ISSN 2949-5598

ДИСКРЕТНЫЙ АНАЛИЗ И ИССЛЕДОВАНИЕ ОПЕРАЦИЙ

Том 31 № 2 2024

Новосибирск
Издательство Института математики

ДИСКРЕТНЫЙ АНАЛИЗ И ИССЛЕДОВАНИЕ ОПЕРАЦИЙ
Апрель–июнь 2024. Т. 31, № 2. С. 80–95

УДК 519.24 DOI: 10.33048/daio.2024.31.785

СРАВНИТЕЛЬНЫЙ АНАЛИЗ АЛГОРИТМОВ ДЛЯ ОЦЕНКИ ДИНАМИКИ ЧИСЛЕННОСТИ РЫБ

В. И. Зоркальцев[a], *А. С. Князев*[b]

Байкальский гос. университет,
ул. Ленина, 11, 664003 Иркутск, Россия
E-mail: [a]`vizork@mail.ru`, [b]`010193@bgu.ru`

Аннотация. Цель представляемой разработки состояла в создании инструмента для сопоставления алгоритмов оценки динамики численности отдельных видов рыбы озера Байкал на основе экспериментальных отловов. Для каждого экземпляра случайно отобранной рыбы по специальной технологии определяется её возраст. Из полученных данных о численности рыбы разных возрастов в выборке оцениваются параметры заданных законов возрастного распределения. Это служит основанием для формирования представления о динамике смертности и изменения популяции рыбы данного вида в предыдущие годы. Для оценки параметров законов распределения могут использоваться разные алгоритмы, приводящие порой к существенно разным результатам.

Обсуждаемая методика сопоставления алгоритмов оценки параметров основывается на многократных вычислительных экспериментах с использованием метода Монте-Карло для имитации случайных выборок рыб. По предлагаемой методике анализируются несколько алгоритмов оценки параметров усечённого экспоненциального закона распределения при различных объёмах выборок. В качестве примера рассмотрена задача оценки динамики смертности голомянки, основной по биомассе рыбы озера Байкал. Ил. 4, библиогр. 13.

Ключевые слова: оценка параметров смертности рыб, озеро Байкал, голомянка, усечённый экспоненциальный закон распределения, метод статистических испытаний.

Введение

В данной статье излагается методика сравнительного анализа алгоритмов обработки данных возрастной структуры случайных выборок рыб озера Байкал для определения показателей динамики их смертности по возрастам. Используя разные алгоритмы, можно получать из одних и тех же исходных данных существенно разные результаты. Возникает проблема, какой из алгоритмов следует выбрать? С этой проблемой тесно связаны и другие вопросы. В том числе: как влияет объём используемой выборки на результаты? Какой объём выборки можно считать минимально достаточным для получения надёжных результатов? В данной статье методику сравнительного анализа алгоритмов будем рассматривать на примерах алгоритмов оценки распределения возрастного состава рыб по усечённому экспоненциальному закону.

По традиции теоретическое обоснование алгоритма оценки параметров закона распределения случайной величины осуществляется на основе исследования свойств этого алгоритма при возрастании объёма выборки до бесконечности. При этом используются такие характеристики получаемых оценок, как состоятельность, асимптотическая несмещённость, асимптотическая эффективность, асимптотическая нормальность [1–3]. На практике оценка параметров закона распределения производится на базе ограниченной выборки. Вместе с тем, свойства алгоритмов, выявляемые при возрастании выборки до бесконечности, малопригодны для сравнения этих алгоритмов при оценке параметров на базе ограниченных выборок. О том, что способы оценки параметров, свойственных большим выборкам, не подходят в реальных ситуациях с ограниченным объёмом выборки, писал ещё Гаусс [4].

Статья подготовлена на основе доклада [5], представленного авторами на 2-м Международном семинаре «Вычислительные технологии и прикладная математика», прошедшем 12–16 июня 2023 г. в Благовещенске. Авторы надеются, что представленная в данной статье технология сравнительного анализа методов оценки параметров законов распределения случайных величин на базе вычислительных экспериментов будет полезна не только для исследования и выбора методов оценки динамики смертности рыб озера Байкал, но и в других случаях.

1. Исходные положения

Излагаемая методика базируется на многократной имитации методом статистических испытаний (методом Монте-Карло) заданного объёма случайных реализаций по рассматриваемому закону распределения. Для каждой имитации осуществляется оценка параметров исследуемым алгоритмом. Тем самым получается множество расчётных реализаций параметров, которое можно рассматривать как набор случайных чисел. По их отклонениям от истинного значения параметров можно определять различные статистические характеристики результатов использования алгоритма оценки параметров. Выбор наилучшего алгоритма, исследования влияния объёмов используемых выборок могут осуществляться путём сопоставления рассчитываемых характеристик для разных алгоритмов, для разных объёмов выборок.

