соответствующей величины $R(n)$ до тех пор, пока суммарная вероятность, содержащаяся в доверительной области, не достигнет заданного доверительного уровня.

Такой рецепт построения доверительных интервалов, как отмечено в самой работе [6], не решает важной проблемы, возникающей в случае меньшего, чем фон, числа событий: доверительный интервал оказывается тем меньше, чем меньшее число событий получено в результате эксперимента. Следовательно, независимо от величины фона в эксперименте может получиться сколь угодно сильное ограничение на величину сигнала.

Аналогичная проблема возникает при применении рецепта [6] и для ограниченных параметров непрерывных распределений. Чем дальше в нефизической области оказывается экспериментальное значение оценки, тем меньший доверительный интервал такому значению соответствует. Рецепт Фельдмана и Казинса дает в этом случае парадоксальный результат — наименее надежные результаты (оценки, лежащие в нефизической области) дают наиболее сильное ограничение на оцениваемый параметр.

Обе указанные проблемы построения [6] приводят к тому, что невозможно провести сравнение не только результатов разных экспериментов, но и результатов одного эксперимента (например, двух разных сеансов), если результаты представлены в форме доверительных интервалов, полученных по рецепту Фельдмана и Казинса.

Невозможность сравнивать доверительные интервалы характерна для всех решений из первой группы (см., например, попытку дальнейшей модификации самих интервалов и порядка построения доверительной области в работе [18]).

Этого существенного недостатка лишены решения второй группы.

4. КОНСТРУКЦИЯ МАНДЕЛЬКЕРНА И ШУЛЬЦА

Предыдущие исследователи использовали изменение порядка добавления точек в доверительную область для получения доверительных интервалов с учетом дополнительной информации о параметрах.

Однако их построения приводят к нефизически коротким доверительным интервалам вблизи границы физически допустимых значений параметра. Это происходит, оттого что способ построения оценки фактически не учитывает физического ограничения на интересующие параметры, и используется та же оценка, что и в случае без априорных ограничений.

Манделькern и Шульц [8] предложили использовать в решении задачи другую оценку. В частном случае, рассматриваемом в работе [8], подходящая оценка находится с использованием метода максимального правдоподобия.

В функцию правдоподобия вводится фактор — функция Хевисайда, явным образом отражающая условие ограничения параметра, и получается оценка, всегда лежащая в физической области для данной задачи. Между тем, введение такого фактора выглядит искусственным (постулируется), а корректность рецепта становится очевидной лишь постфактум — сравнением распределения для оценки из [8] с общим решением [9].

Дальнейший вывод доверительных интервалов производится без каких-либо дополнительных предположений.

Отметим, что такая процедура не является байесовской: исключение нефизических значений параметра — не аналог введения равномерной байесовской априорной функции распределения.

Частное решение [8], вообще говоря, корректно (хотя, повторим, убедиться в этом проще всего, сравнив его с общим решением): такое построение

1) в отличие от построения Фельдмана и Казинса формально решает проблему меньшего, чем фон, измеренного числа событий (для пуассоновского процесса с фоном) и уменьшения доверительного интервала для отрицательных значений оценки заведомо положительного параметра гауссовского распределения;

2) дает правильное вероятностное содержание доверительной области, в отличие от рецептов flip-flop, нарушающих условия (3) и (10).

С другой стороны, рецепт построения оценки здесь основан на методе максимального правдоподобия и исключает целый класс задач, где изначально выбран другой метод оценивания.

Вероятно, все эти причины и обусловили тот факт, что подход Манделькерна и Шульца почти не использовался на практике.

5. МЕТОД ПРЕДЕЛА ЧУВСТВИТЕЛЬНОСТИ ДЛЯ ПАРАМЕТРА НЕПРЕРЫВНОГО РАСПРЕДЕЛЕНИЯ

Полное решение задачи о доверительных интервалах для параметра непрерывного распределения с учетом априорной информации об ограниченной области значений параметра дается в работе [9].

Как уже отмечалось, ключевым элементом в построении доверительных интервалов является оценка (estimator). Можно поставить вопрос: как правильно выбрать оценку $\hat{\theta} = \hat{\theta}(X)$, если заранее известно, что $\theta \geq 0$?

Определяющим свойством любой оценки является то, что она дает значение, наиболее близкое к неизвестному θ . Тогда определим новую оценку как

$$\tilde{\theta} = \max(\hat{\theta}, 0). \quad (16)$$

Очевидно, что $\tilde{\theta}$ дает оценки, которые заведомо ближе к неизвестному значению θ , чем $\hat{\theta}$. Такая оценка содержит как статистическую информацию, заключенную в обычной оценке $\hat{\theta}$, так и априорную информацию о том, что $\theta \geq 0$. После этого остается построить доверительные интервалы для новой оценки $\tilde{\theta}$.

Прежде чем продолжать построение, обсудим определение (16).

Распределение вероятности для $\tilde{\theta}$ имеет следующий вид:

$$\tilde{d}_{\theta}(\tilde{\theta}) = H(\tilde{\theta}) d_{\theta}(\tilde{\theta}) + c_{\theta} \delta(\tilde{\theta}), \quad (17)$$

где $H(t)$ — обычная функция Хевисайда, $\delta(t)$ — дираковская δ -функция,

$$c_{\theta} = \int_{-\infty}^0 d\hat{\theta} d_{\theta}(\hat{\theta}). \quad (18)$$

Следовательно, возникает дополнительная сложность, связанная с сингулярным вкладом в (17). Для работы с таким вкладом можно использовать стандартные методы либо применить специальный прием. Применение стандартного подхода (регуляризации) описано в работе [9].

Перейдем к описанию специального приема. Воспользуемся следующим наблюдением. Определение (16) означает, что случайные значения обычной оценки $\hat{\theta}$ в итоге переносятся в точку ноль и скапливаются в этой точке. Это означает, что все такие значения становятся неразличимыми: они дадут нулевое значение модифицированной оценки $\tilde{\theta}$ — и, следовательно, один и тот же доверительный интервал. Тогда все неположительные значения $\hat{\theta}$ дают в итоге один и тот же доверительный интервал $[0, \text{const}]$, где константа не зависит от $\hat{\theta}$.

Как только это становится понятным, построение доверительных интервалов может быть произведено целиком в терминах обычной оценки $\hat{\theta}$; модифицированная оценка $\tilde{\theta}$ используется лишь как дополнительное условие: итоговые системы доверительных интервалов должны удовлетворять тому условию, что все значения $\hat{\theta}$ за априорной границей должны давать одинаковый доверительный интервал.

Это дополнительное условие имеет ясный экспериментальный смысл: условие можно сформулировать как требование устойчивости доверительных интервалов к возможным экспериментальным дефектам, что придает всей конструкции дополнительный физический смысл. Следует, однако, помнить, что приведенное рассуждение прозрачно и конкретно и само по себе не требует неких метафизических обоснований: построение интервалов в терминах обычной оценки $\hat{\theta}$ с упомянутым выше дополнительным условием эквивалентно непосредственному построению доверительных поясов для модифицированной оценки (16), которая включает в себя априорную информацию наиболее простым и прозрачным способом.

5.1. Горизонтальные деформации. Горизонтальные деформации — визуализация того, как можно корректно модифицировать доверительные пояса (системы доверительных интервалов), не получая при этом нежелательных свойств: превращение либо недостаток вероятностного содержания в данном поясе. С помощью горизонтальных деформаций можно получить как систему доверительных интервалов Фельдмана и Казинса, так и систему, соответствующую методу предела чувствительности. Горизонтальные деформации как инструмент противопоставляются flip-flop (вертикальным) деформациям доверительных интервалов. Вертикальные деформации приводят к некорректному вероятностному содержанию доверительных поясов (нарушаются условия (3) и (10)), соответствующие интервалы не являются по определению доверительными.

Прием горизонтальных деформаций основан на следующих свойствах доверительных поясов для фиксированного β . Если для данного θ величина $U(\theta)$, определенная формулой (5), смещается к ее нижней границе $U_{1-\beta}(\theta)$, то соответствующее значение величины $L(\theta) \rightarrow -\infty$ (аналогично можно смещать L к ее верхней границе). Если L и U могут быть таким образом деформированы (при условии постоянного сохранения непрерывности и монотонности), то существуют корректно определенные обратные функции $u = U^{-1}$, $l = L^{-1}$ на каждом шаге деформаций, т. е. существует доверительный пояс для уровня доверия β .

На наших графиках такие деформации происходят в горизонтальных направлениях. Это показано на рис. 3: жирные кривые должны лежать вне внутреннего пояса и могут только приближаться к границам этого пояса (в горизонтальном направлении, как показано на рис. 3 стрелками) с одной стороны, при этом с другой стороны кривая удаляется от границы на бесконечность.

