Занятие 2. Методы построения оценок.

Домашнее задание.

Обязательные – глава 19, задачи 119, 120, 122, 124, 126, 127, 129, 131, 132, 133, 134.

Дополнительные – глава 19, задача 105(*).

Задача 2.1.

Пусть $(\xi_1,..., \xi_n)$ – выборка из показательного распределения $E(a)$. Построить оценки неизвестного параметра a методами:

1) моментов,

2) максимального правдоподобия,

3) порядковых статистик.

Решение:

1) Для построения оценки по методу моментов необходимо получить соотношение, в котором параметр a был бы выражен через момент (или моменты) показательного распределения $E(a)$. Легко видеть, что если случайная величина ξ имеет показательное распределение $E(a)$, то математическое ожидание ξ:

$$M_a[\xi] = \frac{1}{a},$$

Отсюда

$$a = \frac{1}{M_a[\xi]}.$$

Теперь для построения оценки параметра a достаточно в выражении справа от равенства заменить математическое ожидание $M_a[\xi]$ его оценкой, построенной по выборке $(\xi_1,..., \xi_n)$. Такой оценкой вполне может быть, например, выборочное среднее $\hat{m}_1(\xi_1,..., \xi_n)$:

$$\hat{m}_1(\xi_1,..., \xi_n) = \frac{1}{n}\sum_{i=1}^{n} \xi_i .$$

Таким образом, в качестве оценки по методу моментов может быть выбрана оценка $a_1(\xi_1,..., \xi_n)$:

$$a_1(\xi_1,..., \xi_n) = \frac{1}{\hat{m}_1(\xi_1,..., \xi_n)} = \frac{1}{\frac{1}{n}\sum_{i=1}^{n} \xi_i}.$$

2) Для построения оценки по методу максимального правдоподобия необходимо прежде всего построить функцию правдоподобия. В данном случае, совокупность наблюдений представляет собой выборку из показательного распределения $E(a)$, поэтому все случайны величины ξ_n являются независимыми и имеют одинаковую плотность вероятности $f_\xi(x \mid a)$:

$$f_\xi(x \mid a) = ae^{-ax}.$$

Отсюда следует, что совместное распределение случайных величин ξ_1, …, ξ_n является произведением плотностей вероятностей $f_\xi(x \mid a)$:

$$L(x_1,..., x_n \mid a) = \prod_{i=1}^{n} f_\xi(x \mid a) = \prod_{i=1}^{n} ae^{-ax} = a^n e^{-a\sum_{i=1}^{n} x_i}.$$

Функцией правдоподобия является функция $L(\xi_1,..., \xi_n \mid a)$:

$$L(\xi_1,..., \xi_n \mid a) = a^n e^{-a\sum_{i=1}^{n} \xi_i},$$

в которой величины ξ_i считаются заданными, а параметр a считается переменным. Суть метода максимального правдоподобия заключается в том, чтобы для каждого набора значений величин ξ_1, …, ξ_n находить такое значение параметра a^* (которое в общем случае, конечно, зависит от значений величин ξ_1, …, ξ_n) при котором функция правдоподобия $L(\xi_1,\ldots,\ \xi_n \mid a)$ принимает наибольшее значение.

В данном случае, как и во многих других, наиболее просто выражение для значений a^* может быть получено в результате перехода к логарифму функции правдоподобия:

$$\ln\ L(\xi_1,\ldots,\xi_n \mid a) = n \ln\ a - a\sum_{i=1}^{n} \xi_i\ ,$$

и выписыванию необходимого условия экстремума (нормального уравнения):

$$\left.\frac{\partial \ln\ L(\xi_1,\ldots,\xi_n \mid a)}{\partial a}\right|_{a=a*} = 0\ ,$$

$$\frac{n}{a^*} - \sum_{i=1}^{n} \xi_i = 0\ ,$$

$$a^* = \frac{1}{\frac{1}{n}\sum_{i=1}^{n} \xi_i}\ .$$

Таким образом, оценкой по методу максимального правдоподобия является статистика $a_2(\xi_1,\ldots,\ \xi_n)$:

$$a_2(\xi_1,\ldots,\xi_n) = \frac{1}{\frac{1}{n}\sum_{i=1}^{n} \xi_i}\ .$$

3) Для построения оценки методом порядковых статистик необходимо выразить параметр a через квантили показательного распределения $E(a)$.

