Известия вузов. Математика *http://www.ksu.ru/journals/izv_vuz/*
2010, №8, с. 89–98 *Гос. номер статьи по НТЦ "Информрегистр" 0421000123\0077*

Е.Д. ШЕРМАН

ЭМПИРИЧЕСКИЕ ОЦЕНКИ С МИНИМАЛЬНЫМ d-РИСКОМ ДЛЯ ДИСКРЕТНЫХ ЭКСПОНЕНЦИАЛЬНЫХ СЕМЕЙСТВ

Аннотация. Развивается эмпирический d-апостериорный подход к построению оценок с равномерно минимальным d-риском в случае полностью неизвестного априорного распределения. Эмпирические оценки, основанные на архиве данных, строятся для скалярного параметра дискретного экспоненциального семейства. Доказывается сходимость эмпирического d-риска к истинному. В качестве примера рассматривается оценка параметра распределения Пуассона. Точность оценок исследуется численно методом статистического моделирования.

Ключевые слова: эмпирический d-апостериорный подход, оценки с равномерно минимальным d-риском, сходимость эмпирического d-риска, дискретные экспоненциальные семейства.

УДК: 519.23

Abstract. We develop a d-posteriori approach to estimations with uniformly minimal d-risk, when the a priori distribution is completely unknown. For a scalar parameter of a discrete exponential family we construct empirical estimates based on archive data and prove the convergence of the empirical d-risk to the true one. As an example we adduce the estimation of the Poisson distribution parameter. We numerically study the accuracy of the estimates by the statistical modeling method.

Keywords: empirical d-posteriori approach, estimates with uniformly minimal d-risk, convergence of empirical d-risk, discrete exponential families.

Введение

Оценки с равномерно минимальным d-риском (ОРМД) рассматривались в работах И.Н. Володина, А.Н. Новикова и С.В. Симушкина ([1]–[4]) как некоторая байесовская альтернатива несмещенным оценкам с равномерно минимальным риском (НОРМР). Общая идеология d-апостериорного подхода к проблеме гарантийности статистического вывода и основные результаты по этой теме содержатся в обзоре [5]. Возможные варианты эмпирического подхода к d-гарантийности при неизвестном априорном распределении, предложенные в статье С.В. Симушкина [6], будут использованы в нашей работе применительно к проблеме построения ОРМД. Такие оценки впервые были рассмотрены в статье [1] (см. также [2]); существование, несмещенность и ряд других свойств ОРМД изучались в [3]. Наконец, в работе [4] было доказано, что оценки максимального правдоподобия имеют асимптотически минимальный d-риск.

Поступила 11.09.2008

В рамках классического подхода при вычислении средних потерь от принятия решения фиксируется некоторое значение параметра, рассматриваются эксперименты, соответствующие этому значению, и именно по этим экспериментам вычисляется величина средних потерь. В d-апостериорном подходе из последовательности экспериментов отбираются те, которые завершились принятием одного и того же решения (в нашем случае — оценки), и при этом величина средних потерь вычисляется только по этим экспериментам.

Естественно, d-апостериорный подход применим только в том случае, когда значение параметра в каждом эксперименте является реализацией случайной величины ϑ с некоторым априорным распределением. В случае известного априорного распределения построение ОРМД основано на минимизации апостериорного риска по значениям случайной выборки (сравните с байесовским подходом, при котором минимизация производится по возможным значениям оценки). Как отмечалось в работе [3], такая минимизация возможна лишь при наличии достаточной статистики той же размерности, что и параметр, — в противном случае требуется редукция выборочных данных к некоторой статистике (например, байесовской оценке). Аналогичная проблема возникает и в случае неизвестного априорного распределения при построении эмпирических аналогов ОРМД (так же, как и при эмпирическом байесовском подходе Г. Роббинса [7]).

В данной работе рассматривается дискретное экспоненциальное семейство с функцией плотности

$$f(x|\theta) = h(x)\theta^x b(\theta), \quad x \in \mathcal{X} \subseteq \{0, 1, 2, \dots\}, \quad \theta \in \mathrm{R}_+, \quad h(x) > 0 \ \ \forall x \in \mathcal{X}. \tag{1}$$

Для случая квадратичной функции потерь разрабатывается метод построения эмпирического аналога ОРМД по архиву данных $\mathrm{X}^{[n]}$ объема n, основанный на минимизации оценки апостериорного риска. Доказывается, что минимум оценки апостериорного риска по возможным значениям текущего эксперимента сходится к минимальному значению d-риска по вероятности, соответствующей безусловному распределению $\mathrm{X}^{[n]}$. В качестве примера рассматривается задача построения эмпирической ОРМД параметра пуассоновского распределения. Для случая априорного гамма-распределения методом статистического моделирования исследуются точностные свойства эмпирической ОРМД в сравнении с ОРМД, а также с эмпирической байесовской оценкой.