Как только одна из жирных кривых пересекает границу симметричного доверительного пояса (сплошные наклонные прямые), вторая жирная кривая

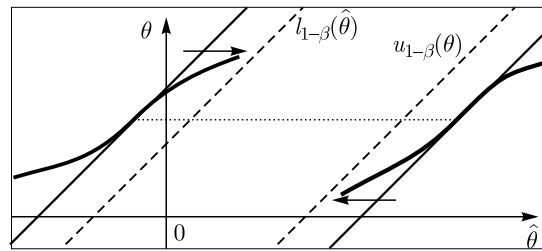


Рис. 3. Пары сплошных и штриховых наклонных прямых ограничивают симметричные доверительные пояса для доверительных уровней $\beta = 1 - 2\alpha$ и $\tilde{\beta} = 1 - 2(1 - \beta) = 1 - 4\alpha$ (ср. рис. 1). Две жирные кривые показывают возможный выбор l , u для доверительного уровня $\beta = 1 - 2\alpha$. Стрелки указывают направление горизонтальных деформаций, обсуждаемых в тексте

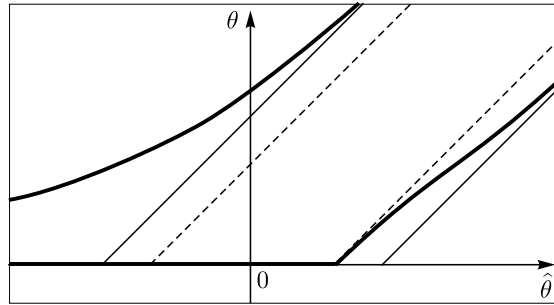


Рис. 4. Доверительный пояс (ограниченный жирными линиями), определенный в работе [6] (ср. с рис. 10 из указанной работы). Остальные линии те же, что и на рис. 3. Правой жирной кривой, приближающейся к штриховой границе, соответствует левая кривая, уходящая на бесконечность с приближением к горизонтальной оси (ср. рис. 3)

пересекает другую границу, как это показано с помощью горизонтальной пунктирной линии.

Описанная свобода была использована в работе [6], где пара L, U была выбрана на основе веры в «магию правдоподобий». Выбор системы интервалов из работы [6] показан на рис. 4.

Наконец, можно перейти к пределу, деформируя L таким образом, что ее часть становится частью кривой $L_{1-\beta}(\theta)$ (рис. 5). Это может привести к тому, что L (и U) перестанут быть непрерывными на границе этой части. Однако если соответствующие обратные функции $u = U^{-1}$, $l = L^{-1}$ непрерывно приближаются в пределе к хорошо определенным непрерывным монотонным

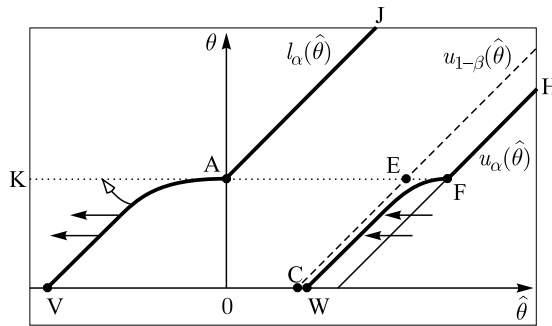


Рис. 5. Жирные линии ограничивают допустимый доверительный пояс для $\hat{\theta}$ и доверительного уровня β . Остальные линии и точки совпадают с показанными на рис. 2. Жирные кривые описываются двумя функциями l и u . Черными стрелками обозначена горизонтальная деформация, применяемая для получения доверительного пояса, который удовлетворяет дополнительному условию $\theta \geq 0$

(неубывающим) функциям величины $\hat{\theta}$, то система доверительных интервалов (6) будет непрерывно приближаться к хорошо определенному результату, а предельный доверительный пояс будет настолько же хорош, как и любой допустимый доверительный пояс для оценки параметров.

Корректное построение доверительных интервалов по методу предела чувствительности можно провести с использованием описанного приема горизонтальных деформаций. В обозначениях рис. 2 и для фиксированного доверительного уровня β рассмотрим доверительный пояс, соответствующий двум функциям l, u , выбранным так, как показано на рис. 5 (ср. также рис. 3 и 4).

Как обсуждалось выше, можно проводить горизонтальную деформацию, обозначенную на рис. 5 черными стрелками, пока сегмент кривой WF не достигнет ломаной CEF. Сегмент кривой AV фактически выпрямится, превратившись в прямую АК (белая стрелка). Предельный доверительный пояс хорошо определен и показан на рис. 6.

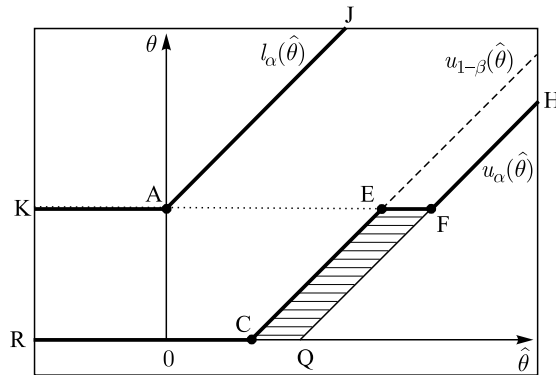


Рис. 6. Система доверительных интервалов в методе предела чувствительности. Две ломаные жирные линии KAJ и RCFH ограничивают окончательный доверительный пояс с уровнем доверия β для оценки $\hat{\theta}$, который удовлетворяет дополнительному условию и, следовательно, корректно отображается на доверительный пояс для $\hat{\theta}$, определенной уравнением (16). Область CEFQ, вырезанная из немодифицированного доверительного пояса, представляет собой выигрыш, полученный за счет априорной информации

5.2. Доверительные интервалы в методе предела чувствительности.

Окончательная система доверительных интервалов, построенных по методу предела чувствительности для параметров непрерывных распределений, представлена на рис. 6 и может быть описана аналитически следующим образом (обозначения соответствуют рис. 6):

— В области $\hat{\theta} \geq \hat{\theta}_F$ доверительный пояс $[u_{\alpha}(\hat{\theta}), l_{\alpha}(\hat{\theta})]$ совпадает с симметричным доверительным интервалом для $\hat{\theta}$ и доверительного уровня $\beta = 1 - 2\alpha$, когда параметр не ограничен снизу.

— В области $\hat{\theta}_E \leq \hat{\theta} \leq \hat{\theta}_F$ доверительный интервал имеет вид $[\theta_A, l_\alpha(\hat{\theta})]$, т. е. верхняя граница имеет тот же вид, что и при $\hat{\theta} \geq \hat{\theta}_F$, а нижняя граница неизменна и равна θ_A .

— В области $\hat{\theta}_C \leq \hat{\theta} \leq \hat{\theta}_E$ доверительный интервал имеет вид $[u_{1-\beta}(\hat{\theta}), l_\alpha(\hat{\theta})]$, т. е. верхняя граница сохраняется, а нижняя граница совпадает с нижней границей симметричного доверительного интервала для доверительного уровня $\tilde{\beta} = 1 - 2(1 - \beta) = 1 - 4\alpha$.

— В области $0 \leq \hat{\theta} \leq \hat{\theta}_C$ доверительный интервал имеет вид $[0, l_\alpha(\hat{\theta})]$, т. е. верхняя граница сохраняется, а нижняя равна 0.

— Наконец, для всех $\hat{\theta} \leq 0$ доверительный интервал одинаков и имеет вид $[0, l_\alpha(0)]$.

Подчеркнем для ясности, что конструкция использует два обычных симметричных доверительных интервала для обычной (не учитывающей априорную информацию) оценки $\hat{\theta}$:

1) симметричный пояс для уровня доверия $\beta = 1 - 2\alpha$, в наших обозначениях имеющий вид $[u_\alpha(\hat{\theta}), l_\alpha(\hat{\theta})]$;

2) симметричный пояс для уровня доверия $\tilde{\beta} = 1 - 2(1 - \beta) = 1 - 4\alpha$, имеющий вид $[u_{1-\beta}(\hat{\theta}), l_{1-\beta}(\hat{\theta})]$.

Поэтому для построения системы доверительных интервалов по методу предела чувствительности не требуется дополнительных громоздких вычислений.

Верхняя граница доверительного интервала $[0, l_\alpha(0)]$, отвечающего всем нефизическим значениям $\hat{\theta}$, и определяет так называемый *предел чувствительности* $\theta_A = l_\alpha(0)$. Величина предела чувствительности, очевидно, не зависит от конкретной величины $\hat{\theta}$, а задается только ошибкой определения $\hat{\theta}$. Таким образом, предел чувствительности отражает величину экспериментальной ошибки и дает объективное представление результатов эксперимента, если экспериментальная оценка параметра лежит в нефизической области.