В данной статье излагаются результаты разработки и апробации данной методики на примерах алгоритмов оценки параметров усечённого экспоненциального закона. Согласно этому закону вероятность случайно пойманной рыбе иметь возраст t выражается зависимостью

$$P_t = AR^t, \quad t = 1, \dots, T, \tag{1}$$

где T — заданный максимальный возраст дожития,

$$R = \exp(-\Lambda) \tag{2}$$

— условная вероятность дожития до возраста $t+1$ для рыбы, дожившей до возраста $t = 1, \dots, T-1$. Здесь Λ — положительный параметр, который принято называть применительно к динамике жизни биологических организмов коэффициентом смертности. Из условия, что сумма вероятностей P_t, $t = 1, \dots, T$, равна единице, следует, что

$$A = \frac{1-R}{1-R^{T+1}}. \tag{3}$$

Отметим, что при фиксированном T все величины определяются однозначно, если задать показатель $R \in (0,1)$ или параметр $\Lambda > 0$.

Оценки указанных параметров будем обозначать теми же, только строчными буквами. В этой статье производится сопоставление по четырём возможным методам, каждый из которых применяется в трёх вариантах. Итого будет рассмотрено двенадцать алгоритмов оценки параметров усечённого экспоненциального закона распределения.

Методика использует m случайных выборок рыб одного и того же объёма N с заданным законом распределения случайной величины возраста рыб в выборке. Обозначим через N_{ti} численность рыб возраста t в выборке $i \in \{1, \dots, m\}$,

$$\sum_{t=1}^{T} N_{ti} = N, \quad i = 1, \dots, m. \tag{4}$$

В проведённых ниже расчётах используется $m = 50\,000$ выборок, что является достаточным для получения устойчивых результатов. В первых двух методах сначала оценивается условная вероятность r дожития до следующего года для рыбы, дожившей до года $t = 1, \dots, T-1$. Затем на основе (1)–(3) определяются оценки λ, a, p_t. В третьем и четвёртом методах сначала вычисляется коэффициент смертности λ, а затем, исходя из (1)–(3), определяются значения r, a, p_t, $t = 1, \dots, T$.

В рамках каждого из рассматриваемых методов алгоритмы отличаются используемыми в расчётах положительными весовыми коэффициентами h_t. Рассматриваются следующие весовые коэффициенты:

1) неизменные во времени

$$h_t = \frac{1}{T-1}, \quad t = 1, \ldots, T, \tag{5}$$

2) пропорциональные оценкам вероятности p_t

$$h_t = p_t / \sum_{\tau=1}^{T-1} p_\tau, \quad t = 1, \ldots, T, \tag{6}$$

3) пропорциональные обратным значениям оценок дисперсии случайных величин p_t

$$h_t = \frac{1}{p_t(1-p_t)} / \sum_{\tau=1}^{T-1} \frac{1}{p_\tau(1-p_\tau)}, \quad t = 1, \ldots, T. \tag{7}$$

Во всех трёх вариантах сумма первых $T-1$ весовых коэффициентов равна единице. В правилах (6), (7) весовые коэффициенты зависят от величин p_t, которые предстоит определить с использованием этих же весовых коэффициентов. Это достигается применением итеративной процедуры, включающей на каждой итерации оценку вероятностей p_t (при некоторых заданных в начале итерации весовых коэффициентах), а затем пересчёт весовых коэффициентов по формулам (6) или (7) с использованием полученных оценок p_t. Этот процесс повторяется, пока оценки вероятностей p_t на соседних итерациях не совпадут с заданной точностью. В качестве начального приближения используются одинаковые весовые коэффициенты, определяемые по формуле (4).

Приведём расчётные формулы методов оценки параметров. В этих описаниях индекс — номер выборки i опущен. Используются относительные величины

$$n_t = \frac{N_t}{N}, \quad t = 1, \ldots, T,$$
$$q_t = \frac{N_{t+1}}{N_t}, \quad t = 1, \ldots, T-1.$$

1. Оценка условной вероятности дожития до следующего года в виде взвешенного среднего арифметического от соотношений количеств рыб соседних возрастов в выборке равна

$$r = \sum_{t=1}^{T-1} h_t q_t. \tag{8}$$

2. Оценка условной вероятности дожития до следующего года в виде взвешенного среднего геометрического от соотношений количеств рыб

соседних возрастов в выборке равна

$$r = \prod_{t=1}^{T-1} q_t^{h_t}. \tag{9}$$

В обоих методах на основе найденного по формулам (8) или (9) значения r вычисляются

$$\lambda = -\ln(r), \quad a = \frac{1-r}{1-r^{T+1}},$$
$$p_t = ar^t, \quad t = 1, \ldots, T.$$

3. Оценка коэффициента смертности с использованием метода наименьших квадратов в логарифмической шкале. Сначала определяются значения переменных α, λ в результате безусловной минимизации квадратичной выпуклой функции от двух переменных

$$\varphi(\alpha, \lambda) = \sum_{t=1}^{T} h_t(\ln n_t - \alpha - t\lambda)^2. \tag{10}$$

Затем, исходя из (1)–(3), определяются значения

$$r = \exp(-\lambda), \quad a = \frac{1-r}{1-r^{T+1}},$$
$$p_t = ar^t, \quad t = 1, \ldots, T. \tag{11}$$

4. Оценка методом наименьших квадратов в исходной шкале. Необходимо найти минимум функции одной переменной

$$f(\lambda) = \sum_{t=1}^{T} h_t(n_t - a(\lambda)\exp(-\lambda t))^2, \tag{12}$$

$$a(\lambda) = \frac{1-r(\lambda)}{1-r(\lambda)^{T+1}}, \quad r(\lambda) = \exp(-\lambda). \tag{13}$$

Остальные параметры вычисляются по формулам (11). Функция f минимизируется методом золотого сечения [6].

