EESTI NSV TEADUSTE AKADEEMIA TOIMETISED. 32. KÖIDE
FÜÜSIKA * MATEMAATIKA. 1983, NR. 3

ИЗВЕСТИЯ АКАДЕМИИ НАУК ЭСТОНСКОЙ ССР. ТОМ 32
ФИЗИКА * МАТЕМАТИКА. 1983, № 3

https://doi.org/10.3176/phys.math.1983.3.05

УДК 519.214.5

Я. КУКС

ПРИБЛИЖЕННО МИНИМАКСНАЯ ОЦЕНКА СРЕДНЕГО ОДНОМЕРНОГО НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ

(Представил Н. Алумяэ)

Введение

Пусть X — одномерная случайная величина, распределенная по нормальному закону $N(\mu, \sigma^2)$ с дисперсией $\sigma^2 = 1$ и средним μ, не превышающим по модулю известной положительной константы a, т. е. $\mu \in [-a, a]$. Рассмотрим проблему построения оценки среднего μ, минимизирующей максимум риска при квадратичной потере. Известно, что минимаксная оценка является байесовской оценкой относительно априорного распределения на интервале $[-a, a]$, при котором байесовский риск максимален. Тем самым построение минимаксной оценки сводится к нахождению априорного распределения, максимизирующего байесовский риск.

В [1] показано, что априорное распределение с максимальным байесовским риском, являющееся симметричным и единственным, при $a \leqslant 1{,}05$ сосредоточено в граничных точках интервала $[-a, a]$, а при $a \leqslant 2$ — в этих же точках и в нуле. Вероятности точек сосредоточения при каждом $a = 1{,}1\,(0{,}1)\,2$ найдены численным путем. В [2] построены оценки среднего μ, являющиеся при $a \to \infty$ асимптотически минимаксными второго порядка на интервале $[-a, a]$. Такие же оценки рассмотрены в [3].

В настоящей статье приводятся построенные автором численными методами априорные распределения на интервале $[-a, a]$ при $a = 1{,}1\,(0{,}1)\,4\,(0{,}2)\,6\,(0{,}4)\,8\,(1)\,10$, обладающие почти максимальным байесовским риском. Максимальный риск соответствующих байесовских оценок с точностью проведенных вычислений совпадает с байесовским риском. Поскольку максимальный риск не может быть меньше байесовского риска, то эти байесовские оценки являются приближенно минимаксными, и с точки зрения применений проблема построения минимаксных оценок решена для вышеуказанных значений a. Приводимые результаты до $a = 2$ совпадают с результатами [1], за исключением случая, когда $a = 1{,}5$. Есть основания надеяться, что построенные априорные распределения близки к наименее благоприятным не только в смысле байесовского риска, но и в смысле точек сосредоточения распределения и вероятностей этих точек.

Постановка задачи и известные предложения

Пусть $X \sim N(\mu, 1)$, где μ — число из заданного интервала $[-a, a]$. Для оценки $d(X)$ среднего μ определим функцию риска в виде

$$R(d, \mu) = (2\pi)^{-1/2} \int\limits_{-\infty}^{+\infty} [d(x) - \mu]^2 \, e^{-(x-\mu)^2/2} \, dx.$$

Задача заключается в построении минимаксной оценки $d_a^*(X)$, минимизирующей функционал

$$q_a(d) = \sup_{\mu \in [-a, a]} R(d, \mu).$$

Пусть Ξ_a — класс априорных вероятностных мер ξ, определенных на σ-алгебре борелевских подмножеств интервала $[-a, a]$, r_ξ — байесовский риск, соответствующий априорному распределению ξ, т. е.

$$r_\xi = \inf_d \int_{-a}^{a} R(d, \mu) \, d\xi(\mu),$$

а $d_\xi(X)$ — байесовская оценка относительно априорного распределения ξ, минимизирующая функционал $\int_{-a}^{a} R(d, \mu) \, d\xi(\mu)$. Пусть ξ^*_a — наименее благоприятное распределение в классе Ξ_a, т. е. такое, что

$$r_{\xi_a^*} = \sup_{\xi \in \Xi_a} r_\xi.$$

Приведем здесь некоторые известные предложения, имеющие отношение к построению минимаксной оценки.

Предложение 1. *Существует наименее благоприятное распределение в классе* Ξ_a.