6. МЕТОД ПРЕДЕЛА ЧУВСТВИТЕЛЬНОСТИ В СЛУЧАЕ ДИСКРЕТНЫХ РАСПРЕДЕЛЕНИЙ

Рассмотрим снова случай, когда в эксперименте детектируется число событий n , причем n имеет распределение Пуассона (9) с параметром μ . Учтем теперь наличие фоновых событий. Число фоновых событий — измеряемая величина, следовательно, в общем случае известно распределение $P_\beta(b)$. Здесь β — неизвестное истинное значение для среднего числа фоновых событий.

В работе [6] рассматривался случай, когда среднее число фоновых событий известно точно и равно b . Тогда число событий, зарегистрированных в

эксперименте, будет распределено по Пуассону со средним $(\mu + b)$:

$$P_{\mu}(n) = \frac{(\mu + b)^n}{n!} e^{-(\mu + b)}. \quad (19)$$

Используем теперь дополнительную информацию о фоне и построим доверительные интервалы для параметра μ по методу предела чувствительности, в духе рассуждений [9]. Основная идея состоит в том, что априорная информация о параметре может быть снова учтена в выборе оценки (estimator) для данного параметра. После этого построение интервалов производится автоматически (ср. с предыдущими попытками [6] модификации непосредственно доверительных интервалов, содержащими в себе элемент произвола). В частном случае, когда среднее значение для фона известно точно, можно применить прием из [9] и выбрать в качестве оценки (estimator) следующую величину:

$$(\mu + b) = \max(n, b). \quad (20)$$

Очевидно, что измеряемая в эксперименте величина n может быть выбрана в качестве оценки для $(\mu + b)$, но, в отличие от (20), допускает отрицательные значения для μ . Измеряя n и используя оценку (20), в результате вычитания постоянного фона b получим заведомо неотрицательную оценку параметра μ .

Распределение для оценки (20) позволяет строить доверительные интервалы для $(\mu + b)$ и, следовательно, для величины μ . Дискретность распределения для величины n (19) позволяет провести следующее рассуждение. Для любого заданного μ известна вероятность получить при измерении значение $n \leq b$, она равна

$$P(n \leq b) = \sum_{n=0}^{[b]} \frac{(\mu + b)^n}{n!} e^{-(\mu + b)}, \quad (21)$$

где $[b]$ обозначает, как обычно, целую часть числа. Тогда, используя оценку (20), в результате измерений будем получать значение этой оценки, равное b с вероятностью (21).

Таким образом, распределение вероятностей для случайной величины $\max(n, b)$ будет состоять из распределения (19) при значениях $n > b$, и вероятности (21) получить при измерениях величину b (рис. 7).

Используя полученное распределение вероятностей, можно, следуя п. 1.2, построить доверительный интервал для величины $(\mu + b)$ и, соответственно, интересующего нас параметра μ .

Для практического применения удобно указать доверительные интервалы в переменных μ и n (рис. 8). В таком представлении при получении в эксперименте значения n_0 можно непосредственно указать доверительный интервал для параметра μ . Заметим, что все значения $n \leq b$ неразличимы после перехода к оценке вида (20). Поэтому всем этим значениям соответствует один и

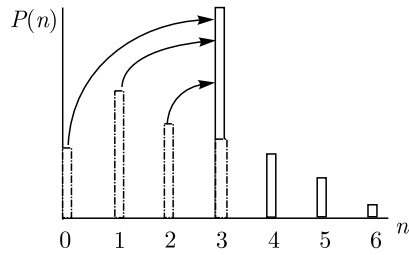


Рис. 7. Схема изменения распределения вероятностей при переходе от оценки $(\mu + b) = n$ (штриховой график в области $n \leq b$) к оценке вида $(\mu + b) = \max(n, b)$ (сплошной график) при величине фона $b = 3$. При $n > b$ значения вероятностей для обычной и модифицированной оценок совпадают

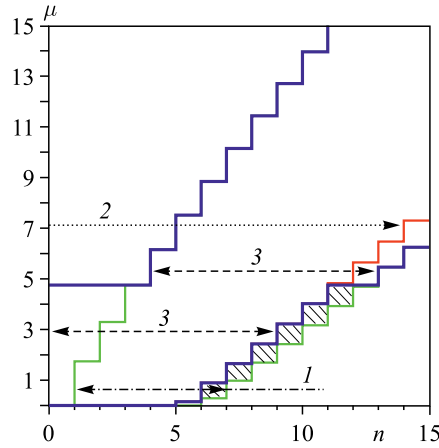


Рис. 8 (цветной в электронной версии). Доверительные интервалы для неизвестного сигнала μ при наличии пуассоновского фона со средним $b = 3$ (зеленый (1) — 90 %-й симметричный доверительный интервал без учета априорной информации; красный (2) — односторонний 90 %-й интервал без учета информации о фоне; синий (3) — 90 %-й доверительный интервал, полученный по методу предела чувствительности, с учетом априорной информации в оценке (estimator) параметра μ)

тот же доверительный интервал. Верхняя граница такого интервала соответствует так называемому *пределу чувствительности*. Она несет информацию о величине фона в конкретном эксперименте.

Рассмотрим доверительные интервалы для параметра дискретного распределения в методе предела чувствительности. Система доверительных интервалов для параметра пуассоновского распределения с известным фоном, построенная по методу предела чувствительности, показана на рис. 8 (доверительный уровень $\beta = 90\%$). Она может быть описана аналитически следующим образом:

- Верхняя граница интервала при $n > b$ совпадает с границей симметричного доверительного интервала для $\beta = 90\%$.

- В области $n \leq b$ верхняя граница доверительного интервала неизменна и равна μ_b .

Нижняя граница разбивается (по аналогии со случаем непрерывных распределений) на четыре участка:

- при малых n нижняя граница интервала равна 0;
- следующий участок совпадает с границей одностороннего интервала с доверительным уровнем $\beta = 90\%$;

— переходный участок между границами одностороннего и симметричного интервала: нижняя граница равна μ_b ;

— при больших n нижняя граница доверительного интервала совпадает с нижней границей симметричного доверительного интервала с доверительным уровнем $\beta = 90\%$.

Расчет системы доверительных интервалов по методу предела чувствительности для пуассоновского процесса с фоном необходимо производить с использованием специально разработанного для этого программного обеспечения [20], так как простых аналитических формул здесь нет (см. неявную формулу (13); в программе для нахождения μ^* используется некоторый алгоритм поиска).

На рис. 9 приведены интервалы, построенные по методу Фельдмана и Казинса (серый, 2), и интервалы, построенные с учетом априорной информации в оценке по методу предела чувствительности (синий, 1). Очевидно, что доверительные интервалы с корректным учетом априорной информации (синий график, 1)

- 1) дают корректную оценку в области $n \leq b$;
- 2) имеют участок, аналогичный участку СЕФ на рис. 6, дают наилучшую оценку нижней границы доверительного интервала;
- 3) по построению обеспечивают близкое к 90% вероятностное содержание доверительной области (для дискретных распределений получить 90%-е содержание в общем случае невозможно).

Для удобства сравнения с конструкцией [6] мы приводим таблицу, аналогичную табл. II–IX в работе [6], в которой даются интервалы для доверительного уровня 90% и для различных значений фона и измеренного числа событий (b от 0 до 10, $n_0 = 0–20$).

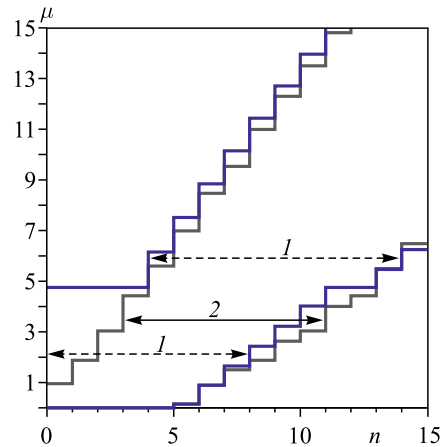


Рис. 9 (цветной в электронной версии). Доверительные интервалы для неизвестного сигнала μ при наличии пуассоновского фона со средним $b = 3$ (синий (1) — 90%-й доверительный интервал с учетом априорной информации (аналогично рис. 8), серый (2) — 90%-й доверительный интервал, построенный по методу Фельдмана и Казинса [6])

90 %-е доверительные интервалы с учетом априорной информации для параметра μ , при известном значении фона b в интервале от 0 до 10 и при измеренном числе событий $n_0 = (0, \dots, 20)$ (аналогично табл. II–IX в работе [6])