В общем, случае весьма затруднительно указать какие квантили нужно выбрать, чтобы получить заведомо «самую лучшую» оценку. Тем не менее, если не ставить перед собой задачу получения «самой лучшей» оценки, то достаточным бывает использование таких квантилей, через которые параметр выражается некоторым «простым» образом.

В данном случае, например, параметр a весьма просто выражается через медиану $x_{1/2}$ показательного распределения $E(a)$. Если функция $F_\xi(x \mid a)$ является функцией распределения $E(a)$:

$$F_\xi(x \mid a) = 1 - e^{-ax}\ ,$$

то медиана $x_{1/2}$ по определению является таким значением аргумента x функции $F_\xi(x \mid a)$, при котором значение функции равно $\frac{1}{2}$:

$$F_\xi(x_{1/2} \mid a) = \frac{1}{2}\ ,$$

$$1 - e^{-ax_{1/2}} = \frac{1}{2}\ ,$$

$$a = \frac{\ln\ 2}{x_{1/2}}\ .$$

Теперь для получения оценки необходимо заменить точное значение медианы $x_{1/2}$ её оценкой, построенной по выборке, где в качестве оценки, например, может быть использована соответствующая порядковая статистика $\xi_{\left(\left[\frac{n}{2}\right]+1\right)}$.

Таким образом, окончательно оценка по методу порядковых статистик $a_3(\xi_1,..., \xi_n)$ имеет следующий вид:

$$a_3(\xi_1,..., \xi_n) = \frac{\ln 2}{\xi_{\left(\left[\frac{n}{2}\right]+1\right)}}.$$

Ответ:

1) $\dfrac{1}{\frac{1}{n}\sum\limits_{i=1}^{n}\xi_i}$,

2) $\dfrac{1}{\frac{1}{n}\sum\limits_{i=1}^{n}\xi_i}$,

3) $\dfrac{\ln 2}{\xi_{\left(\left[\frac{n}{2}\right]+1\right)}}$.

Задача 2.2.

Из партии приборов наугад выбирают n приборов и ставят на испытания. Времена работы приборов до первого отказа оказываются равными ξ_1, …, ξ_n. Оценить вероятность того, что время работы прибора в партии не меньше заданного значения t, считая, что времена работы всех приборов в партии до первого отказа независимы и имеют показательное распределение $E(a)$, где a неизвестный параметр.

Решение:

Прежде всего, необходимо выяснить каким образом искомая вероятность зависит от неизвестного параметра a. Представим, что некоторая случайная величина ξ, характеризующая время работы одного прибора, имеет показательное распределение $E(a)$ и выразим неизвестную вероятность P_t события $\{\xi \geq t\}$ через параметр a:

$$P_t = P\{\xi \geq t\} = 1 - F_\xi(t \mid a) = 1 - (1 - e^{-at}) = e^{-at}.$$

Легко видеть, что оценку неизвестной вероятности P_t можно получить, если заменить в полученном выражении неизвестный параметр a, его оценкой, полученной на основе измерений ξ_1, …, ξ_n. По условию задачи измерения ξ_1, …, ξ_n являются независимыми и одинаково распределенными, поэтому совокупность величин $(\xi_1,..., \xi_n)$ можно считать выборкой, откуда в качестве оценки параметра a может быть использована любая из оценок, полученных при решении задачи 3.1. Например, используя оценку по методу максимального правдоподобия, получим следующую оценку искомой вероятности:

$$\hat{P}(\xi_1,..., \xi_n) = e^{-\frac{1}{\frac{1}{n}\sum_{i=1}^{n}\xi_i}t}.$$

Ответ:

$e^{-\frac{1}{\frac{1}{n}\sum_{i=1}^{n}\xi_i}t}$.

Задача 2.3.