Предложение 2. *Наименее благоприятное распределение* ξ_a^* *в классе* Ξ_a *сосредоточено на множестве точек интервала* $[-a, a]$, *где байесовская функция риска* $R(d_{\xi_a^*}, \cdot)$ *достигает максимального значения.*

Предложение 3. *Наименее благоприятное распределение в классе* Ξ_a *сосредоточено на конечном множестве точек.*

Предложение 4. *Наименее благоприятное распределение в классе* Ξ_a *присваивает равные вероятности симметричным относительно нуля точкам.*

Предложение 5. *Наименее благоприятное распределение в классе* Ξ_a *единственно.*

Предложение 6. *Существует минимаксная оценка* $d_a^*(X)$, *и она является байесовской оценкой относительно наименее благоприятного распределения в классе* Ξ_a.

Предложение 7. *Байесовская оценка относительно любого априорного распределения* ξ *из класса* Ξ_a *задается на всей прямой (исключая, быть может, множество лебеговской меры нуль) следующей формулой:*

$$d_\xi(x) = \frac{\int\limits^{a} \mu \, e^{-(x-\mu)^2/2} \, d\xi(\mu)}{\int\limits^{a} e^{-(x-\mu)^2/2} \, d\xi(\mu)}.$$

Предложение 8 ([4]). *Если* $\xi^1, \xi^2, \ldots$ — *такая последовательность априорных распределений в классе* Ξ_a, *что соответствующая последовательность байесовских рисков* $r_{\xi^1}, r_{\xi^2}, \ldots$ *сходится к максимальному байесовскому риску* $r_{\xi_a^*}$, *то последовательность байесовских функций риска* $R(d_{\xi^1}, \cdot), R(d_{\xi^2}, \cdot), \ldots$ *сходится равномерно на интервале* $[-a, a]$ *к единственной минимаксной функции риска* $R(d_\alpha^*, \cdot)$.

О процедуре вычислений

На выбор процедуры вычислений в настоящей работе значительно повлияли идеи работы [4], где для построения минимаксной оценки предложен итерационный метод. Однако стремление обойтись малым количеством вычислительных операций, а также условия работы с ЭВМ «Минск-32», имеющей небольшой объем оперативной памяти, привели к построению нового алгоритма для решения конкретной задачи. Итерационным методом ищется априорное распределение с «большим» байесовским риском, что с учетом предложения 8 дает надежду на «малость» максимума соответствующей байесовской функции риска $R(d_\xi, \cdot)$ на интервале $[-a, a]$. В силу предложений 3 и 4 было разумно построить процедуру поиска на симметричных распределениях, сосредоточенных на конечном множестве точек. Предложения 2 и 8 указывают на то, что при выполнении итераций следует точки сосредоточения распределения сдвигать в сторону возрастания байесовской функции риска, а вероятности этих точек выбирать так, чтобы байесовская функция риска принимала в них одно и то же значение. Последнего естественно добиваться увеличением априорных вероятностей в точках, где байесовская функция риска имеет большие значения.

При разработке алгоритма было сделано предположение, что если для некоторого a наименее благоприятное распределение ξ_a^* сосредоточено в точках $\pm c^*_{a1}, \ldots, \pm c^*_{am_a-1}, \pm a$, то существуют неотрицательные числа b_1 и b_2 такие, что при $0 \leqslant \Delta a \leqslant b_1$ распределение $\xi^*_{a+\Delta a}$ сосредоточено в точках

$$\pm[|c^*_{a1}|+\delta_{a1}(\Delta a)], \ldots, \pm[|c^*_{am_a-1}|+\delta_{am_a-1}(\Delta a)], \pm[a+\Delta a],$$

а при $b_1 < \Delta a \leqslant b_2$ — в точках

$$0, \pm[|c^*_{a1}|+\delta_{a1}(\Delta a)], \ldots, \pm[|c^*_{am_a-1}|+\delta_{am_a-1}(\Delta a)], \pm[a+\Delta a],$$