№	0,0	0,5	1,0	1,5	2,0	2,5	3,0
0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
1	0,11	0,0	0,0	0,0	0,0	0,0	0,0
2	0,54	0,04	0,0	0,0	0,0	0,0	0,0
3	1,11	0,61	0,11	0,0	0,0	0,0	0,0
4	1,75	1,25	0,75	0,25	0,0	0,0	0,0
5	2,44	1,94	1,44	0,94	0,44	0,0	0,0
6	3,0	2,5	2,16	1,66	1,16	0,66	0,16
7	3,29	2,79	2,9	2,4	1,9	1,4	0,9
8	3,99	3,49	3,66	3,16	2,66	2,16	1,66
9	4,7	4,2	3,75	3,25	3,44	2,94	2,44
10	5,43	4,93	4,43	3,93	4,23	3,73	3,23
11	6,17	5,67	5,17	4,67	4,3	3,8	4,03
12	6,93	6,43	5,93	5,43	4,93	4,43	4,76
13	7,69	7,19	6,69	6,19	5,69	5,19	4,76
14	8,47	7,97	7,47	6,97	6,47	5,97	5,47
15	9,25	8,75	8,25	7,75	7,25	6,75	6,25
16	10,04	9,54	9,04	8,54	8,04	7,54	7,04
17	10,84	10,34	9,84	9,34	8,84	8,34	7,84
18	11,64	11,14	10,64	10,14	9,64	9,14	8,64
19	12,45	11,95	11,45	10,95	10,45	9,95	9,45
20	13,26	12,76	12,26	11,76	11,26	10,76	10,26

Продолжение таблицы

№	3,5	4,0	4,5	5,0	5,5	6,0	6,5
0	0,0, 4,26	0,0, 5,16	0,0, 4,66	0,0, 5,52	0,0, 5,02	0,0, 5,85	0,0, 5,35
1	0,0, 4,26	0,0, 5,16	0,0, 4,66	0,0, 5,52	0,0, 5,02	0,0, 5,85	0,0, 5,35
2	0,0, 4,26	0,0, 5,16	0,0, 4,66	0,0, 5,52	0,0, 5,02	0,0, 5,85	0,0, 5,35
3	0,0, 4,26	0,0, 5,16	0,0, 4,66	0,0, 5,52	0,0, 5,02	0,0, 5,85	0,0, 5,35
4	0,0, 5,66	0,0, 5,16	0,0, 4,66	0,0, 5,52	0,0, 5,02	0,0, 5,85	0,0, 5,35
5	0,0, 7,02	0,0, 6,52	0,0, 6,02	0,0, 5,52	0,0, 5,02	0,0, 5,85	0,0, 5,35
6	0,0, 8,35	0,0, 7,85	0,0, 7,35	0,0, 6,85	0,0, 6,35	0,0, 5,85	0,0, 5,35
7	0,4, 9,65	0,0, 9,15	0,0, 8,65	0,0, 8,15	0,0, 7,65	0,0, 7,15	0,0, 6,65
8	1,16, 10,94	0,66, 10,44	0,16, 9,94	0,0, 9,44	0,0, 8,94	0,0, 8,44	0,0, 7,94
9	1,94, 12,21	1,44, 11,71	0,94, 11,21	0,44, 10,71	0,0, 10,21	0,0, 9,71	0,0, 9,21
10	2,73, 13,47	2,23, 12,97	1,73, 12,47	1,23, 11,97	0,73, 11,47	0,23, 10,97	0,0, 10,47
11	3,53, 14,71	3,03, 14,21	2,53, 13,71	2,03, 13,21	1,53, 12,71	1,03, 12,21	0,53, 11,71
12	4,26, 15,95	3,83, 15,45	3,33, 14,95	2,83, 14,45	2,33, 13,95	1,83, 13,45	1,33, 12,95
13	4,26, 17,17	4,65, 16,67	4,15, 16,17	3,65, 15,67	3,15, 15,17	2,65, 14,67	2,15, 14,17
14	4,97, 18,39	5,16, 17,89	4,66, 17,39	4,47, 16,89	3,97, 16,39	3,47, 15,89	2,97, 15,39
15	5,75, 19,6	5,25, 19,1	4,75, 18,6	5,3, 18,1	4,8, 17,6	4,3, 17,1	3,8, 16,6
16	6,54, 20,81	6,04, 20,31	5,54, 19,81	5,52, 19,31	5,02, 18,81	5,14, 18,31	4,64, 17,81
17	7,34, 22,0	6,84, 21,5	6,34, 21,0	5,84, 20,5	5,34, 20,0	5,85, 19,5	5,35, 19,0
18	8,14, 23,2	7,64, 22,7	7,14, 22,2	6,64, 21,7	6,14, 21,2	5,85, 20,7	5,35, 20,2
19	8,95, 24,38	8,45, 23,88	7,95, 23,38	7,45, 22,88	6,95, 22,38	6,45, 21,88	5,95, 21,38
20	9,76, 25,57	9,26, 25,07	8,76, 24,57	8,26, 24,07	7,76, 23,57	7,26, 23,07	6,76, 22,57

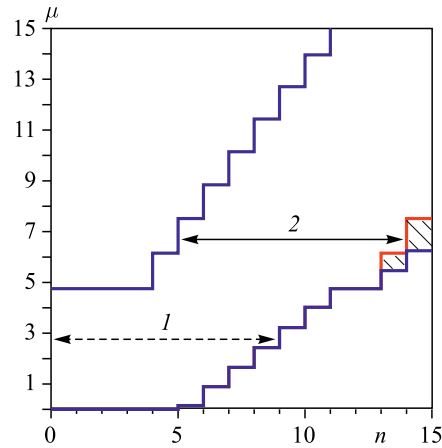
Окончание таблицы

№	7,0	7,5	8,0	8,5	9,0	9,5	10,0					
0	0,0	6,15	0,0	5,65	0,0	6,44	0,0	6,71	0,0	6,21	0,0	6,97
1	0,0	6,15	0,0	5,65	0,0	6,44	0,0	6,71	0,0	6,21	0,0	6,97
2	0,0	6,15	0,0	5,65	0,0	6,44	0,0	6,71	0,0	6,21	0,0	6,97
3	0,0	6,15	0,0	5,65	0,0	6,44	0,0	6,71	0,0	6,21	0,0	6,97
4	0,0	6,15	0,0	5,65	0,0	6,44	0,0	6,71	0,0	6,21	0,0	6,97
5	0,0	6,15	0,0	5,65	0,0	6,44	0,0	6,71	0,0	6,21	0,0	6,97
6	0,0	6,15	0,0	5,65	0,0	6,44	0,0	6,71	0,0	6,21	0,0	6,97
7	0,0	6,15	0,0	5,65	0,0	6,44	0,0	6,71	0,0	6,21	0,0	6,97
8	0,0	7,44	0,0	6,94	0,0	6,44	0,0	6,71	0,0	6,21	0,0	6,97
9	0,0	8,71	0,0	8,21	0,0	7,71	0,0	7,21	0,0	6,71	0,0	6,97
10	0,0	9,97	0,0	9,47	0,0	8,97	0,0	8,47	0,0	7,97	0,0	6,97
11	0,03	11,21	0,0	10,71	0,0	10,21	0,0	9,71	0,0	9,21	0,0	8,21
12	0,83	12,45	0,33	11,95	0,0	11,45	0,0	10,95	0,0	10,45	0,0	9,45
13	1,65	13,67	1,15	13,17	0,65	12,67	0,15	12,17	0,0	11,67	0,0	10,67
14	2,47	14,89	1,97	14,39	1,47	13,89	0,97	13,39	0,47	12,89	0,0	11,89
15	3,3	16,1	2,8	15,6	2,3	15,1	1,8	14,6	1,3	14,1	0,8	13,1
16	4,14	17,31	3,64	16,81	3,14	16,31	2,64	15,81	2,14	15,31	1,64	14,31
17	4,98	18,5	4,48	18,0	3,98	17,5	3,48	17,0	2,98	16,5	2,48	15,5
18	5,83	19,7	5,33	19,2	4,83	18,7	4,33	18,2	3,83	17,7	3,33	16,7
19	6,15	20,88	5,65	20,38	5,68	19,88	5,18	19,38	4,68	18,88	4,18	17,88
20	6,26	22,07	5,76	21,57	6,44	21,07	5,94	20,57	5,53	20,07	5,03	19,57

Программное обеспечение, позволяющее рассчитывать интервалы для любых значений доверительного уровня параметров μ , b и n_0 , а также рисовать соответствующие картинки, свободно доступно с адреса [20].