Можно считать, что изложенные методы представлены в последовательности от наиболее простых к более сложным в отношении восприятия и вычислений. В созданном программно-вычислительном комплексе предусмотрена возможность добавления других методов.

Результаты оценки параметров для каждого варианта случайных выборок $i = 1, \ldots, m$ используются для вычислений обобщённых характеристик данного метода. В реализованном к настоящему времени программно-вычислительном комплексе рассчитываются следующие характеристики.

1. Математическое ожидание коэффициента смертности, абсолютное и относительное смещение математического ожидания от истинного значения коэффициента смертности:

$$\mathbb{M}\lambda = \frac{1}{m}\sum_{i=1}^{m}\lambda_i, \quad \Delta = \mathbb{M}\lambda - \Lambda, \quad \delta = \frac{\Delta}{\Lambda}. \tag{14}$$

2. Абсолютные и относительные средние квадратические отклонения оценок коэффициента смертности от математического ожидания этих оценок и от истинного значения этих оценок:

$$D = \left(\frac{1}{m}\sum_{i=1}^{m}(\lambda_i - \mathbb{M}\lambda)^2\right)^{1/2}, \quad \delta D = \frac{D}{\mathbb{M}\lambda}, \tag{15}$$

$$Q = \left(\frac{1}{m}\sum_{i=1}^{m}(\lambda_i - \Lambda)^2\right)^{1/2}, \quad \delta Q = \frac{Q}{\Lambda}.$$

3. Вероятности (частоты) попадания значений λ_i в заданные интервалы (например в 20% или 10%) от математического ожидания $\mathbb{M}\lambda$ и от истинного значения Λ.

4. Медиана оценок коэффициента смертности: значение λ, при котором половина оценок λ_i будет меньше этой величины. Абсолютные и относительные отклонения медианы от истинного значения Λ.

Рассматриваемая методика позволяет использовать и другие показатели качества алгоритмов оценки параметров. Например, приведённые показатели можно дополнить аналогичными характеристиками для условной вероятности r дожития до следующего года рыб, доживших до данного возраста.

2. Обсуждение

Можно выделить три источника погрешностей в оценках динамики смертности рыб на основе данных экспериментальных отловов.

1. Неадекватность рассматриваемой модели при сопоставлении с действительностью. В данном случае — возможное несоответствие реального распределения возраста рыб усечённому экспоненциальному закону. Об экспоненциальном распределении можно говорить, если есть веские основания полагать, что ежегодно в водоёме появляется примерно одно и то же число рыб-сеголеток исследуемого вида, доля выживающих за год рыб примерно одинаковая для каждого возраста, и эта доля не изменяется во времени.

Для указанных условий стационарности в Байкале наиболее подходящей является рыба голомянка. Имеются два вида — большая и малая голомянка, суммарно составляющих по имеющимся оценкам около

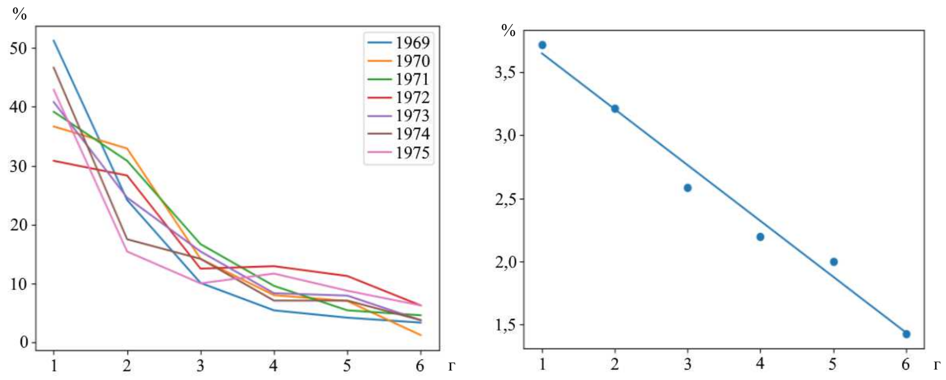

Рис. 1. Возрастной состав (%) большой голомянки в отловах 1969–75 гг. [7] и линейная аппроксимация усреднённых данных в логарифмической шкале

70% рыб Байкала [7]. Голомянка — живородящая рыба (вынашивающая внутри себя личинки из икры), не собирающаяся в стаи — распределённая по всей территории и по всей глубине озера. Оба вида голомянок всю свою жизнь проводят непосредственно в акватории озера, поэтому не испытывают особого влияния колебаний приточности воды в реках, впадающих в Байкал. Конечно, влияние на численность голомянки могут оказывать изменения кормовой базы (биомассы отдельных видов зоопланктона) и, вероятно, некоторые отмечаемые изменения температуры верхних слоёв воды озера Байкал. К факторам возможной нестационарности условий существования голомянки следует отнести и колебание численности байкальской нерпы, для которой голомянка является источником питания. Представленные на рис. 1 данные Старикова [7] о возрастной структуре большой голомянки в отловах 1969–75 гг. можно рассматривать как подтверждение возможности использования усечённого экспоненциального закона распределения для описания структуры возрастного состава.