1. Оценка с равномерно минимальным d-риском для параметра дискретного распределения

Семейство распределений с функцией плотности (1) обладает достаточной статистикой, распределение которой имеет тот же вид, поэтому, не ограничивая общности, будем полагать объем выборки равным единице. Пусть $g(\theta)$, $\theta \in \mathrm{R}_+$, — функция плотности априорного распределения G параметра ϑ. В проблеме оценки θ при функции потерь $L(\theta, d)$ d-апостериорный риск (или, коротко, d-риск) $\Re(d|\delta)$ решающей функции (оценки) $\delta(x)$ определяется как условное математическое ожидание $L(\vartheta, d)$ относительно σ-алгебры, порожденной статистикой $\delta(x)$. Очевидно, $\Re(d|\delta)$ можно представить в виде условного математического ожидания апостериорного риска

$$R(d|x) = \frac{\int\limits_{\Theta} L(\theta, d) f(x|\theta) g(\theta) d\theta}{f(x)},$$

где $f(x)$ — маргинальная функция плотности случайной величины x.

Известно [1], что оценка, принимающая некоторое заданное значение d_0 лишь на точках выборочного пространства, доставляющих минимум апостериорного риска $R(d_0|x)$ (обозначим множество таких точек $X(d_0)$), обладает минимальным d-риском в точке $d = d_0$. Это означает, что для построения ОРМД $\delta^*(x)$ нужно для каждого фиксированного $d \in \mathrm{R}_+$ найти соответствующее множество $X(d)$ и принимать затем решение $\delta^*(x) = d$ лишь в том случае, когда результат наблюдения $x \in X(d)$. Отметим, что d-риск построенной таким образом оценки $\Re(d|\delta^*) = \min\limits_x R(d|x)$. Поэтому фактически задача построения ОРМД сводится к поиску множеств $X(d)$.

Может оказаться, что существует подмножество $D \in \mathrm{R}_+$, на котором

$$Z = \bigcap_{d \in D} X(d) \neq \emptyset,$$

— в этом случае при $x \in Z$ принимается любое решение $d \in D$ в соответствии с произвольным рандомизированным правилом. Существование таких подмножеств D типично для дискретных распределений, причем для любого $d \in D$ множества $X(d)$ совпадают и состоят из одной точки.

Поясним эту ситуацию на примере оценки параметра θ пуассоновского распределения при априорном гамма-распределении с параметром формы λ и параметром масштаба, равным единице. В этом случае апостериорный риск $R(d|x)$ при квадратичной функции потерь имеет вид

$$R(d|x) = (\lambda + x)(1 + \lambda + x)/4 - d(\lambda + x) + d^2.$$

Проекция графика этой функции при $\lambda = 1$ и $x = 0, 1, \ldots$ на плоскость (d, R) приведена на рис. 1 вверху.

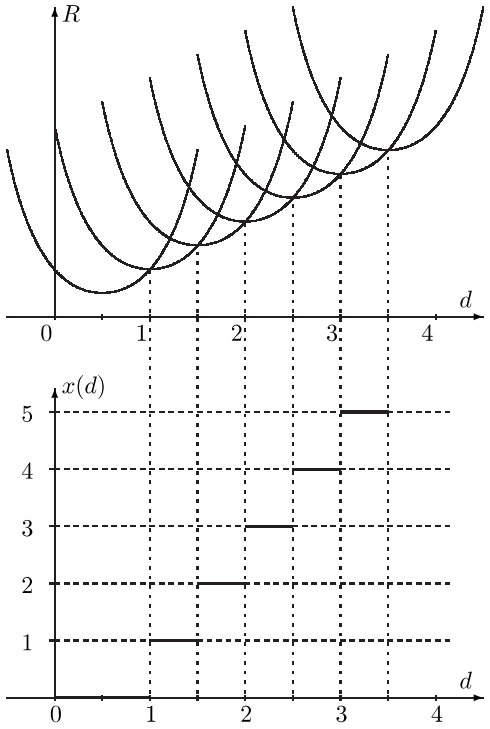

Рис. 1

Вид графика указывает на то, что существует счетное множество интервалов $D_i = [(i + \lambda)/2, (i + \lambda + 1)/2)$, на которых $X_i(d) = i$, $i = 1, 2, \dots$, и $X_0(d) = 0$ на $D_0 = [0, (i + \lambda)/2)$.