Благодаря дополнительной свободе, возникающей при построении доверительных интервалов для параметров дискретных распределений (замена точного равенства (11) на \geq (10)), можно минимизировать вероятностное содержание доверительной области, приблизить его к требуемому доверительному уровню (в наших примерах — 90%). Например, на рис. 10 область правее/ниже верхней границы удовлетворяет условию $P_{\hat{\mu}}(n_1(\mu) \geq n) \geq 0,95$. Эта верхняя граница задает односторонний 95%-й доверительный интервал. Как указывалось выше (п. 1.2), нижнюю границу можно построить формально, исходя из условия $P_{\hat{\mu}}(n_2(\mu) \leq n) \geq 0,95$ (синий график (1) на рис. 10). При этом заведомо выполняется условие $P_{\hat{\mu}}(n_1(\mu) \geq n \geq n_2(\mu)) \geq 0,90$.

Рис. 10 (цветной в электронной версии). Доверительные интервалы для неизвестного сигнала μ при наличии пуассоновского фона со средним $b = 3$ (синий (1) — 90%-й доверительный интервал с учетом априорной информации, красный (2) — 90%-й доверительный интервал с учетом априорной информации о фоне, построенный для случая, когда важно найти ограничение сверху на исследуемый параметр μ)



Если верхняя граница доверительных интервалов фиксирована, то можно строить нижнюю границу $n_2(\mu)$ сразу исходя из условия $P_{\hat{\mu}}(n_1(\mu) \geq n \geq n_2(\mu)) \geq 0,90$ (красный график (2) на рис. 10). Вероятностное содержание доверительной области внутри красного графика (2) будет меньше либо равно содержанию области внутри синего графика (1). Таким образом, отказавшись от симметричности, можно получить доверительные интервалы, более близкие к требуемому доверительному уровню по вероятностному содержанию.

Для полноты картины можно также сравнить (рис. 11) доверительный интервал, построенный по методу предела чувствительности (синий график, 1), и интервал, построенный в соответствии с процедурой Стерна, Кроу и Гарднера (темно-серый график, 2). Построение Стерна, Кроу и Гарднера производилось для оценки вида (20). В области $n < b = 3$ серый график дает пустой доверительный интервал, поскольку оценка (20) не принимает значе-

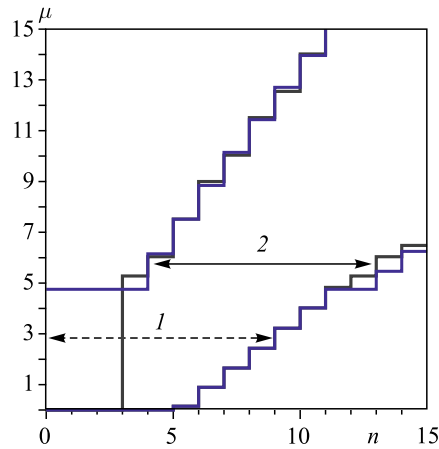


Рис. 11 (цветной в электронной версии). Доверительные интервалы для неизвестного сигнала μ при наличии пуассоновского фона со средним $b = 3$ (синий (1) — 90%-й доверительный интервал с учетом априорной информации по методу предела чувствительности, темно-серый (2) — 90%-й доверительный интервал для оценки (estimator) вида (20), построенный с помощью процедуры Стерна, Кроу и Гарднера (наиболее короткая система доверительных интервалов))

ний в этой области. Такие доверительные интервалы по построению имеют наименьшую длину при заданном доверительном уровне.

В результате учета информации о фоне по методу предела чувствительности получено множество доверительных интервалов, что привело к следующему:

- решена обозначенная в [6] проблема меньшего числа зарегистрированных событий, чем ожидаемый фон (в области, где число событий меньше фона, доверительный интервал естественным образом дает оценку сверху для измеряемого параметра μ , не зависящую от значения n , как в области $n \leq 3$ на рис. 8);

- нижняя граница интервала имеет участок, аналогичный участку CEF на рис. 6;

- благодаря неоднозначности определения интервалов для дискретных распределений (условие (10)) возможны различные варианты построения доверительных интервалов (фиксация нижней или верхней границы, построение СКГ, их комбинации). Выбор может зависеть от конкретной ситуации, например, от необходимости получить более строгое ограничение на параметры сверху или снизу.

Таким образом, корректный учет априорной информации о наличии фона при выборе оценки позволяет строить доверительные интервалы, отвечающие требованиям физической реалистичности и лишенные недостатков построений [6, 7] и [18]. По аналогии со случаем непрерывных распределений рассматриваемый метод предела чувствительности позволяет сравнивать доверительные интервалы, полученные в разных экспериментах.

7. НАИЛУЧШИЙ ВЕРХНИЙ ПРЕДЕЛ (НЕПРЕРЫВНЫЙ И ДИСКРЕТНЫЙ СЛУЧАИ) В МЕТОДЕ ПРЕДЕЛА ЧУВСТВИТЕЛЬНОСТИ

7.1. Непрерывный случай. Рассмотрим теперь в рамках метода предела чувствительности важный для экспериментальных приложений случай несимметричного доверительного интервала.

В разд. 5 обычная оценка $\hat{\theta}$ параметра θ переопределялась таким образом, что оценка учитывала априорное неравенство $\theta \geq 0$. После этого для преопределенной оценки $\hat{\theta} = \max(\hat{\theta}, 0)$ (уравнение (16)) доверительный пояс строился обычным способом. Работа с δ -образным вкладом в распределение вероятностей для $\hat{\theta}$ упростилась благодаря сведению задачи к построению доверительного пояса для обычной оценки $\hat{\theta}$ таким образом, чтобы итоговый пояс удовлетворял дополнительному условию (см. разд. 5). Построение было проведено с помощью приема так называемых горизонтальных деформаций, полученная система интервалов представлена на рис. 6.

В разд. 5 был модифицирован стандартный симметричный доверительный пояс, соответствующий выбору $\alpha = \alpha' = (1 - \beta)/2$ в обозначениях рис. 1. Естественным развитием данного построения будет такая же модификация для асимметричного случая $\alpha' = 0$, $\beta = 1 - \alpha$, который соответствует верхней границе для доверительного уровня β :

$$P(\theta < l_{1-\beta}(\hat{\theta})) = \beta. \quad (22)$$

Такой вариант удобно использовать, если необходимо измерить положительный сигнал, в то время как точность не позволяет доказать наличие ненулевого сигнала с достаточной доверительной вероятностью. В этом случае желательно указать наиболее компактный верхний предел. Такая постановка задачи и ее решение рассмотрены в работе [12].

Следуя работе [12], модифицируем доверительный интервал (22) так, чтобы он удовлетворял априорному неравенству $\theta \geq 0$.

Необходимое геометрическое построение дается на рис. 12, который отличается от рис. 2 добавленными точками пересечения (точки MBDN на горизонтали LG).

В последующем построении будут участвовать только точки M, B, C, D, N; остальные точки указаны для явной связи с рис. 2.

Величина θ_B есть ордината точки B (а также точек M, D и N):

$$\theta_B = l_{1-\beta}(0). \quad (23)$$

Величины $\theta_C < \theta_D$ суть абсциссы точек C и D:

$$\theta_C = U_{1-\beta}(0), \quad \theta_D = U_{1-\beta}(\theta_B). \quad (24)$$

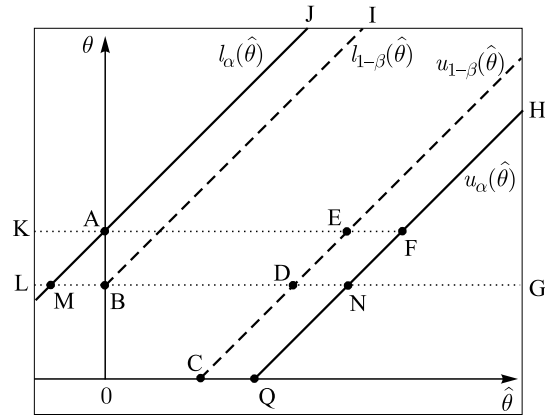


Рис. 12. Пары сплошных и штриховых наклонных линий задают симметричные доверительные пояса для доверительных уровней $\beta = 1 - 2\alpha$ и $\tilde{\beta} = 1 - 2(1 - \beta) = 1 - 4\alpha$. Функции, соответствующие этим линиям, обозначены на рисунке. Точки А и В суть пересечения кривых $\theta = l_\alpha(\hat{\theta})$ и $\theta = l_{1-\beta}(\hat{\theta})$ с вертикальной осью. Точки А и В задают горизонтальные линии KF и LG с другими точками пересечения

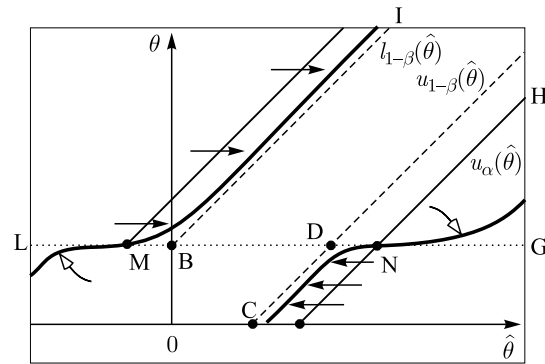


Рис. 13. Две жирные кривые задают допустимый доверительный пояс для доверительного уровня β . Жирные линии «закреплены» в точках М и N. Черные стрелки указывают допустимые горизонтальные деформации. Белыми стрелками указано распрямление соответствующих сегментов

Немодифицированная граница (22) соответствует доверительным интервалам (с уровнем доверия β), которые начинаются на штриховой линии ВI и продолжаютсЯ вниз до бесконечности.