Проведено тестирование n приборов, в котором каждый прибор испытывался до первого отказа. По результатам тестирования были получены данные о том, что прибор с номером i $(i=\overline{1,n})$ отказал в испытании с номером ξ_i. Считая, что испытания всех приборов проводятся независимо, на основании полученных данных построить оценку максимального правдоподобия вероятности отказа прибора в одном испытании p, $p \in (0,1)$.

Решение:

1) Поскольку прибор в одном испытании отказывает с вероятностью p и все испытания по условию являются независимыми, то все величины ξ_i независимы в совокупности и имеют одинаковое геометрическое распределение:

$$P\{\xi_i = k \mid p\} = (1-p)^{k-1} p\ ,\ k = 1,2,\ldots\ .$$

2) Считая, что совокупность величин $(\xi_1,\ldots,\xi_n)$ является выборкой, построим функцию совместного распределения $L(k_1,\ldots,k_n \mid p)$ величин ξ_1, …, ξ_n, которая в данном случае является произведением вероятностей (в силу независимости) отдельных геометрических распределений:

$$L(k_1,\ldots,k_n \mid p) = \prod_{i=1}^{n} P\{\xi_i = k_i \mid p\} = \prod_{i=1}^{n} (1-p)^{k_i-1} p\ .$$

С целью определения значений параметра p, при которых достигается наибольшее значение функции правдоподобия $L(\xi_1,\ldots,\xi_n \mid p)$, удобней перейти к логарифму от функции правдоподобия:

$$\ln L(\xi_1,\ldots,\xi_n \mid p) = \sum_{i=1}^{n} \left\{(\xi_i - 1)\ln(1-p) + \ln p\right\}.$$

Необходимое условие экстремума (нормальное уравнение) в данном случае имеет следующий вид:

$$\left.\frac{\partial \ln L(\xi_1,\ldots,\xi_n \mid p)}{\partial p}\right|_{p=p^*} = 0\ ,$$

$$\sum_{i=1}^{n} \left\{(\xi_i - 1)\frac{(-1)}{1-p^*} + \frac{1}{p^*}\right\} = 0\ ,$$

$$\frac{-\sum\limits_{i=1}^{n}(\xi_i - 1)}{1-p^*} + \frac{n}{p^*} = 0\ ,$$

$$\frac{-p^*\sum\limits_{i=1}^{n}(\xi_i - 1) + n(1-p^*)}{(1-p^*)p^*} = 0\ ,$$

$$-p^*\sum_{i=1}^{n}\xi_i + np^* + n - np^* = 0\ ,$$

$$p^* = \frac{1}{\frac{1}{n}\sum\limits_{i=1}^{n}\xi_i}\ .$$

Таким образом, оценка максимального правдоподобия параметра p имеет вид:

$$p^*(\xi_1,\ldots,\xi_n) = \frac{1}{\frac{1}{n}\sum\limits_{i=1}^{n}\xi_i}$$

Ответ:

$$\frac{1}{\frac{1}{n}\sum_{i=1}^{n}\xi_i}.$$

Задача 2.4.

Пусть $(\xi_1,\ldots,\xi_n)$ – выборка из нормального распределения $N(m,\sigma^2)$, где m и σ неизвестные параметры. Построить оценки параметров m и σ методом порядковых статистик.

Решение:

1) Известно, что в силу симметрии для всякого нормального распределения $N(m,\sigma^2)$ величины математического ожидания m и медианы $x_{1/2}$ совпадают:

$$m = x_{1/2}.$$

Откуда заменой медианы $x_{1/2}$ её оценкой, например порядковой статистикой $\xi_{\left(\left[\frac{n}{2}\right]+1\right)}$, получим оценку для математического ожидания следующего вида:

$$\hat{m}(\xi_1,\ldots,\xi_n) = \xi_{\left(\left[\frac{n}{2}\right]+1\right)}.$$

2) Теперь получим оценку для среднеквадратичного отклонения σ, для этого представим выражения для нижней и верхней квартилей $x_{1/4}$ и $x_{3/4}$:

$$\begin{cases} \Phi\left(\dfrac{x_{1/4}-m}{\sigma}\right) = \dfrac{1}{4} \\ \Phi\left(\dfrac{x_{3/4}-m}{\sigma}\right) = \dfrac{3}{4} \end{cases}, \quad \begin{cases} x_{1/4} = m + \sigma\Phi^{-1}\left(\dfrac{1}{4}\right) \\ x_{3/4} = m + \sigma\Phi^{-1}\left(\dfrac{3}{4}\right) \end{cases}.$$