Внизу на рис. 1 приведена функция $x(d)$, $d \in \mathrm{R}_+$, представляющая точки достижения минимума по $x \in \mathcal{X} = \{0, 1, 2, \dots\}$ апостериорного риска $R(d|x)$ при каждом фиксированном $d \in \mathrm{R}_+$.

Таким образом, для любого результата текущего наблюдения x_0 можно принимать произвольное решение d, для которого $x(d) = x_0$, и возникает проблема выбора конкретного решения. Для рассмотренного примера представляется разумным выбирать в качестве оценки значение $d_0 = \min\{d : x(d) = x_0\}$, поскольку функция апостериорного риска $R(d|x_0)$ достигает минимума в точке $d = d_0$ (см. рис. 1).

Если выбрать другой детерминированный способ принятия решения, то получим иное правило с равномерно минимальным d-риском, сосредоточенное на другом подмножестве пространства решений. Поскольку эти два правила сосредоточены на разных носителях, они не сравнимы по d-риску.

Универсальным с этой точки зрения является рандомизированное правило, в соответствии с которым решение из интервала $\{d : x(d) = x_0\}$ выбирается случайным образом в соответствии с некоторым (например, равномерным) распределением, носителем которого является весь этот интервал.

2. Оценка апостериорного риска и ее сходимость

При полностью неизвестном априорном распределении для построения эмпирической оценки с равномерно минимальным d-риском (ЭОРМД) вместо функции $x(d)$ следует искать ее оценку $x_n(d)$, заменив в предложенной схеме минимизацию апостериорного риска на минимизацию некоторой его оценки $R_n(d|x)$, построенной по архиву данных $\mathrm{X}^{[n]} = (x_1, \dots, x_n, x)$. В связи с этим возникает вопрос о сходимости минимума $R_n(d|x)$ по всевозможным значениям эксперимента к соответствующему минимуму апостериорного риска (совпадающему с минимальным значением d-риска). Условие такой сходимости дает следующее предложение ([8], с. 45, теорема 5.7), где $\mathbf{P}$ – вероятность, соответствующая маргинальному распределению $\mathrm{X}^{[n]}$.

Предложение 1. *Если при каждом фиксированном $d \in \mathrm{R}_+$*

$$\lim_{n\to\infty} \mathbf{P}\big(\sup_{x\in\mathcal{X}} |R_n(d|x) - R(d|x)| > \varepsilon\big) = 0, \tag{2}$$

то $\inf\limits_{x\in\mathcal{X}} R_n(d|x)$ сходится при $n \to \infty$ по вероятности $\mathbf{P}$ к минимальному значению d-риска.

Это предложение носит общий характер: оно справедливо для произвольных семейств распределений случайных величин и для любых функций потерь.

В данной работе рассматривается задача построения ЭОРМД параметра семейства (1) при квадратичной функции потерь и полностью неизвестном априорном распределении. Нетрудно убедиться, что апостериорный риск можно представить в виде

$$R(d|x) = \frac{h(x)}{h(x+2)}\frac{f(x+2)}{f(x)} - 2d\frac{h(x)}{h(x+1)}\frac{f(x+1)}{f(x)} + d^2.$$

Подставив в это выражение вместо неизвестной априорной плотности $f(y)$ ее частотную оценку $f_n(y) = m/(n+1)$, где m — количество элементов архива $\mathrm{X}^{[n]} = \{x_1, \dots, x_n, x\}$, равных y, получаем оценку апостериорного риска

$$R_n(d|x) = \frac{h(x)}{h(x+2)}\frac{f_n(x+2)}{f_n(x)} - 2d\frac{h(x)}{h(x+1)}\frac{f_n(x+1)}{f_n(x)} + d^2. \tag{3}$$