Для получения модификации границы (22) начнем с доверительного пояса $[u(\hat{\theta}), l(\hat{\theta})]$, обозначенного на рис. 13 жирными линиями.

После этого проводим горизонтальные деформации u и l , как показано черными стрелками (подробнее прием описан в разд. 5). В результате сегмент u ниже точки N прижимается к прямой CD, одновременно часть l выше точки M прижимается к прямой VI. Деформации с другой стороны обозначены белыми стрелками. Полученный таким образом доверительный пояс показан на рис. 14.

Рассмотрим несимметричные доверительные интервалы в методе предела чувствительности.

Система доверительных интервалов, представляющая собой наилучшие верхние пределы, модифицированные учетом априорной информации по методу предела чувствительности, представлена на рис. 14. Аналитически построение можно описать следующим образом (мы рассматриваем доверительный уровень β , абсциссы точек B, C и D на рис. 14 обозначены $\hat{\theta}_B$, $\hat{\theta}_C$ и $\hat{\theta}_D$ соответственно):

— для $\hat{\theta} \geq \hat{\theta}_D$ доверительный интервал есть $[\theta_B, l_{1-\beta}(\hat{\theta})]$, т. е. верхняя граница та же самая, что и в немодифицированном случае (уравнение (22)), но ограничен снизу величиной θ_B . Область под ломаной CDG является выигрышем за счет учета априорной информации;

— для $\hat{\theta}_C \leq \hat{\theta} \leq \hat{\theta}_D$ доверительный интервал — $[u_{1-\beta}(\hat{\theta}), l_{1-\beta}(\hat{\theta})]$, т. е. в точности немодифицированный симметричный доверительный интервал для доверительного уровня $\tilde{\beta} = 1 - 2(1 - \beta) = 1 - 4\alpha$;

— для $0 \leq \hat{\theta} \leq \hat{\theta}_C$ доверительный интервал — $[0, l_{1-\beta}(\hat{\theta})]$, т. е. немодифицированная верхняя граница (22), ограниченная снизу физическим условием $\theta \geq 0$;

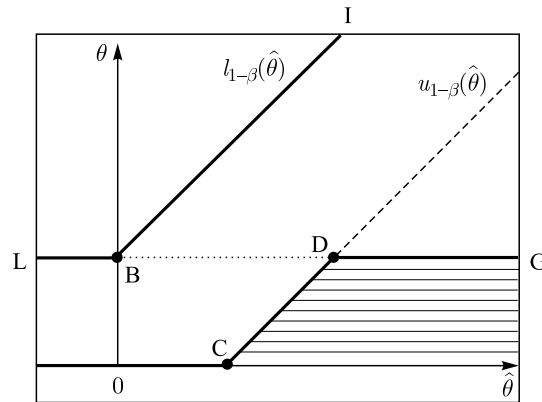


Рис. 14. Наилучший верхний предел в методе предела чувствительности. Доверительный пояс получен из немодифицированной верхней границы для доверительного уровня β с учетом априорной информации $\theta \geq 0$. Область под ломаной CDG является выигрышем за счет учета априорной информации

— наконец, для $\hat{\theta} \leq 0$ доверительный интервал фиксирован и имеет вид $[0, \theta_B]$.

Перечислим примечательные свойства такого доверительного пояса:

— оценивание устойчиво для нефизических значений оценки параметра, т. е. для $\hat{\theta} < 0$;

— верхняя граница интервала для физических значений $\hat{\theta}$ та же, что и для немодифицированного случая (22) и является наименьшей возможной для данного доверительного уровня β ;

— нижняя граница интервала отходит от нуля в точке с наименьшей возможной (для данного доверительного уровня) абсциссой $\hat{\theta}_C$, кроме того, нижняя граница является максимальной возможной для данного доверительного уровня в интервале $\hat{\theta}_C \leq \hat{\theta} \leq \hat{\theta}_D$;

— для вычисления доверительного интервала не требуется сложных алгоритмов или таблиц сверх стандартных процедур построения доверительного интервала для доверительного уровня $\tilde{\beta} = 1 - 2(1 - \beta) = 1 - 4\alpha$;

— важной особенностью такого верхнего предела оказывается отсутствие перекрывания (превышения доверительного уровня), характерного для искусственных рецептов типа [7].

7.2. Дискретный случай. Метод предела чувствительности для параметра пуассоновского распределения с фоном: случай несимметричной системы доверительных интервалов. Для параметров дискретных распределений построение верхнего предела с учетом априорной информации по методу предела чувствительности производится по аналогии с построением в непрерывном случае.

Оценка (estimator) выбирается в виде (20). Далее, аналогично рис. 12 и 13, рассматриваются доверительные интервалы: 90 %-й двусторонний без учета априорной информации, а также верхняя и нижняя границы 90 %-х односторонних интервалов (рис. 15). Хотя для дискретных распределений процедура горизонтальных деформаций не определена, в результате выбора оценки (20) с учетом информации о фоне получим доверительную область (ограничена черными графиками (3) на рис. 15), схожую по конфигурации с доверительным поясом на рис. 14.

Несимметричная система доверительных интервалов (доверительный уровень $\beta = 90\%$), построенная по методу предела чувствительности и отвечающая наилучшему верхнему пределу, показана на рис. 15. Такую систему интервалов можно описать аналитически следующим образом.

Верхняя граница интервала состоит из двух частей:

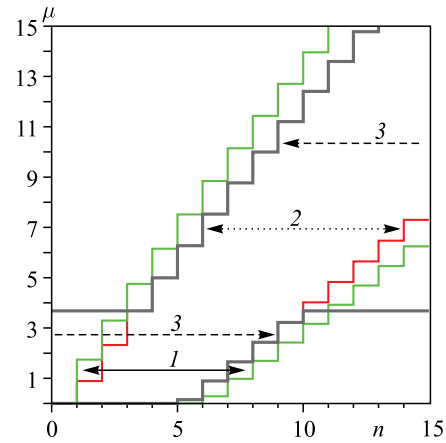
— при $n \leq b$ верхняя граница неизменна и равна μ_b ;

— в области $n > b$ верхняя граница совпадает с границей одностороннего интервала с доверительным уровнем $\beta = 90\%$.

Нижняя граница состоит из трех частей:

— при малых n нижняя граница равна 0;

Рис. 15 (цветной в электронной версии). Наилучший верхний предел в методе предела чувствительности. Верхняя граница доверительного интервала для неизвестного сигнала μ при наличии пуассоновского фона со средним $b = 3$ (зеленый (1) — 90%-й двусторонний доверительный интервал без учета априорной информации, красный (2) — 90%-е верхний и нижние интервалы без учета априорной информации, черный (3) — 90%-й односторонний доверительный интервал для оценки (estimator) вида (20) — наилучшая верхняя граница интервала по аналогии с [12])



— на втором участке нижняя граница совпадает с границей одностороннего интервала с доверительным уровнем $\beta = 90\%$;

— когда нижняя граница достигает величины μ_b , она переходит в константу, равную μ_b .

Как и в случае симметричной системы доверительных интервалов для параметра пуассоновского процесса с фоном (см. разд. 6), расчет наилучшего верхнего предела по методу предела чувствительности удобно производить, используя программное обеспечение [20].

Построенный по методу предела чувствительности наилучший верхний предел для неизвестного сигнала μ при наличии пуассоновского фона со средним b обладает следующими свойствами:

— при измерении в эксперименте любого числа событий $n \leq b$ получаем одинаковый доверительный интервал. Это также решает проблему меньшего, чем фон, числа событий, обозначенную в [6];

— вероятностное содержание доверительной области (между черными графиками (3) на рис. 15) близко к требуемому доверительному уровню. Существенное превышение доверительного уровня (перекрывание в понимании [6]) является недостатком некоторых других построений [19];

— нижняя граница доверительного интервала отрывается от оси n (становится отличной от нуля) при наименьшем возможном значении n . Таким образом, не только указывается наилучший (наиболее сильный) верхний предел для параметра, но и появляется возможность детектировать сигнал (отличная от нуля нижняя граница в некотором смысле обозначает наличие сигнала) при достаточно малых значениях измеренного числа событий;

— как и в непрерывном случае, область под нижней границей является выигрышем от использования априорной информации о фоне.