Теперь остается лишь исключить математическое ожидание m, например, вычитая из второго уравнения первое, получим следующее выражение для параметра σ:

$$\sigma = \frac{x_{3/4}-x_{1/4}}{\Phi^{-1}\left(\frac{3}{4}\right)-\Phi^{-1}\left(\frac{1}{4}\right)}.$$

Заменяя в полученном выражении квартили $x_{1/4}$ и $x_{3/4}$ их оценками $\xi_{\left(\left[\frac{1}{4}n\right]+1\right)}$ и $\xi_{\left(\left[\frac{3}{4}n\right]+1\right)}$, получим оценку для параметра σ следующего вида:

$$\hat{\sigma}(\xi_1,\ldots,\xi_n) = \frac{\xi_{\left(\left[\frac{3}{4}n\right]+1\right)}-\xi_{\left(\left[\frac{1}{4}n\right]+1\right)}}{\Phi^{-1}\left(\frac{3}{4}\right)-\Phi^{-1}\left(\frac{1}{4}\right)}.$$

Ответ:

Оценка параметра m – $\xi_{\left(\left[\frac{n}{2}\right]+1\right)}$, оценка параметра σ – $\dfrac{\xi_{\left(\left[\frac{3}{4}n\right]+1\right)}-\xi_{\left(\left[\frac{1}{4}n\right]+1\right)}}{\Phi^{-1}\left(\frac{3}{4}\right)-\Phi^{-1}\left(\frac{1}{4}\right)}$.

Задача 2.5.

Пусть $(\xi_1,\ldots,\xi_n)$ – выборка из равномерного распределения $R[a,b]$. Построить оценки неизвестного двумерного параметра (a,b) методами:

1) моментов,

2) максимального правдоподобия.

Решение:

1) Пусть случайная величина ξ имеет равномерное распределение $R[a,b]$, тогда математическое ожидание $M_{(a,b)}[\xi]$ и дисперсия $D_{(a,b)}[\xi]$ случайной величины ξ выражаются через параметр (a,b) следующим образом:

$$\begin{cases} M_{(a,b)}[\xi] = \dfrac{b+a}{2} \\ D_{(a,b)}[\xi] = \dfrac{(b-a)^2}{12} \end{cases}.$$

Разрешая полученную систему относительно a и b, в результате преобразования,

$$\begin{cases} 2M_{(a,b)}[\xi] = b+a \\ \sqrt{12\,D_{(a,b)}[\xi]} = b-a \end{cases}$$

Окончательно получим следующую систему:

$$\begin{cases} a = M_{(a,b)}[\xi] - \sqrt{3D_{(a,b)}[\xi]} \\ b = M_{(a,b)}[\xi] + \sqrt{3D_{(a,b)}[\xi]} \end{cases}.$$

Заменяя в полученной системе математическое ожидание $M_{(a,b)}[\xi]$ выборочным средним $\hat{m}_1(\xi_1,..., \xi_n)$, а дисперсию $D_{(a,b)}[\xi]$ исправленной выборочной дисперсией $\tilde{\mu}_2(\xi_1,..., \xi_n)$, получим оценку по методу моментов параметра (a,b):

$$\begin{pmatrix} a_1(\xi_1,..., \xi_n) \\ b_1(\xi_1,..., \xi_n) \end{pmatrix} = \begin{pmatrix} \hat{m}_1(\xi_1,..., \xi_n) - \sqrt{3\tilde{\mu}_2(\xi_1,..., \xi_n)} \\ \hat{m}_1(\xi_1,..., \xi_n) + \sqrt{3\tilde{\mu}_2(\xi_1,..., \xi_n)} \end{pmatrix},$$

где $\hat{m}_1(\xi_1,..., \xi_n) = \dfrac{1}{n}\sum_{i=1}^{n} \xi_i$, $\tilde{\mu}_2(\xi_1,..., \xi_n) = \dfrac{1}{(n-1)}\sum_{i=1}^{n} \left(\xi_i - \hat{m}_1(\xi_1,..., \xi_n)\right)^2$.