Ещё более оправданным выглядит такое допущение при рассмотрении усреднённых за весь указанный семилетний период данных о возрастной структуре большой голомянки. Как видно из рис. 1, усреднённая за семь лет возрастная структура в логарифмической шкале хорошо аппроксимируется линейной зависимостью. В представленных ниже результатах расчётов будем рассматривать усечённый экспоненциальный закон с параметрами $T = 6$ и $\Lambda = 0,44$. Согласно рис. 1 это примерно соответствует ожидаемым значениям параметров усечённого экспоненциального закона распределения для большой голомянки.

Помимо экспоненциального могут рассматриваться и другие законы распределения случайной величины. В том числе, более общие, которые учитывают возможные зависимости коэффициента Λ от возраста. К ним относится распределение Вейсбула [1], в котором коэффициент Λ задаётся в виде линейной функции от возраста.

2. Погрешности экспериментов, непредставительность используемых выборок. Причиной этого могут быть случайные и систематические отклонения в возрастной структуре выборки, влияние на используемые в расчётах исходные данные времени, места и технологии проведения отловов. Для промысловых рыб (в частности для байкальского омуля [8]) отмечается сокращение удельного веса рыб младших возрастов, поскольку эти рыбы из-за меньших размеров чаще проходят сквозь ячейки рыболовецких сетей, а также из-за мотивации рыболовов к вылову крупных рыб.

Погрешности возможны и в определении возраста отдельных экземпляров рыб, что оценивается на основе специальной трудоёмкой технологии анализа рыбной чешуи, отолитов и костей рыб [9]. Одним из планируемых направлений развития излагаемой в этой статье методики оценки параметров является включение в неё методов, учитывающих возможные погрешности в самих исходных данных в том числе в оценках возраста отдельных экземпляров рыб.

3. Погрешности алгоритмов расчёта. На одних и тех же исходных данных при оценке параметров одного и того же закона распределения случайной величины можно получать существенно разные результаты разными алгоритмами оценки параметров. Излагаемая здесь методика исследования алгоритмов нацелена в основном на уменьшение роли этих погрешностей.

3. Результаты расчётов

На рис. 2–4 представлены результаты расчётов для трёх из приведённых выше характеристик качества алгоритмов оценки параметров в зависимости от объёма выборки N. Рассматриваются следующие характеристики: относительное смещение δ математического ожидания оценки коэффициента смертности от его истинного значения (14) (рис. 2); относительное среднеквадратическое отклонение δD оценки коэффициента смертности от её математического ожидания (15) (рис. 3); частота попадания оценки коэффициента смертности в 10%-й интервал от его истинного значения (рис. 4). На каждом рисунке отражены расчёты для всех четырёх изложенных методов с тремя видами весовых коэффициентов: а) метод взвешенного среднего арифметического (8); б) метод взвешенного среднего геометрического (9); в) метод наименьших квадратов в логарифмической шкале (10)–(11); г) метод наименьших квадратов в исходной шкале (12)–(13). Способ взвешивания обозначен цветом: зелёный — равные неизменные веса (5); жёлтый — веса, пропорциональные оценкам