где $\delta_{a1}(\Delta a), \ldots, \delta_{am_a-1}(\Delta a)$ — непрерывные монотонно неубывающие функции. Другими словами, при расширении интервала $[-a, a]$ точки сосредоточения наименее благоприятного распределения непрерывно удаляются от нуля, а в нуле появляется новая точка сосредоточения распределения, которая затем раздваивается и т. д. Результаты вычислений подтвердили справедливость такого предположения. Таким образом, если при некотором $a = a_0$ находилось распределение ξ_{a_0} с байесовским риском, близким к максимальному, то при $a = a_0 + \Delta a$ в качестве начального приближения выбиралось распределение, сосредоточенное в точках, близких к тем, в которых сосредоточено распределение ξ_{a_0}. Аналогично выбирались и вероятности этих точек. Вычисления начинались с $a = 1$, когда наименее благоприятное распределение сосредоточено в точках -1 и $+1$, и продолжались до $a = 10$.

Результаты вычислений

Для каждого промежутка $[-a, a]$ при $a = 1{,}1\,(0{,}1)\,4\,(0{,}2)\,6\,(0{,}4)\,8\,(1)\,10$ определено распределение ξ_a с почти максимальным байесовским риском $r_{\xi_a} \approx r_{\xi^*_a}$, присваивающее точкам $\pm c_{a1}, \ldots, \pm c_{am_a}$ соответственно вероятности $\xi_{a1}/2, \ldots, \xi_{am_a}/2$, где $\sum_{j=1}^{m_a} \xi_{aj} = 1$. В таблице приведены для каждого a модули $+c_{a1}, \ldots, +c_{am_a}$ точек сосредоточения распределения ξ_a, двукратные вероятности этих точек $\xi_{a1}, \ldots, \xi_{am_a}$ и

Априорные распределения ξ_a с байесовским риском, близким к максимальному

a	$q_a(d_{\xi_a})$	$+c_{aj}$	ξ_{aj}
1,1	0,471	0	0,0499
		1,1	0,9501
1,2	0,492	0	0,1435
		1,2	0,8565
1,3	0,513	0	0,2145
		1,3	0,7855
1,4	0,535	0	0,2692
		1,4	0,7308
1,5	0,556	0	0,3117
		1,5	0,6883
1,6	0,577	0	0,3450
		1,6	0,6550
1,7	0,596	0	0,3711
		1,7	0,6289
1,8	0,615	0	0,3916
		1,8	0,6084
1,9	0,631	0	0,4077
		1,9	0,5923
2,0	0,645	0	0,4204
		2,0	0,5796
2,1	0,657	0,361	0,4585
		2,1	0,5415
2,2	0,670	0,502	0,4926
		2,2	0,5074
2,3	0,681	0,602	0,5221
		2,3	0,4779
2,4	0,693	0,681	0,5473
		2,4	0,4527
2,5	0,704	0,747	0,5687
		2,5	0,4313
2,6	0,715	0,804	0,5868
		2,6	0,4132
2,7	0,725	0,854	0,6020
		2,7	0,3980
2,8	0,734	0,901	0,6151
		2,8	0,3849
2,9	0,743	0	0,0040
		0,947	0,6222
		2,9	0,3738
3,0	0,751	0	0,1122
		1,121	0,5333
		3,0	0,3545
3,1	0,758	0	0,1679
		1,260	0,4952
		3,1	0,3369
3,2	0,766	0	0,2013
		1,375	0,4776
		3,2	0,3211
3,3	0,773	0	0,2242
		1,476	0,4690
		3,3	0,3068
3,4	0,780	0	0,2404
		1,565	0,4656
		3,4	0,2940

a	$q_a(d_{\xi_a})$	$+c_{aj}$	ξ_{aj}
3,5	0,787	0	0,2523
		1,645	0,4649
		3,5	0,2828
3,6	0,793	0	0,2615
		1,718	0,4656
		3,6	0,2729
3,7	0,799	0	0,2687
		1,786	0,4672
		3,7	0,2641
3,8	0,805	0,313	0,2974
		1,887	0,4483
		3,8	0,2543
3,9	0,810	0,487	0,3350
		2,008	0,4212
		3,9	0,2438
4,0	0,815	0,598	0,3668
		2,126	0,3996
		4,0	0,2336
4,2	0,825	0,730	0,4094
		2,334	0,3749
		4,2	0,2157
4,4	0,834	0,815	0,4358
		2,514	0,3634
		4,4	0,2008
4,6	0,842	0	0,0372
		0,937	0,4210
		2,684	0,3536
		4,6	0,1882
4,8	0,850	0	0,1455
		1,268	0,3581
		2,909	0,3220
		4,8	0,1744
5,0	0,857	0	0,1779
		1,463	0,3584
		3,116	0,3013
		5,0	0,1624
5,2	0,864	0	0,1937
		1,602	0,3660
		3,307	0,2883
		5,2	0,1520
5,4	0,870	0,278	0,2196
		1,749	0,3600
		3,490	0,2776
		5,4	0,1428
5,6	0,876	0,585	0,2849
		2,010	0,3234
		3,701	0,2582
		5,6	0,1335
5,8	0,881	0,712	0,3225
		2,224	0,3101
		3,908	0,2423
		5,8	0,1251
6,0	0,886	0,782	0,3426
		2,387	0,3083
		4,101	0,2314
		6,0	0,1177