Условия, при которых данная оценка равномерно по вероятности сходится к апостериорному риску, устанавливает

Теорема. *Если случайная величина x принимает бесконечное множество значений с положительной вероятностью ($\mathcal{X}$ счетно), $h(x) \searrow 0$ при $x \to \infty$ и существует такая последовательность $c_n \to \infty$, что*

$$\lim_{n\to\infty} \frac{n^\alpha e^{\beta c_n} h(c_n+2)}{c_n^\lambda} = +\infty \quad \forall \alpha > 0,\ \beta > 0,\ \lambda > 0, \tag{4}$$

то минимум по x оценки апостериорного риска (3) *сходится по вероятности* $\mathbf{P}$ *к минимальному значению d-риска.*

Доказательство. Достаточно показать выполнение условия (2). Очевидно, для оценки (3) справедливо асимптотическое $((f_n(x) - f_g(x))/f_g(x) \to 0)$ представление

$$\begin{aligned} R_n(d|x) = \frac{h(x)}{h(x+2)} \left(\frac{f_n(x+2) - f(x+2)}{f(x)} - \frac{f_n(x+2)(f_n(x) - f(x))}{f^2(x)} \right) - \\ - 2d \frac{h(x)}{h(x+1)} \left(\frac{f_n(x+1) - f(x+1)}{f(x)} - \frac{f_n(x+1)(f_n(x) - f(x))}{f^2(x)} \right) + \\ + o\left(\frac{f_n(x) - f(x)}{f(x)} \right) + R(d|x). \end{aligned}$$

Вероятность в условии (2) мажорируется выражением

$$\begin{aligned} \mathbf{P}\big(\sup_x \big| R_n(d|x) - R(d|x) \big| > \varepsilon \big) \leq \\ \leq \mathbf{P}\left(\sup_x \frac{h(x)}{h(x+2)f(x)} |f_n(x+2) - f(x+2)| > \frac{\varepsilon}{5} \right) + \\ + \mathbf{P}\left(\sup_x \frac{h(x) f_n(x+2)}{h(x+2)f^2(x)} |f_n(x) - f(x)| > \frac{\varepsilon}{5} \right) + \\ + \mathbf{P}\left(\sup_x 2d \frac{h(x)}{h(x+1)f(x)} |f_n(x+1) - f(x+1)| > \frac{\varepsilon}{5} \right) + \\ + \mathbf{P}\left(\sup_x 2d \frac{h(x) f_n(x+1)}{h(x+1)f^2(x)} |f_n(x) - f(x)| > \frac{\varepsilon}{5} \right) + \\ + \mathbf{P}\left(\sup_x o\left(\frac{f_n(x) - f(x)}{f(x)} \right) > \frac{\varepsilon}{5} \right). \end{aligned} \tag{5}$$

Покажем, что первое слагаемое в (5) (обозначим его $\mathbf{P}_1$) сходится к нулю при $n \to \infty$; для остальных слагаемых это делается аналогично.

Для любой последовательности c_n, удовлетворяющей (4),

$$\begin{aligned} \mathbf{P}_1 = \mathbf{P}\left(\sup_x |f_n(x+2) - f(x+2)| > \frac{\varepsilon h(x+2) f(x)}{5h(x)} \right) \leq \\ \leq \mathbf{P}\left(\left\{ \sup_x |f_n(x+2) - f(x+2)| > \frac{\varepsilon h(x+2) f(x)}{5h(x)} \right\} \bigcap \{x < c_n\} \right) + \mathbf{P}\left(x > c_n \right). \end{aligned}$$

В силу счетности $\mathcal{X}$ плотность $f(x) \to 0$ при $x \to \infty$. Следовательно, существует такое N, что для всех $n > N$ при $x < c_n$ имеет место неравенство $f(x) > f(c_n)$. Используя это, а

также неравенство Чебышева, получаем для $\mathbf{P}_1$ оценку сверху

$$\mathbf{P}_1 \le \mathbf{P}\left(\left\{\sup_x |f_n(x+2) - f(x+2)| > \frac{\varepsilon h(x+2) f(c_n)}{5h(x)}\right\} \bigcap \{x < c_n\}\right) + \frac{\mathbf{E}X}{c_n}.$$