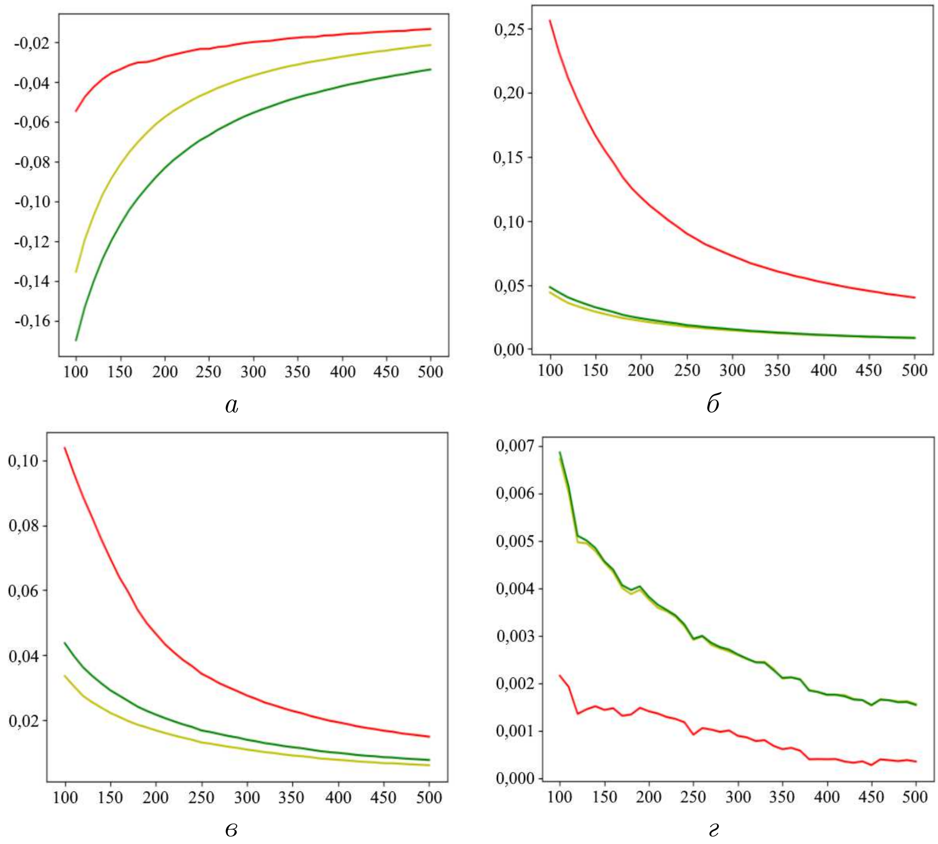

Рис. 2. Относительное смещение δ математического ожидания оценки коэффициента смертности от его истинного значения

вероятности (6); красный — веса, пропорциональные обратным значениям оценок дисперсии (7). В представленных расчётах объём выборки варьируется в диапазоне 100–500 экземпляров рыб.

Из представленных расчётов видно, что выбор метода и правила задания весовых коэффициентов и объёмов используемых выборок имеет очень существенное значение. Все методы имеют значительно худшие характеристики при малых объёмах используемых выборок, скажем, при $N \leqslant 300$, чем при больших значениях N. Этот факт важен для выработки рекомендации о минимально необходимом числе рыб в выборках.

Качество методов по всем рассматриваемым характеристикам возрастает от более простых к более сложным. Если для метода среднего геометрического (второго из описанных выше) смещение математического ожидания λ достигает 15–25% при выборке в 100–200 экземпляров, то для четвёртого метода смещение имеет малозначительную величину

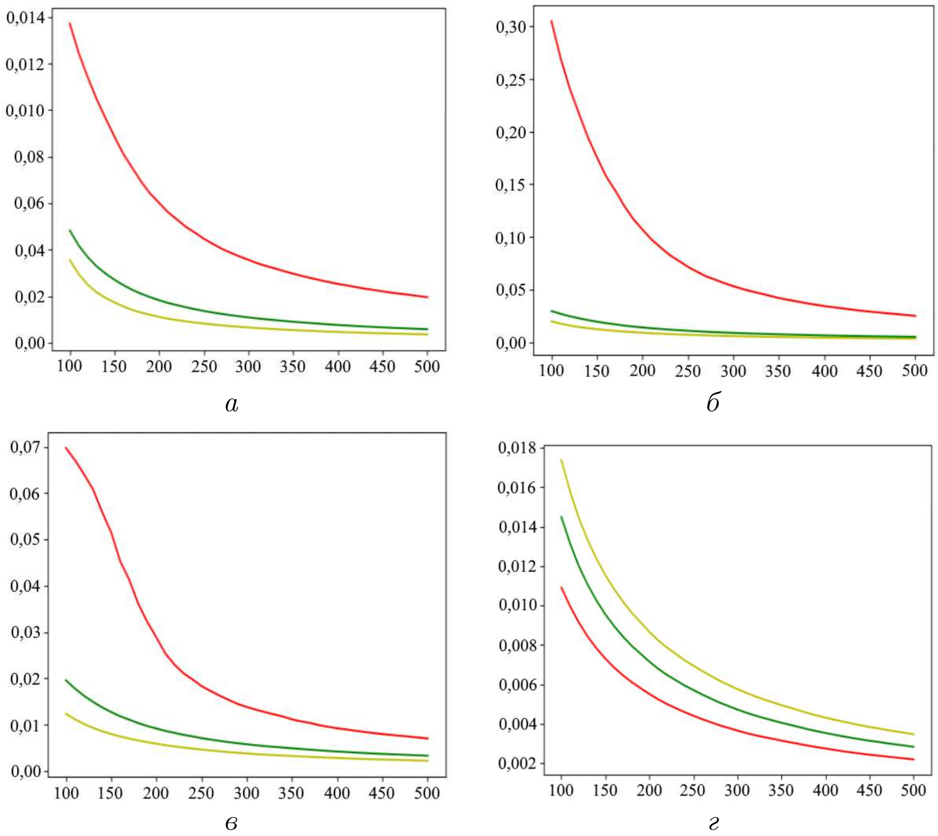

Рис. 3. Относительное среднеквадратическое отклонение δD оценки коэффициента смертности от её математического ожидания

в 0,7–0,4% при таких же объёмах выборки. Малая величина смещения для метода (12), (13), выражаемая тысячными долями, объясняет, почему в этом случае построенная расчётная зависимость относительного смещения λ не имеет монотонного и гладкого характера. Чтобы достичь строгой монотонности и гладкости графика, надо многократно увеличить число выборок m.