a	$q_a(d_{\xi_a})$	$+c_{aj}$	ξ_{aj}
6,4	0,896	0	0,1270
		1,275	0,2842
		2,795	0,2768
		4,500	0,2077
		6,4	0,1043
6,8	0,904	0	0,1574
		1,559	0,3018
		3,174	0,2601
		4,897	0,1878
		6,8	0,0929
7,2	0,911	0,634	0,2501
		2,021	0,2631
		3,585	0,2347
		5,298	0,1691
		7,2	0,0830
7,6	0,917	0,774	0,2871
		2,341	0,2648
		3,966	0,2198
		5,695	0,1538
		7,6	0,0745
8,0	0,923	0	0,1179
		1,336	0,2419
		2,793	0,2351
		4,379	0,1988
		6,097	0,1392
		8,0	0,0671
9,0	0,935	0,738	0,2420
		2,228	0,2293
		3,764	0,2030
		5,373	0,1626
		7,096	0,1107
		9,0	0,0524
10,0	0,945	0	0,1181
		1,551	0,2288
		3,113	0,2085
		4,705	0,1767
		6,349	0,1362
		8,089	0,0899
		10,0	0,0418

максимум функции риска при байесовской оценке $q_a(d_{\xi_a}) = \sup\limits_{\mu \in [-a, a]} R(d_{\xi_a}, \mu)$. В таблице значения функционала $q_a(d_{\xi_a})$ приведены с точностью до трех знаков после запятой, а вычисления проводились с точностью до четырех знаков без учета погрешностей вычисления интеграла при определении значений функции риска $R(d_{\xi_a}, \cdot)$.

При вычислении значений функции $R(d_{\xi_a}, \cdot)$ использовалась квадратурная формула наивысшей алгебраической степени точности для интегрирования по оси $-\infty < x < +\infty$ с весом $\exp(-x^2)$. Вычисления проводились по формуле, точной для многочленов порядка 35, и проверялись по формуле, точной для многочленов порядка 33 [5]. Результаты вычислений по обеим формулам совпали с точностью до четвертого знака после запятой.

Поскольку вычисленные значения байесовского риска r_{ξ_a} и максимума байесовской функции риска $q_a(d_{\xi_a})$ совпали с точностью до четвертого знака после запятой, а для минимаксной оценки $d_a^*(X)$ имеют место неравенства

$$r_{\xi_a} \leqslant q_a(d_a^*) \leqslant q_a(d_{\xi_a}),$$

то в пределах точности проведенных вычислений решением поставленной задачи, т. е. минимаксной оценкой, можно считать байесовскую оценку $d_{\xi_a}(X)$. По предложению 7 эта оценка представима в виде

$$d_{\xi_a}(X) = \frac{\sum\limits_{j=1}^{m_a} +c_{aj}\xi_{aj}\, e^{-c_{aj}^2/2}\left(e^{+c_{aj}X} - e^{-c_{aj}X}\right)}{\sum\limits_{j=1}^{m_a} \xi_{aj}\, e^{-c_{aj}^2/2}\left(e^{+c_{aj}X} - e^{-c_{aj}X}\right)},$$

где для фиксированного значения a соответствующие числа $+c_{aj}$ и и ξ_{aj} $(j = 1, \ldots, m_a)$ выбираются из таблицы.