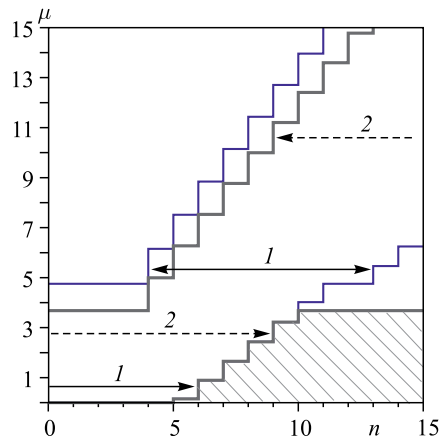


Рис. 16 (цветной в электронной версии). Доверительные интервалы для неизвестного сигнала μ при наличии пуассоновского фона со средним $b = 3$ (синий (1) — 90%-й двусторонний доверительный интервал с учетом априорной информации, серый (2) — 90%-й интервал для оценки вида (20) — наилучшая верхняя граница доверительного интервала по аналогии с [12])

Для сравнения приведем также двусторонний (симметричный) 90%-й интервал с учетом априорной информации (синий график (1), рис. 16) и 90%-й наилучший верхний предел (серый график (2), рис. 16) для оценки вида (20). Видно, что наилучшая верхняя граница доверительного интервала для параметра μ лежит ниже верхней границы двустороннего интервала. Нижние границы совпадают до точки $n = 10$. Очевидно, что двусторонний интервал дает более универсальную оценку — для любого измеренного числа событий. Система доверительных интервалов, определяющая наилучшую верхнюю границу для параметра, по построению также существует для любого значения n . Однако наибольший интерес такая система интервалов представляет в случае, когда n мало (например, $n < 10$ на рис. 16). Именно в этой области с одной стороны верхний предел наиболее строгий, с другой стороны возможно детектирование сигнала (в описанном выше смысле: отрыв нижней границы от нуля).

Еще раз подчеркнем, что корректное построение верхнего предела с учетом априорной информации в непрерывном и дискретном случаях по методу предела чувствительности дает возможность осмысленно сравнивать результаты экспериментов, представленные непосредственно в виде доверительных интервалов.

8. ОЦЕНКИ МАССЫ ЭЛЕКТРОННОГО АНТИНЕЙТРИНО В ЭКСПЕРИМЕНТАХ В МАЙНЦЕ И ТРОИЦКЕ

Хорошей иллюстрацией для свойств системы доверительных интервалов (верхнего предела) служат результаты экспериментов по измерению массы нейтрино в β -распаде трития в Майнце и Троицке. В этих экспериментах

измеряется (фитируется в ходе обработки данных) квадрат массы электронного антинейтрино. Этот параметр заведомо неотрицателен, но его оценка может флуктуировать в отрицательную область. Поэтому возникает необходимость в корректной процедуре построения верхнего предела для массы нейтрино.

В эксперименте в Майнце [10] получены оценки квадрата массы нейтрино $\tilde{m}_\nu^2 = -1,2 \text{ эВ}^2$ и $\tilde{m}_\nu^2 = -0,6 \text{ эВ}^2$ (в зависимости от выбора величины одного из параметров установки — вероятности возбуждения соседних молекул трития в источнике). Очевидно, что отрицательные значения \tilde{m}_ν^2 связаны с недооценкой систематических факторов или со статистической флуктуацией. Суммарная неопределенность составила $\Delta m_\nu^2 = 3,04 \text{ эВ}^2$, поэтому можно получить оценки сверху на массу нейтрино по рецепту Фельдмана и Казинса, например, на доверительном уровне 95%. Для этого в табл. X из работы [6] необходимо найти верхние границы доверительного интервала для $\tilde{m}_\nu^2 = -0,4\Delta m_\nu^2$ и $\tilde{m}_\nu^2 = -0,2\Delta m_\nu^2$: 1,77 и 1,58 соответственно (все величины поделены на величину ошибки). Тогда, умножая эти значения на величину ошибки и извлекая квадратный корень, получим следующие оценки для массы нейтрино: $m_\nu < 2,2 \text{ эВ}$ и $m_\nu < 2,3 \text{ эВ}$.

В работе [10] утверждается, что оценка $\tilde{m}_\nu^2 = -0,6 \text{ эВ}^2$ получена в результате самосогласованного анализа данных и является более надежной, чем $\tilde{m}_\nu^2 = -1,2 \text{ эВ}^2$. Тем не менее, получился парадоксальный результат: рецепт [6] дал лучшее ограничение (2,2 эВ) для более отрицательной и менее надежной оценки \tilde{m}_ν^2 .

Оказывается, что при использовании рецепта [6] невозможно обойтись без дополнительной информации об оценке, а конкретное значение оценки из нефизической области сильно влияет на величину верхнего предела. Тем самым указание двух величин 2,2 и 2,3 эВ в качестве верхнего предела на массу нейтрино недостаточно, чтобы выбрать лучшую из этих оценок при том, что цель представления результатов в виде доверительных интервалов состоит ровно в том, чтобы указать величину, которая позволяла бы непосредственно сравнивать разные эксперименты (а в нашем примере — разные способы обработки одних и тех же данных экспериментов), не требуя при этом дополнительных сведений о полученных в эксперименте оценках.

Воспользуемся теперь методом предела чувствительности для эксперимента в Майнце. Для метода предела чувствительности оценки, лежащие в нефизической области $\tilde{m}_\nu^2 < 0$, неразличимы, и верхняя граница доверительного интервала зависит только от величины ошибки (см. рис. 6 и 14). При доверительном уровне 95% эта верхняя граница равна $1,96\Delta m_\nu^2$. Для эксперимента в Майнце получаем

$$m_\nu < 2,4 \text{ эВ} = \sqrt{1,96 \cdot 3,04 \text{ эВ}^2}$$

для обеих оценок квадрата массы $\tilde{m}_\nu^2 = -1,2 \text{ эВ}^2$ и $\tilde{m}_\nu^2 = -0,6 \text{ эВ}^2$. Дополнительной информации о способе фитирования и величинах других параметров при этом не требуется. Заметим, что если оценка квадрата массы нейтрино оказывается положительной, то доверительный интервал можно получить непосредственно из построенной по методу предела чувствительности системы доверительных интервалов (см., например, рис. 14 в области $\tilde{\theta} > 0$).

Чтобы продемонстрировать удобство метода предела чувствительности для сравнения результатов различных экспериментов, обратимся к результату эксперимента «Троицк-ню-масс» [2]. Оценка квадрата массы нейтрино в [2] равна $\tilde{m}_\nu^2 = -0,67 \text{ эВ}^2$, а величина ошибки $\Delta m_\nu^2 = 2,53 \text{ эВ}^2$. Так как оценка снова лежит в нефизической области, то верхняя граница доверительного интервала зависит только от величины ошибки и равна (доверительный уровень 95 %, рис. 17):

$$m_\nu < 2,2 \text{ эВ} = \sqrt{1,96 \cdot 2,53 \text{ эВ}^2} = \sqrt{4,96 \text{ эВ}^2}.$$

Из двух значений, 2,4 эВ (Майнц) и 2,2 эВ (Троицк), полученных с помощью метода предела чувствительности, можно сделать вывод о том, что эксперимент в Троицке получил лучшее ограничение на массу электронного антинейтрино. При этом нет необходимости обращаться к вопросу об отрицательных значениях оценок квадрата массы нейтрино, не требуется информации о способе фитирования и параметрах фита. Возможность непосредственного сравнения результатов различных экспериментов является несомненным достоинством метода предела чувствительности. При использова-

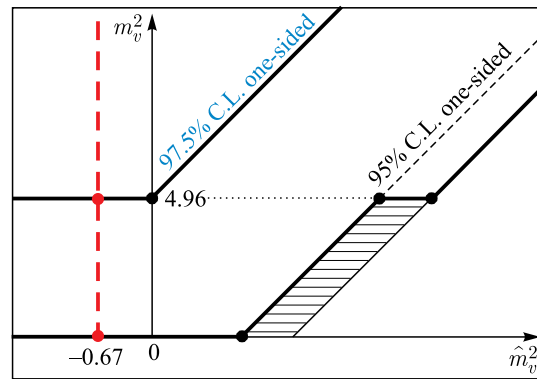


Рис. 17. Использование системы доверительных интервалов, построенных по методу предела чувствительности. Жирные черные линии соответствуют доверительному поясу с уровнем доверия 95 % (см. также рис. 6). Доверительный интервал для экспериментальной оценки определяется ординатами точек пересечения штриховой линии $\tilde{m}_\nu^2 = -0,67 \text{ эВ}^2$ с границами доверительного пояса: $0 \leq m_\nu^2 \leq 4,96 \text{ эВ}^2$

нии рецепта Фельдмана и Казинса такое сравнение некорректно, более того, оно может привести к заведомо неправильной интерпретации результатов экспериментов.