Для первых трёх методов явно лучшим оказался второй способ взвешивания (6), а не третий (7). Проблема выбора весов при оценке параметров усечённого экспоненциального метода распределения возраста большой и малой голомянки третьим из изложенных здесь методов обсуждалась в статье [10]. Обычно в литературе по методам математической статистики рекомендуется использовать веса, пропорциональные обратным значениям точности измерений показателей, под которой понимается оценка дисперсии измеряемой величины. Это соответствует третьему

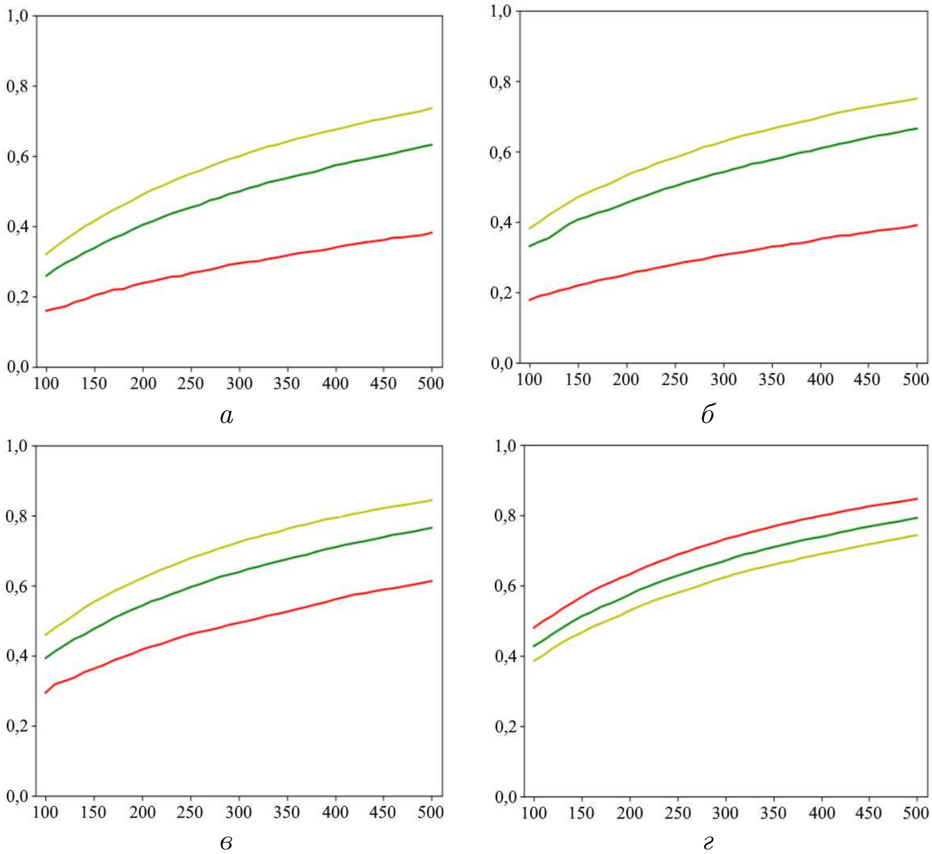

Рис. 4. Вероятность попадания оценки коэффициента смертности в 10%-й интервал от его истинного значения

способу взвешивания, выражаемому формулой (7). В таком случае веса h_t должны возрастать с увеличением t. В то же время имелись весомые аргументы, изложенные в [10], по использованию весов h_t, убывающих с увеличением t, а именно весовых коэффициентов (6). Стремление строго доказать преимущество коэффициентов (6) привело к реализации первоначального варианта излагаемой здесь методики [10]. В [10] представлены результаты исследования первых трёх из рассматриваемых здесь четырёх методов для усечённого экспоненциального распределения при $T = 6$ и $\Lambda = 0{,}65$. Это примерно соответствует параметрам малой голомянки [7]. Соотношения исследовавшихся показателей оказались примерно такими же, как и представленные в настоящей статье для данных, примерно адекватных для большой голомянки.

Оценка динамики смертности (по различным причинам) отдельных видов организмов озера Байкал необходима для построения автономных моделей функционирования этих организмов в Байкале, что является одним из этапов создания модели экосистемы озера с учётом взаимодействий разных видов между собой [12, 13].

Наилучшие характеристики из приведённых здесь имеет четвёртый метод, базирующийся на минимизации функции (12) и расчётных формулах (13). Причём для этого метода лучшим оказался третий вариант взвешивания, а увеличение объёма используемой выборки до 300–500 экземпляров позволяет увеличить вероятность попадания оценки коэффициента смертности в 10%-й интервал от истинного значения до 70–80%, что не достижимо другими методами из рассмотренных.