Чтобы убедиться в том, что первое слагаемое в правой части этого неравенства сходится к нулю, оценим снизу маргинальную плотность. Пусть $a(\theta)$ — некоторая функция, удовлетворяющая следующим двум условиям: $a(\theta)$ имеет единственный максимум в точке θ_0, причем $a(\theta_0) > 0$, и $a(\theta) \le \ln\theta \quad \forall\, \theta \in \mathrm{R}_+$. Тогда

$$f(c_n) = h(c_n) \int_\Theta \mathrm{e}^{c_n \ln\theta} b(\theta) g(\theta) d\theta \ge h(c_n) \int_\Theta \mathrm{e}^{c_n a(\theta)} b(\theta) g(\theta) d\theta.$$

При $c_n \to \infty$ для последнего интеграла справедливо асимптотическое представление ([9], с. 55)

$$\int_\Theta \mathrm{e}^{c_n a(\theta)} b(\theta) g(\theta) d\theta \sim \sqrt{-\frac{2\pi}{c_n a''(\theta_0)}}\ \mathrm{e}^{c_n a(\theta_0)} b(\theta_0) g(\theta_0).$$

Используя эту формулу, а также тот факт, что в силу монотонности убывания функции $h(x)$

$$\frac{h(c_n) h(x+2)}{h(x)} \ge \frac{h^2(c_n+2)}{h(0)}$$

при $x < c_n$, убеждаемся в справедливости неравенства

$$\mathbf{P}\left(\left\{\sup_x |f_n(x+2) - f(x+2)| > \frac{\varepsilon h(x+2) f(c_n)}{5h(x)}\right\} \bigcap \{x < c_n\}\right) \le$$
$$\le \mathbf{P}\left(\max_{x: x < c_n} |f_n(x+2) - f(x+2)| > \frac{\varepsilon K_1 h^2(c_n+2) \mathrm{e}^{K_2 c_n}}{\sqrt{c_n}}\right),$$

где K_1 и K_2 — положительные постоянные. Наконец, применяя к полученному выражению известную оценку

$$\mathbf{P}\left(|f_n(x) - f(x)| > \gamma\right) < \exp\left(-2n\gamma^2\right),$$

находим

$$\mathbf{P}_1 \le \frac{\mathbf{E}X}{c_n} + \exp\left(-\varepsilon^2 K_1^2 \frac{n h^4(c_n+2) \mathrm{e}^{2K_2 c_n}}{c_n}\right).$$

В силу выбора последовательности c_n имеем $\mathbf{P}_1 \to 0$ при $n \to \infty$.

Как показывает следующее предложение, в случае конечного $\mathcal{X}$ при доказательстве утверждения теоремы можно обойтись без достаточно громоздкого условия (4).

Предложение 2. *Если носитель распределения случайной величины x с плотностью* (1) *является конечным множеством, то условие* (2) *выполняется без дополнительных ограничений на функцию h.*

Доказательство. Пусть $\mathcal{X} = \{0, 1, \ldots, k\}$, тогда $\forall x \in \mathcal{X} : C_1 \le h(x) \le C_2$, где C_1 и C_2 — некоторые положительные постоянные. В этом случае оценку снизу маргинальной функции плотности $f(x)$ можно модифицировать следующим образом:

$$f(x) = h(x) \int_\Theta \mathrm{e}^{x \ln\theta} b(\theta) g(\theta) d\theta \ge C_1 \int_\Theta \mathrm{e}^{x a(\theta)} b(\theta) g(\theta) d\theta,$$

отсюда

$$f(x) \ge C_1 \sqrt{-\frac{2\pi}{x a''(\theta_0)}}\ \mathrm{e}^{x a(\theta_0)} b(\theta_0) g(\theta_0) \ge C_1 \sqrt{-\frac{2\pi}{k a''(\theta_0)}}\ b(\theta_0) g(\theta_0) = C_3.$$

Тогда оценка сверху вероятности $\mathbf{P}_1$ примет вид

$$\mathbf{P}_1 \le \mathbf{P}\left(\max_{x\in\mathcal{X}}|f_n(x+2)-f(x+2)| > \frac{\varepsilon h(x+2)f(x)}{5h(x)}\right) \le$$
$$\le \mathbf{P}\left(\max_{x\in\mathcal{X}}|f_n(x+2)-f(x+2)| > \frac{\varepsilon C_1 C_3}{5C_2}\right) \le \exp\left(-2n\varepsilon^2 C\right),$$

где C — положительная постоянная, т. е. $\mathbf{P}_1 \to 0$ при $n\to\infty$.