9. ВЫВОДЫ

Таким образом, задача об учете априорной информации о параметрах при построении доверительных интервалов допускает полноценное, физически корректное решение, которое мы называем методом предела чувствительности, так как оно уточняет известный рецепт предела чувствительности.

Это решение имеет прозрачное и потому убедительное обоснование и оказывается простым и удобным для практического применения (см. аналитические описания систем доверительных интервалов в разд. 5, 6 и 7, а также рис. 6, 8, 14 и 15, представляющие эти рецепты в удобном графическом виде).

Перечислим важнейшие особенности рассмотренной конструкции.

Непрерывные распределения. Система доверительных интервалов, построенная на основе выбора оценки с учетом априорной информации об ограниченном параметре непрерывного распределения, обладает следующими примечательными свойствами (как для симметричного интервала (рис. 6, разд. 5), так и для наилучшего верхнего предела (рис. 14, п. 7.1)):

— оценка устойчива для нефизических значений оценки параметра, т. е. для $\hat{\theta} < 0$;

— верхняя граница интервала для физических значений $\hat{\theta}$ та же, что и для немодифицированного случая, и является наименьшей возможной для данного доверительного уровня β (для наилучшего верхнего предела). В случае симметричного интервала для физических значений $\hat{\theta}$ верхняя граница всегда лежит ниже верхней границы для построения [6];

— нижняя граница интервала отходит от нуля в точке с наименьшей возможной (для данного доверительного уровня) абсциссой $\hat{\theta}_C$ (рис. 6 и 14), кроме того, нижняя граница является максимально возможной для данного доверительного уровня на отрезках $\hat{\theta}_C \leq \hat{\theta} \leq \hat{\theta}_E$ (рис. 6), $\hat{\theta}_C \leq \hat{\theta} \leq \hat{\theta}_D$ (рис. 14);

— для вычисления доверительного интервала не требуется сложных алгоритмов или таблиц помимо стандартных процедур построения доверительного интервала для заданного доверительного уровня.

Дискретные распределения. В результате учета информации о фоне по методу предела чувствительности строится система доверительных интервалов (рис. 8, 15, разд. 6, п. 7.2), обладающих следующими свойствами:

— решается указанная в [6] проблема меньшего числа зарегистрированных событий, чем ожидаемый фон (в области, где число событий меньше фона, построенная система доверительных интервалов естественным образом

дает оценку сверху для измеряемого параметра, не зависящую от значения n , как в области $n \leq 3$ на рис. 8);

— нижняя граница интервала имеет участок, аналогичный участку СЕФ на рис. 6; благодаря учету ограничения на параметр отрыв нижней границы доверительного интервала от оси n происходит раньше;

— благодаря неоднозначности определения интервалов для дискретных распределений (условие (10)) возможны различные варианты построения доверительных интервалов (фиксация нижней или верхней границы, построение Стерна, Кроу и Гарднера, их комбинации). Выбор может зависеть от конкретной ситуации, например, при необходимости можно получить более строгое ограничение на параметры сверху или снизу.

Наилучший верхний предел для неизвестного сигнала μ при наличии пуассоновского фона со средним b с учетом априорной информации о фоне в виде оценки (20) обладает следующими свойствами:

— при измерении в эксперименте любого числа событий $n \leq b$ получим одинаковый доверительный интервал. Это, в частности, решает проблему меньшего, чем фон, числа событий, обозначенную в [6];

— вероятностное содержание доверительной области близко к требуемому доверительному уровню, перекрывание мало (для дискретных распределений полностью его избежать не удастся). Перекрывание (в понимании [6]) является недостатком некоторых других построений [19, 7];

— нижняя граница доверительного интервала отходит от оси n (становится отличной от нуля) при наименьшем возможном значении n . Таким образом, не только указывается наилучший (наиболее сильный) верхний предел для параметра, но и появляется возможность детектировать сигнал (отличная от нуля нижняя граница в некотором смысле обозначает наличие сигнала) при достаточно малых значениях измеренного числа событий;

— как и в непрерывном случае, область под нижней границей является выигрышем от использования априорной информации о фоне.

Следует помнить, что суть и смысл обработки данных состоит в том, чтобы сама величина получаемых чисел непосредственно отражала ту информацию, которую требуется извлечь из данных посредством обработки. В этом смысле методы Фельдмана и Казинса и Кауэна и др. просто не достигают цели, можно сказать, шифруя и тем самым скрывая некоторую существенную часть информации. Метод предела чувствительности лишен этого недостатка и эффективно решает задачу построения доверительных интервалов с учетом априорной информации о параметрах.

Благодарности. Авторы благодарят сотрудников отдела экспериментальной физики ИЯИ РАН за обсуждение результатов работы. Отдельная благодарность А. И. Берлеву за постоянный интерес к работе. Авторы также благодарят А. С. Барабаша за критические замечания по программному обеспечению.

Работа была поддержана грантом Российского фонда фундаментальных исследований 14-02-31055 и грантом НШ-2835.2014.2.

СПИСОК ЛИТЕРАТУРЫ

1. *Neyman J.* Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability // *Phil. Trans. Roy. Soc. London. A.* 1937. V. 236. P. 333–380.
2. *Aseev V.N. et al.* // *Phys. Rev. D.* 2011. V. 84. P. 112003.
3. *Klapdor-Kleingrothaus H.V. et al.* // *Eur. Phys. J. A.* 2001. V. 12. P. 147;
Andreotti E. et al. // *Astropart. Phys.* 2011. V. 34. P. 822; arXiv:1012.3266;
Arnold R. et al. // *Nucl. Phys. A.* 2006. V. 765. P. 483; arXiv:hep-ex/0601021;
Barabash A. et al. // *Phys. At. Nucl.* 2011. V. 74. P. 312; arXiv:1002.2862.
4. *Abe K. et al.* // *Phys. Rev. Lett.* 2011. V. 107. P. 041801; arXiv:1106.2822;
Adamson P. et al. // *Phys. Rev. Lett.* 2011. V. 107. P. 181802; arXiv:1108.0015.
5. *Sterne T.E.* // *Biometrika.* 1954. V. 41. P. 275;
Crow E.L. // *Biometrika.* 1956. V. 43. P. 423;
Crow E.L., Gardner R.S. // *Biometrika.* 1959. V. 46. P. 441.
6. *Feldman G.J., Cousins R.D.* // *Phys. Rev. D.* 1998. V. 57. P. 3873.
7. *Cowan G. et al.* Power-Constrained Limits. arXiv:1105.3166.
8. *Mandelkern M., Schultz J.* The Statistical Analysis of Gaussian and Poisson Signals near Physical Boundaries // *J. Math. Phys.* 2000. V. 41. P. 5701–5709; arXiv:hep-ex/9910041v3.
9. *Tkachov F.* Optimal Confidence Intervals for Bounded Parameters. arXiv:0911.4271.
10. *Kraus Ch. et al.* // *Eur. Phys. J. C.* 2005. V. 40. P. 447.
11. *Lobashev V.M.* // *Prog. Part. Nucl. Phys.* 2002. V. 48. P. 123; *Nucl. Phys. A.* 2003. V. 719. P. 153;
Lobashev V.M. et al. // *Nucl. Phys. B. Proc. Suppl.* 2001. V. 91. P. 280.
12. *Tkachov F.* Optimal Upper Bounds for Non-Negative Parameters. arXiv:0912.1555.
13. *Тяпкин А. А.* По поводу трактовки основных проблем теории оценок. Статистические методы в экспериментальной физике. М.: Атомиздат, 1976. 335 с.
14. *Read A.L.* Modified Frequentist Analysis of Search Results (the CLs Method) // *Proc. of the Workshop on Confidence Limits / Eds.: James F., Lyons L., Perrin Y.* CERN, 2000. P. 81.
15. *The ATLAS Collab.* // *Phys. Lett. B.* 2012. V. 716. P. 1;
The CMS Collab. // *Ibid.* P. 30.
16. *Eadie W.T. et al.* *Statistical Methods in Experimental Physics.* North-Holland, 1971.
17. *Tkachov F.V.* Transcending the Least Squares. arXiv:physics/0604127.
18. *Giunti C.* // *Phys. Rev. D.* 1999. V. 59. P. 053001.
19. *Cousins R.D.* Negatively Biased Relevant Subsets Induced by the Most-Powerful One-Sided Upper Confidence Limits for a Bounded Physical Parameter. arXiv:1109.2023.
20. <http://www.inr.ac.ru/~blackbox/stat/intervals/>