Можно отметить, что и по другим характеристикам, не приведённым здесь из-за ограниченного объёма статьи, также отмечается существенное превосходство четвёртого из рассматриваемых методов при небольшом преимуществе третьего способа взвешивания. Конечно, необходимо более подробное исследование соотношений расчётных характеристик алгоритмов оценки параметров, устойчивости этих соотношений к варьированию исходного значения Λ и теоретическое объяснение получаемых результатов.

4. Некоторые выводы

1. Представленные результаты расчётов иллюстрируют работоспособность развиваемой методики исследования алгоритмов оценки параметров законов распределения на базе вычислительных экспериментов, важность учёта объёма используемых выборок, важность выбора правила задания весовых коэффициентов. Можно утверждать, что анализ алгоритмов оценки параметров по рассматриваемой здесь технологии должен быть необходимым этапом при сопоставлениях и обосновании методов оценки параметров законов распределения случайных величин на базе ограниченных выборок.

2. Расчёты иллюстрируют преимущество второго способа взвешивания для рассмотренных здесь первых трёх методов оценки параметров и преимущество четвёртого из рассмотренных методов оценки параметров на этот раз с третьим способом взвешивания. Можно отметить, что такие соотношения преимуществ алгоритмов имеют место и по другим рассчитанным характеристикам, не приведённым здесь из-за ограниченности объёма статьи.

3. Данная методика может использоваться не только для усечённого экспоненциального, но и для других законов распределения случайных величин. Может быть расширен состав рассчитываемых характеристик алгоритмов оценки параметров, расширен набор методов оценки параметров (целесообразно, в частности, включение в этот набор метод максимального правдоподобия).

В дальнейшем планируется осуществить развитие данной методики в том числе в направлении учёта возможных ошибок в определении возраста рыб, использования других законов распределения и выявления возможных возмущений, приводящих к увеличению или сокращению численности рыб в отдельные годы прошлых лет. Требуется развитие технологии в направлении исследования устойчивости получаемых характеристик алгоритмов в окрестности ожидаемых значений параметров законов распределения.

Финансирование работы

Исследование выполнено за счёт бюджета Байкальского государственного университета. Дополнительных грантов на проведение или руководство этим исследованием получено не было.

Конфликт интересов

Авторы заявляют, что у них нет конфликта интересов.

Литература

1. **Кобзарь А. И.** Прикладная математическая статистика. М.: Физматлит, 2012. 816 с.
2. **Трофимова Е. А., Кисляк Н. В., Гилёв Д. В.** Теория вероятностей и математическая статистика. Екатеринбург: Изд-во Урал. ун-та, 2018. 160 с.
3. **Козлов М. В., Прохоров А. К.** Введение в математическую статистику. М.: МГУ, 1987. 264 с.
4. **Гаусс К. Ф.** Избранные геодезические сочинения. Т. 1. М.: Геодезиздат, 1957. 152 с.
5. **Зоркальцев В. И., Князев А. С.** Сравнительный анализ методов измерения динамики смертности рыб озера Байкал на основе вычислительных экспериментов // Вычислительные технологии и прикладная математика. Мат. II Междунар. семин. (Благовещенск, Россия, 12–16 июня 2023 г.). Благовещенск: Изд-во Амур. гос. ун-та, 2023. С. 91–94.
6. **Зоркальцев В. И.** Элементы оптимизации. Иркутск: ИСЭМ СО РАН, 2014. 100 с.
7. **Стариков Г. В.** Голомянки Байкала. Новосибирск: Наука, 1977. 93 с.
8. **Соколов А. В., Петекрфельд В. А., Бобков А. И.** [и др.]. Омуль байкальский // Материалы, обосновывающие общие допустимые уловы водных биологических ресурсов в озере Байкал на 2018 год. М.: ВНИРО, 2017. С. 5–24.
9. **Кафанова В. В.** Методы определения возраста и роста рыб. Томск: Изд-во Томск. ун-та, 1984. 56 с.
10. **Бычков И. В., Зоркальцев В. И., Казазаева А. В.** Весовые коэффициенты в методе взвешенных наименьших квадратов // Сиб. журн. вычисл. математики. 2015. Т. 18, № 3. С. 275–288.

11. **Бычков И. В., Зоркальцев В. И., Мокрый И. В.** Сопоставление методов оценки параметров усечённого экспоненциального закона распределения на основе вычислительных экспериментов // Вычисл. технологии. 2018. Т. 23, № 5. С. 3–20.
12. **Аношко П. Н., Афанасьева Э. Л., Дзюба Е. В.** [и др.]. Формирование системы моделей трофических взаимоотношений в экосистеме пелагиали оз. Байкал. Иркутск, 1998. 21 с. (Препр. / ИСЭМ СО РАН; № 11).
13. **Зоркальцев В. И., Мокрый И. В., Казазаева А. В.** Моделирование пелагического сообщества экосистемы озера Байкал // Вычисл. технологии. 2011. Т. 16, № 1. С. 48–60.

Зоркальцев Валерий Иванович
Князев Александр Сергеевич

Статья поступила
6 сентября 2023 г.

После доработки —
13 ноября 2023 г.

Принята к публикации
22 декабря 2023 г.