2) Пусть случайная величина ξ имеет равномерное распределение $R[a,b]$, тогда плотность вероятности $f_\xi(x \mid a,b)$ величины ξ имеет следующий вид:

$$f_\xi(x \mid a,b) = \begin{cases} \dfrac{1}{b-a} & , a \le x \le b \\ 0 & \textit{иначе} \end{cases}.$$

Функцию плотности вероятности $f_\xi(x \mid a,b)$ можно также представить и в следующем виде:

$$f_\xi(x \mid a,b) = I(x-a)\frac{1}{b-a}I(b-x),$$

где функция $I(t)$ играет роль индикаторной функции:

$$I(t) = \begin{cases} 1 & , t \ge 0 \\ 0 & , t < 0 \end{cases}.$$

Совместная плотность вероятности совокупности величин $(\xi_1,..., \xi_n)$, являющейся выборкой из распределения $R[a,b]$, являет произведением функций плотности вероятности:

$$L(x_1,..., x_n \mid a,b) = \prod_{i=1}^{n} f_\xi(x \mid a,b) = \prod_{i=1}^{n} I(x_i - a)\frac{1}{b-a}I(b-x_i) =$$

$$= \prod_{i=1}^{n} I(x_i - a)\cdot\left(\frac{1}{b-a}\right)^n \cdot \prod_{i=1}^{n} I(b-x_i) = I\left(\min_{1\le i\le n} x_i - a\right)\left(\frac{1}{b-a}\right)^n I\left(b - \max_{1\le i\le n} x_i\right).$$

Таким образом, функция правдоподобия $L(\xi_1,..., \xi_n \mid a,b)$ принимает вид:

$$L(\xi_1,\dots,\xi_n \mid a,b) = I\left(\min_{1\le i\le n} \xi_i - a\right)\left(\frac{1}{b-a}\right)^n I\left(b - \max_{1\le i\le n} \xi_i\right).$$

При фиксированных величинах ξ_1, …, ξ_n функция правдоподобия $L(\xi_1,\dots,\xi_n \mid a,b)$ принимает наибольшее значение, когда центральный множитель $\left(\frac{1}{b-a}\right)^n$ принимает наибольшее значение, то есть разность $b-a$ должна быть как можно меньше, и при этом $\min\limits_{1\le i\le n} \xi_i - a \ge 0$ и $b - \max\limits_{1\le i\le n} \xi_i \ge 0$ поскольку в противном случае значение функции правдоподобия сразу становится равным нулю. Следовательно, требуется таким образом выбрать значения a^* и b^* для параметров a и b, чтобы:

$$b - a \to \min,$$

$$a \le \min_{1\le i\le n} \xi_i,\ \max_{1\le i\le n} \xi_i \le b.$$

Очевидно, что такими значениями являются величины $a^* = \min\limits_{1\le i\le n} \xi_i$ и $b^* = \max\limits_{1\le i\le n} \xi_i$. Таким образом, оценка по методу максимального правдоподобия имеет вид:

$$\begin{pmatrix} a_2(\xi_1,\dots,\xi_n) \\ b_2(\xi_1,\dots,\xi_n) \end{pmatrix} = \begin{pmatrix} \min\limits_{1\le i\le n} \xi_i \\ \max\limits_{1\le i\le n} \xi_i \end{pmatrix} = \begin{pmatrix} \xi_{(1)} \\ \xi_{(n)} \end{pmatrix}.$$

Ответ:

1) $\begin{pmatrix} \hat{m}_1(\xi_1,\dots,\xi_n) - \sqrt{3\tilde{\mu}_2(\xi_1,\dots,\xi_n)} \\ \hat{m}_1(\xi_1,\dots,\xi_n) + \sqrt{3\tilde{\mu}_2(\xi_1,\dots,\xi_n)} \end{pmatrix}$,

$\hat{m}_1(\xi_1,\dots,\xi_n) = \frac{1}{n}\sum\limits_{i=1}^{n} \xi_i$, $\tilde{\mu}_2(\xi_1,\dots,\xi_n) = \frac{1}{(n-1)}\sum\limits_{i=1}^{n} \left(\xi_i - \hat{m}_1(\xi_1,\dots,\xi_n)\right)^2$;

2) $\begin{pmatrix} \xi_{(1)} \\ \xi_{(n)} \end{pmatrix}$.