3. Эмпирическая оценка с равномерно минимальным d-риском для параметра пуассоновского распределения

Рассмотрим в качестве примера построения ЭОРМД задачу нахождения оценки параметра θ распределения Пуассона с плотностью

$$f(x|\theta) = \frac{\theta^x e^{-\theta}}{x!}, \quad x = 0,1,\ldots,$$

при полностью неизвестном априорном распределении.

Здесь $h(x) = (x!)^{-1}$ удовлетворяет условию (4) с последовательностью $c_n = \ln\ln n$, а оценка апостериорного риска (3), основанная на архиве данных $\mathrm{X}^{[n]}$, имеет вид

$$R_n(d|x) = (x+1)(x+2)\frac{f_n(x+2)}{f_n(x)} - 2d(x+1)\frac{f_n(x+1)}{f_n(x)} + d^2. \tag{6}$$

Как было отмечено выше, алгоритм отыскания ОРМД сводится к построению функции $x(d)$, для этого апостериорный риск $R(d|x)$ при каждом фиксированном $d\in\mathrm{R}_+$ минимизируется по $x\in\mathcal{X}$.

Следовательно, для построения эмпирического аналога ОРМД необходимо найти функцию $x_n(d)$ — оценку $x(d)$, заменив минимизацию апостериорного риска на минимизацию его оценки $R_n(d|x)$, вообще говоря, по всем $x\in\mathrm{X}^{[n]}$. Однако из представления (6) видно, что ЭОРМД можно построить лишь в том случае, когда архив данных вместе с результатом текущего наблюдения x содержит также значения $x+1$ и $x+2$ (похожее условие возникает при построении эмпирической байесовской оценки).

Пусть $x^{[n]} = \{x : x,(x+1),(x+2)\in\mathrm{X}^{[n]}\}$ — множество элементов архива, удовлетворяющих этому условию. Тогда функцию $R_n(d|x)$ при каждом фиксированном d следует минимизировать по всем $x\in x^{[n]}$.

В этом случае функция $x_n(d)$, $d\in\mathrm{R}_+$, представляет точки достижения минимума по $x\in x^{[n]}$ для оценки апостериорного риска $R_n(d|x)$ при каждом фиксированном $d\in\mathrm{R}_+$, и в качестве оценки параметра θ при текущем результате наблюдения $x\in x^{[n]}$ выбирается значение

$$\widehat{\theta}_n = \min_{\{d:x_n(d)=x\}} R_n(d|x).$$

Чтобы найти функцию $x_n(d)$, поставим в соответствие каждому $x_i\in x^{[n]}$ интервал $D_i = \{d : R_n(d|x_i) \le R_n(d|x) \quad \forall x\in x^{[n]}\}$, на котором $x_n(d) = x_i$. Может оказаться, что для некоторых значений $x_0\in x^{[n]}$ соответствующий интервал $D_0 = \emptyset$, и функция $x_n(d)\ne x_0$ ни при каких значениях $d\in\mathrm{R}_+$. В этом случае, а также тогда, когда $x_0\notin x^{[n]}$, ЭОРМД не существует, и в качестве оценки параметра θ при таких результатах текущего наблюдения можно использовать эмпирическую байесовскую оценку (если она существует) или оценку максимального правдоподобия.

В предположении, что априорное распределение параметра θ принадлежит семейству гамма-распределений $G(\lambda,\alpha)$ с параметром масштаба $\alpha = 1$ и параметром формы $\lambda = 1$,

методом статистического моделирования были накоплены архивы данных объема $n = 100$ и $n = 10000$.

Рассмотрим проекцию $R_n(d|x)$ на плоскость (d, R_n) — соответствующие графики при $n = 100$ и $n = 10000$ приведены на рис. 2 и 3 сверху. Очевидно, для результата наблюдения x_i интервал D_i соответствует той части графика $R_n(d|x_i)$, которая расположена ниже всех остальных $R_n(d|x), x \in x^{[n]}$, что указывает на простой способ построения функции $x_n(d)$ (графики $x_n(d)$ приведены в нижней части рис. 2 и 3).