DISKRETNYI ANALIZ I ISSLEDOVANIE OPERATSII
/DISCRETE ANALYSIS AND OPERATIONS RESEARCH/
April–June 2024. Vol. 31, No. 2. P. 80–95

UDC 519.24 DOI: 10.33048/daio.2024.31.785

COMPARATIVE ANALYSIS OF ALGORITHMS FOR ESTIMATING FISH POPULATION DYNAMICS

V. I. Zorkaltsev[a] and *A. S. Knyazev*[b]

Baikal State University,
11 Lenin Street, 664003 Irkutsk, Russia
E-mail: [a]`vizork@mail.ru`, [b]`010193@bgu.ru`

Abstract. The aim of the present work is to create a tool for comparing algorithms for estimating the dynamics of the abundance of individual fish species in Lake Baikal based on experimental catching. For each instance of a randomly selected fish, its age is determined using a special technology. From the obtained data on the numbers of fish of different ages in the sample, the parameters of the given laws of age distribution are estimated. This serves as a basis for the formation of ideas about the dynamics of mortality and changes in the abundance of fish of this species in previous years. Various algorithms can be used to estimate the parameters of distribution laws, sometimes leading to considerably different results.

The discussed technique for comparing parameter estimation algorithms is based on multiple computational experiments using the Monte Carlo method to simulate random samples of fish. The proposed method analyzes several algorithms for estimating the parameters of a truncated exponential distribution law for different sample sizes. As an example, the problem of estimating the mortality dynamics of the Baikal oilfish (Comephorus), the major biomass fish of Lake Baikal, is considered. Illustr. 4, bibliogr. 13.

Keywords: estimation of fish mortality parameters, Lake Baikal, comephorus, truncated exponential distribution law, method of statistical testing.

References

1. **A. I. Kobzar',** *Applied Mathematical Statistics* (Fizmatlit, Moscow, 2012) [Russian].

English transl.: *Journal of Applied and Industrial Mathematics* **18** (2), 361–370 (2024), DOI: 10.1134/S1990478924020170.

2. **E. A. Trofimova, N. V. Kislyak,** and **D. V. Gilyov,** *Probability Theory and Mathematical Statistics* (Izd. Ural. Univ, Yekaterinburg, 2018) [Russian].
3. **M. V. Kozlov** and **A. K. Prokhorov,** *Introduction to Mathematical Statistics* (MGU, Moscow, 1987) [Russian].
4. **K. F. Gauss,** *Selected Geodesic Works,* Vol. 1 (Geodezizdat, Moscow, 1957) [Russian].
5. **V. I. Zorkaltsev** and **A. S. Knyazev,** Comparative analysis of methods for assessing the dynamics of fish mortality in Lake Baikal based on computational experiments, in *Computation Technologies and Applied Mathematics* (Proc. 2nd Int. Semin., Blagoveshchensk, Russia, June 12–16, 2023) (Izd. Amur. Gos. Univ., Blagoveshchensk, 2023), pp. 91–94 [Russian].
6. **V. I. Zorkaltsev,** Elements of Optimization Theory (ISEhM SO RAN, Irkutsk, 2014) [Russian].
7. **G. V. Starikov,** *Golomyankas of The Baikal* (Nauka, Novosibirsk, 1977) [Russian].
8. **A. V. Sokolov, V. A. Petekrfeld, A. I. Bobkov,** [et al.], The Baikal omul, in *Proceedings Justifying the Total Allowable Catches of Aquatic Biological Resources in Lake Baikal for 2018* (VNIRO, Moscow, 2017), pp. 5–24 [Russian].
9. **V. V. Kafanova,** *Methods for Determining the Age and Growth of Fish* (Izd. Tomsk. Univ., Tomsk, 1984) [Russian].
10. **I. V. Bychkov, V. I. Zorkaltsev,** and **A. V. Kazazaeva,** Weight coefficients in the weighted least squares method, *Sib. Zh. Vychisl. Mat.* **18** (3), 275–288 (2015) [Russian] [*Numer. Anal. Appl.* **18** (3), 223–234 (2015)].
11. **I. V. Bychkov, V. I. Zorkaltsev,** and **I. V. Mokryi,** Estimates of parameters for the truncated exponential law of distribution on the basis of exponential data, *Vychisl. Tekhnol.* **23** (5), 3–20 (2018) [Russian].
12. **P. N. Anoshko, Eh. L. Afanasyeva, E. V. Dzyuba,** [et al.], Formation of a system of models for trophic relationships in the pelagic ecosystem of Lake Baikal. Irkutsk, 1998. (Prepr., ISEhM SO RAN, No. 11) [Russian].
13. **V. I. Zorkaltsev, I. V. Mokryi,** and **A. V. Kazazaeva,** Modeling the pelagic community of Lake Baikal ecosystem, *Vychisl. Tekhnol.* **16** (1), 48–60 (2011) [Russian].

Valery I. Zorkaltsev
Aleksandr S. Knyazev

Received September 6, 2023
Revised November 13, 2023
Accepted December 22, 2023