Задача 2.6.

Серия наблюдений состоит всего лишь из одной случайной величины ξ_1, имеющей биномиальное распределение $Bi(n,p)$ с известным параметром n и неизвестной вероятностью p. Построить несмещенную оценку дисперсии $D_p[\xi_1]$.

Решение:

Дисперсия $D_p[\xi_1] = np(1-p)$, поэтому построив оценку $\hat{p}(\xi_1)$ вероятности p можно получить оценку $\hat{d}(\xi_1)$ дисперсии $D_p[\xi_1]$:

$$\hat{d}(\xi_1) = n\hat{p}(\xi_1)(1-\hat{p}(\xi_1)).$$

Оценку $\hat{p}(\xi_1)$ вероятности p можно построить по методу моментов:

$$\hat{p}(\xi_1) = \frac{\xi_1}{n}.$$

Тогда оценка $\hat{d}(\xi_1)$ примет вид:

$$\hat{d}(\xi_1) = n\frac{\xi_1}{n}\left(1 - \frac{\xi_1}{n}\right) = \xi_1\left(1 - \frac{\xi_1}{n}\right).$$

Для того чтобы оценка $\hat{d}(\xi_1)$ была несмещенной оценкой дисперсии $D_p[\xi_1]$ согласно определению несмещенной оценки достаточно, чтобы:

$$\forall p:\ D_p[\hat{d}(\xi_1)] = D_p[\xi_1] = np(1-p).$$

Вычислим математическое ожидание $M_p[\hat{d}(\xi_1)]$:

$$M_p[\hat{d}(\xi_1)] = M_p\left[\xi_1\left(1-\frac{\xi_1}{n}\right)\right] = M_p\left[\xi_1 - \frac{\xi_1^2}{n}\right] = M_p[\xi_1] - \frac{1}{n}M_p[\xi_1^2] = np - \frac{1}{n}M_p[\xi_1^2] =$$

$$= np - \frac{1}{n}\left(D_p[\xi_1^2] + \left(M_p[\xi_1]\right)^2\right) = np - \frac{1}{n}(np(1-p) + n^2p^2) = np - p + p^2 - np^2 =$$

$$= p(n-1) - p^2(n-1) = (n-1)p(1-p)$$

Легко видеть, что оценка $\hat{d}(\xi_1)$ не является несмещенной оценкой $D_p[\xi_1]$, поскольку:

$$M_p[\hat{d}(\xi_1)] \neq D_p[\xi_1] = np(1-p).$$

Тем не менее, смещение оценки $\hat{d}(\xi_1)$ легко исправить, образовав новую оценку $\tilde{d}(\xi_1)$:

$$\tilde{d}(\xi_1) = \frac{n}{n-1}\hat{d}(\xi_1) = \frac{n}{n-1}\xi_1\left(1-\frac{\xi_1}{n}\right),$$

которая будет несмещенной оценкой $D_p[\xi_1]$. Действительно,

$$\forall p:\ M_p[\hat{d}(\xi_1)] = \frac{n}{n-1}M_p[\hat{d}(\xi_1)] = \frac{n}{n-1}(n-1)p(1-p) = np(1-p) = D_p[\xi_1],$$

что по определению означает несмещенность оценки $\tilde{d}(\xi_1)$.

Ответ:

$$\frac{n}{n-1}\xi_1\left(1-\frac{\xi_1}{n}\right)$$

Задача 2.7.

Некоторый источник с координатами $(1, a)$ излучает n частиц, которые засвечиваются на экране $y = 0$ в точках с координатами $(\xi_i, 0)$. Построить оценку неизвестного значения a, считая, что направления излучения частиц независимы, и что источник испускает частицы одинаково во всех направлениях.

(Рисунок).

Решение:

1) Прежде всего, необходимо установить распределение координат по оси абсцисс ξ_i для засвечиваемых на экране частиц.