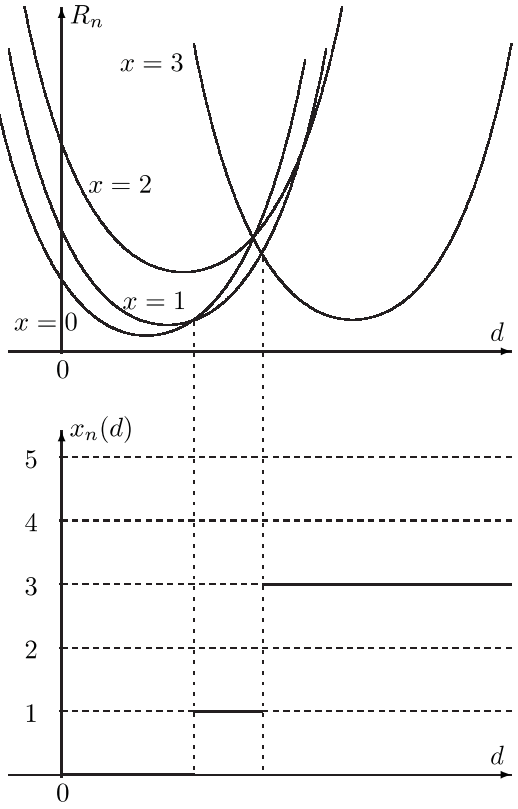

Рис. 2. $n = 100$

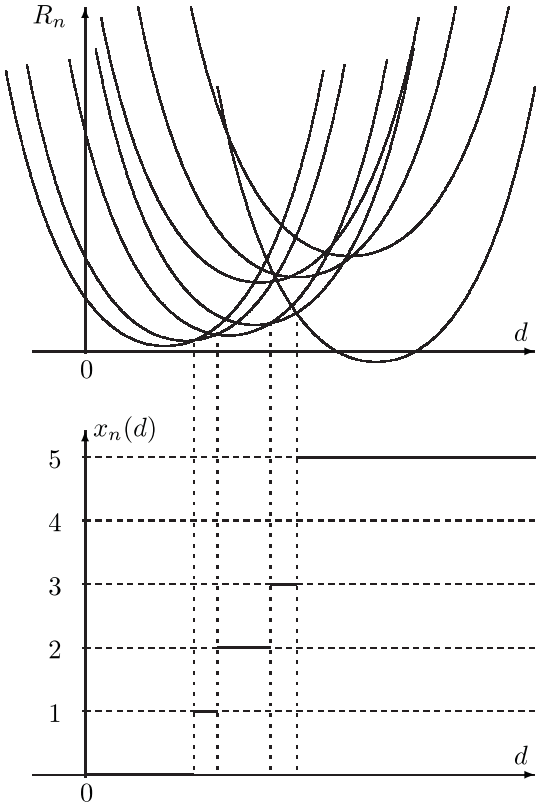

Рис. 3. $n = 10000$

При $n = 100$ имеем множество $x^{[n]} = \{0, 1, 2, 3\}$, причем в том случае, когда результат текущего наблюдения $x = 2$, ЭОРМД не существует, так как $\{d : x_n(d) = 2\} = \emptyset$ (рис. 2). При увеличении объема архива до 10000 множество $x^{[n]}$ также увеличивается: $x^{[n]} = \{0, 1, 2, 3, 4, 5, 6, 7\}$, при этом ЭОРМД не существует при $x = 4, 6, 7$.

В предположении, что априорное распределение принадлежит семейству гамма-распределений $G(\lambda, \alpha)$ с параметром масштаба $\alpha = 1$ и со значениями параметра формы λ, равными 1, 2, 4 и 10, методом статистического моделирования были вычислены различные оценки параметра θ распределения Пуассона для архивов данных разного объема n. С помощью датчика случайных чисел, имеющих гамма-распределение, были получены значения θ, которые приведены в соответствующем столбце таблицы. Для этих значений θ строились архивы данных, включающие результаты текущих экспериментов. По этим данным вычислялись значения оценок с равномерно минимальным d-риском при известном априорном

распределении (столбец θ_g), ЭОРМД при полностью неизвестном априорном распределении (столбец $\widehat{\theta}_n$), а также эмпирические байесовские оценки (столбец δ_n) и оценки максимального правдоподобия (столбец θ_n^*). Прочерк (–) означает, что оценки не существует.