Заключение

Если задача состоит в построении минимаксной оценки среднего μ по выборке $X_1, \ldots, X_n$ из нормального распределения $N(\mu, \sigma^2)$

с известной дисперсией σ^2 при условии, что μ принадлежит заданному интервалу $[a_1, a_2]$, то эта задача сводится к вышеизложенной в силу того, что $\bar{X} = \sum_{i=1}^{n} X_i/n$ является достаточной статистикой. Принимая $a = (a_2 - a_1) n^{1/2}/(2\sigma)$ и $\mu_0 = (a_1 + a_2)/2$, можно записать оценку, приближенно минимизирующую максимум риска при квадратичной потере, в следующем виде:

$$d_{a_1, a_2, \sigma}(X_1, \ldots, X_n) = \mu_0 + (\sigma/n^{1/2}) d_{\xi_a}((\bar{X} - \mu_0) n^{1/2}/\sigma).$$

В настоящей работе построены минимаксные оценки при $a \leqslant 10$ с достаточной для приложений точностью. Ранее такие оценки при $a \leqslant 2$ были получены в [1]. При больших a можно пользоваться асимптотически минимаксными оценками, предложенными в [2, 3]. Представление о том, насколько минимаксная оценка лучше обычно используемой оценки $\bar{X}$ с постоянной функцией риска $R(\bar{X}, \cdot) = \sigma^2/n$, можно получить, используя для сравнения значения максимального риска из таблицы, умноженные на σ^2/n. В случае $a = 1{,}1$ минимальный выигрыш от использования минимаксной оценки, измеряемый относительным уменьшением риска, составляет 53%. С возрастанием a он, естественно, монотонно уменьшается и при $a = 10$ равняется 5,5%.

ЛИТЕРАТУРА

1. *Casella, G., Strawderman, W.* Ann. Statist., **9**, № 4, 870—878 (1981).
2. *Левит Б. Я.* Теория вероятностей и ее применения, **25**, № 3, 561—576 (1980).
3. *Bickel, P. J.* Ann. Statist., **9**, № 6, 1301—1309 (1981).
4. *Nelson, W.* Ann. Math. Statist., **37**, № 6, 1643—1657 (1966).
5. *Крылов В. И., Шульгина Л. Т.* Справочная книга по численному интегрированию. М., «Наука», 1966.

Институт кибернетики
Академии наук Эстонской ССР

Поступила в редакцию
30 сентября 1982

J. KUKS

ÜHEMÕÕTMELISE NORMAALJAOTUSE KESKVÄÄRTUSE LIGIKAUDNE MINIMAKSHINNANG

Olgu $X_1, \ldots, X_n$ valim ühemõõtmelisest normaaljaotusest $N(\mu, \sigma^2)$, mille dispersioon σ^2 on teada ja keskväärtus μ kuulub etteantud vahemikku $[a_1, a_2]$. Numbrilisel teel on leitud keskväärtuse μ ligikaudne minimakshinnang $d_{a_1, a_2, \sigma}(X_1, \ldots, X_n)$ neil juhtudel, kui $(a_2 - a_1) n^{1/2}/(2\sigma) \leqslant 10$.

J. KUKS

AN APPROXIMATELY MINIMAX ESTIMATOR OF THE MEAN OF AN ONE-DIMENSIONAL NORMAL DISTRIBUTION

Let $X_1, \ldots, X_n$ be a sample from one-dimensional normal distribution $N(\mu, \sigma^2)$ with known variance σ^2 and unknown mean μ that belongs to a given interval. The problem considered is to construct a minimax estimator of the parameter μ in the case of quadratic loss.

As $\bar{X} = \sum_{i=1}^{n} X_i/n$ is a sufficient statistic, so the problem reduces to a simpler form where $\sigma^2 = 1$, $n = 1$, and $a_1 = -a_2$. An approximately minimax solution has been numerically found by an iterative procedure in the cases when $(a_2 - a_1) n^{1/2}/(2\sigma) = 1.1(0.1)4(0.2)6(0.4)8(1)10$. The maximum risk of constructed estimators differs from minimax risk no more than 0.01 per cent and it is lower than the risk of the usual estimator $\bar{X}$, for example about 53 per cent if $(a_2 - a_1) n^{1/2}/(2\sigma) = 1.1$, and about 5.5 per cent if $(a_2 - a_1) n^{1/2}/(2\sigma) = 10$.