Пусть величины α_i являются углами излучения частиц, засвечиваемых на экране. Примем направление отсчета углов таким образом, чтобы все $\alpha_i \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. Поскольку по условию направления излучения независимы и частицы излучаются одинаково во всех направлениях, то можно считать, что все углы α_i являются случайными величинами независимыми в совокупности, и что все α_i имеют равномерное распределение $R\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ с плотностью вероятности $p_\alpha(y)$:

$$p_\alpha(y) = \frac{1}{\pi}.$$

В принятых условиях координаты по оси абсцисс ξ_i определяются через углы следующим соотношением:

$$\xi_i = a + tg\,\alpha_i.$$

Обратное преобразование от координат ξ_i к углам α_i имеет вид:

$$\alpha_i = arctg\ (\xi_i - a),$$
$$f(x \mid a) = arctg\ (x - a),$$

поэтому плотность вероятности $p_\xi(x \mid a)$ всех случайных величин ξ_i с учетом равномерного распределения величин α_i имеет вид:

$$p_\xi(x \mid a) = p_\alpha(f^{-1}(x \mid a)) \cdot \left| \frac{d}{dx} f^{-1}(x \mid a) \right| = \frac{1}{\pi} \frac{1}{1 + (x - a)^2}.$$

Таким образом, все случайные величины ξ_i имеют одинаковое распределение Коши и являются независимыми в совокупности (в силу независимости α_i), то есть совокупность величин $(\xi_1, ..., \xi_n)$ является выборкой из распределения Коши с плотностью $p_\xi(x \mid a)$, где a – неизвестный параметр, и требуется построить оценку параметра a.

2) Попробуем воспользоваться методом моментов: для этого определим подходящий момент распределения Коши с тем, чтобы связать момент с параметром a. К сожалению, у распределения Коши нет ни одного момента, поэтому воспользоваться методом моментов для построения оценки в данном случае невозможно.

3) Попробуем воспользоваться методом максимального правдоподобия: для этого определим функцию правдоподобия $L(\xi_1, ..., \xi_n \mid a)$ (в данном случае она будет равна произведению плотностей вероятностей, поскольку $(\xi_1, ..., \xi_n)$ является выборкой):

$$L(\xi_1, ..., \xi_n \mid a) = \left(\frac{1}{\pi} \right)^n \prod_{i=1}^{n} \frac{1}{1 + (x_i - a)^2}.$$

Для определения оценки по методу максимального правдоподобия $a^*(\xi_1, ..., \xi_n)$ перейдем к логарифму функции правдоподобия:

$$\ln L(\xi_1, ..., \xi_n) = -n \ln \pi - \sum_{i=1}^{n} \ln(1 + (\xi_i - a)^2),$$

и составим нормальное уравнение (необходимое условие экстремума):

$$\left. \frac{\partial \ln L(\xi_1, ..., \xi_n)}{\partial a} \right|_{a = a^*} = 0,$$

$$\sum_{i=1}^{n} \frac{2(\xi_i - a^*)}{1 + (\xi_i - a^*)^2} = 0.$$

После приведения к общему знаменателю в числителе окажется многочлен от величин ξ_i степени $2(n-1) + 1 = 2n - 1$, поэтому найти выражения для величины a^* окажется весьма затруднительно.

Таким образом, методом максимального правдоподобия в данном случае оказывается сложным и приводит к весьма трудной задаче.

4) Попробуем воспользоваться методом порядковых статистик: легко видеть, что в силу симметричности плотности вероятности $p_\xi(x \mid a)$ относительно прямой $x = a$ медиана $x_{1/2}$ совпадает с параметром a:

$$a = x_{1/2}.$$

Отсюда в качестве оценки параметра a может быть использована выборочная медиана, в качестве которой может быть использована порядковая статистика $\xi_{\left(\left[\frac{n}{2}\right]+1\right)}$:

$$\hat{a}(\xi_1, ..., \xi_n) = \xi_{\left(\left[\frac{n}{2}\right]+1\right)}.$$

Ответ:

$$\xi_{\left(\left[\frac{n}{2}\right]+1\right)}.$$