Таблица

n	$\lambda=1$					$\lambda=2$				
	θ	θ_g	$\widehat{\theta}_n$	δ_n	θ_n^*	θ	θ_g	$\widehat{\theta}_n$	δ_n	θ_n^*
100	1.2	1	0.95	0.79	1	1.53	2	–	2.65	2
500	0.52	0.5	0.49	0.49	0	0.84	1	0.95	0.95	0
1000	0.47	0.5	0.5	0.5	0	1.56	1.5	1.63	1.63	1
5000	1.5	0.5	0.49	0.49	0	4.46	2.5	2.46	2.36	3
10000	1.44	1	1.03	0.99	1	1.3	1.5	1.55	1.55	1

n	$\lambda=4$					$\lambda=10$				
	θ	θ_g	$\widehat{\theta}_n$	δ_n	θ_n^*	θ	θ_g	$\widehat{\theta}_n$	δ_n	θ_n^*
100	4.73	3.5	7.25	7.25	3	8.37	9.5	–	7.14	9
500	2.18	3.5	–	4.27	3	9.26	8.5	10.23	10.23	7
1000	5.2	3.5	3.63	3.63	3	7.07	10.5	–	9.75	11
5000	4.41	4.5	4.44	4.44	5	10.54	10.5	–	11.25	11
10000	2.91	3	2.93	2.93	2	12.56	11	12.76	10.99	12

Данные в таблице указывают на то, что ЭОРМД часто совпадает с эмпирической байесовской оценкой. Для объяснения этого факта достаточно преобразовать оценку апостериорного риска (6) к виду

$$R_n(d|x) = (d-\delta_n(x))^2 + (x+1)(x+2)\frac{f_n(x+2)}{f_n(x)} - \delta_n^2(x),$$

где $\delta_n(x) = (x+1)f_n(x+1)/f_n(x)$ — эмпирическая байесовская оценка. Если при результате текущего наблюдения x точка $d=\delta_n(x)$ принадлежит множеству $D=\{d : x_n(d)=x\}$, то, очевидно, именно в этой точке достигается $\min R_n(d|x)$ по всем $d\in D$.

Данные в таблице также показывают, что с увеличением λ уменьшается вероятность существования ЭОРМД. Подобный феномен будет наблюдаться и в том случае, когда объем выборки намного больше единицы, что типично для практических приложений. Следовательно, данный метод оценки параметра может быть рекомендован только при небольших объемах выборки и в предположении о малых значениях параметра θ.

Литература

[1] Simushkin S.V., Volodin I.N. *Statistical inference with a minimal d-risk*, Lect. Note in Math. **1021**, 107–114 (1983).

[2] Володин И.Н., Симушкин С.В. *Статистические выводы с минимальным d-риском*, Исследования по приклад. матем. (Изд-во КГУ, Казань, 1984), Вып. 11, 25–39.

[3] Володин И.Н., Симушкин С.В. *Несмещенность и байесовость*, Изв. вузов. Математика, № 1, 3–7 (1987).

[4] Володин И.Н., Новиков А.Н. *Статистические оценки с асимптотически минимальным d-риском*, Теория вероят. и ее применен. **38** (1), 20–32 (1993).

[5] Володин И.Н., Новиков А.Н., Симушкин С.В. *Гарантийный статистический контроль качества: апостериорный подход*, Обозрение приклад. и промыш. математ. (Изд-во ТВП, М., 1994) **1** (2), 1–32 (1994).

[6] Симушкин С.В. *Эмпирический d-апостериорный подход к проблеме гарантийности статистического вывода*, Изв. вузов, Математика, № 11, 42–58 (1983).

[7] Robbins H. *An empirical Bayes approach to statistic*, Proc. Third Berkeley Symp. Math. Statist. Probab. Univ. of Calif. Press **1**, 157–164 (1955).
[8] van der Vaart A.W. *Asymptotic statistics* (Cambridge University Press, Cambridge, 1998).
[9] Федорюк М.В. *Асимптотика: интегралы и ряды* (Наука, М., 1987).

Е.Д. Шерман

ассистент, кафедра математической статистики,
Казанский государственный университет,
ул. Кремлевская, д. 18, г. Казань, 420008,

`e-mail:` sedgold@mail.ru

E.D. Sherman

Junior Researcher, Chair of Mathematical Statistics,
Kazan State University,
18 Kremlyovskaya str., Kazan, 420008 Russia,

`e-mail:` sedgold@mail.ru