УДК 519.856 **DOI** 10.15622/sp.2020.19.1.7

В.Р. Соболь, Р.О. Торишный

ПРИМЕНЕНИЕ ГЛАДКОЙ АППРОКСИМАЦИИ ФУНКЦИЙ ВЕРОЯТНОСТИ И КВАНТИЛИ ПРИ РЕШЕНИИ ЗАДАЧ СТОХАСТИЧЕСКОГО ПРОГРАММИРОВАНИЯ

Соболь В.Р., Торишный Р.О. **Применение гладкой аппроксимации функций вероятности и квантили при решении задач стохастического программирования.**

Аннотация. Исследуется один из возможных вариантов гладкой аппроксимации вероятностных критериальных функций в задачах стохастического программирования, позволяющий получить оценки градиента функции вероятности и функции квантили в форме объемного интеграла. Исследование проведено в приложении к задачам максимизации функции вероятности и минимизации функции квантили для функционала потерь, зависящего от вектора управления и одномерной абсолютно непрерывной случайной величины.

Основная идея аппроксимации – замена разрывной функции Хевисайда в интегральном представлении функции вероятности на гладкую функцию, обладающую такими свойствами, как непрерывность, гладкость, а также имеющую легко вычислимые производные. Примером такой функции является функция распределения случайной величины, распределенной по логистическому закону с нулевым средним и конечной дисперсией – сигмоида. Величина, обратно пропорциональная корню из дисперсии, при этом является параметром, который обеспечивает близость исходной функции и ее аппроксимации. Такая замена позволяет получить гладкое приближение функции вероятности, для которого легко могут быть найдены производные по вектору управления и иным параметрам задачи.

Основным результатом статьи являются полученные выражения для аппроксимации производных функции вероятности по вектору управления и по допустимому уровню потерь, а также выражения для аппроксимации градиента функции квантили в форме объемных интегралов. Доказана сходимость аппроксимации функции вероятности, полученной при замене функции Хевисайда на сигмоидальную функцию, к исходной функции вероятности, и получена оценка погрешности такой аппроксимации. Также доказана сходимость аппроксимации производных функции вероятности к истинным производным при выполнении ряда условий на функционал потерь.

Рассмотрены примеры, демонстрирующие возможность применения предложенных оценок к решению задач стохастического программирования с критериальными функциями в форме функции вероятности и функции квантили, в том числе в случае многомерной случайной величины.

Ключевые слова: стохастическое программирование, вероятностный критерий, квантильный критерий, аппроксимация, численные методы, сигмоидальная функция

1. Введение. В прикладных задачах управления и оптимизации нередко возникают проблемы, требующие принятия решений в условиях неопределенности. Существуют различные подходы ее учета, основывающиеся на специфике исследования. В самом простом случае неопределенность учитывают как реализацию какого-то исхода, рассматривая только результаты эксперимента или наблюдения без применения каких-либо теоретических моделей; это видится единственно верным вариантом, если природа неопределенности не может быть установлена. Таким случаем учета неопределенности является, например, наблюдение некоторых случайных искажений в генетических данных, описанных в [1]. Во многих других задачах используются различные теоретические аппараты учета специфических неопределенностей, причем достаточно широко распространен принцип учета неопределенности с использованием аппаратов нечетких множеств или его производных. Например, в [2] автором предлагается подход к решению задачи оптимизации ресурсов в мелкосерийном производственном цикле с помощью нечеткой логики, описывающей неопределенности стохастической природы. В [3] апробируется модель поддержки принятия решений о выборе портфеля проектов в рамках стратегического развития университета, где степени желания изменений действующих лиц задаются нечеткими правилами вывода, и ограничения модели также являются нечеткими. В [4] представлена идея перехода к нечетким критериям в конкретной задаче анализа и оптимизации параметров приемников ГНСС, а в [5] делается упор на альтернативные меры исчисления истинности и на конструкции более сложные, нежели классическая вероятность.

Одним из основных подходов к учету неопределенности является подход на основе аппарата теории вероятностей. Это порождает задачи стохастического программирования, представляющие большую сложность в сравнении с обычными оптимизационными задачами. Наиболее сложные и интересные из них — задачи с вероятностными критериями оптимизации, например, задачи, связанные с обеспечением приемлемого уровня надежности или минимального уровня потерь при фиксированном уровне надежности. В монографии [6] приведено большое количество прикладных задач такого типа и описаны способы их решения.

Необходимо отметить, что в последнее время ведутся исследования задач стохастического программирования, критерии которых сложнее обычной функции вероятности или квантили, и при некоторых условиях такие задачи могут быть решены. Например, в задачах, в которых критериальная функция представляет собой сложный нелинейный функционал риска, решение может быть найдено с помощью статистических оценок [7]. Также для некоторых задач может быть применима теория о представлении функции риска как комбинации ее предельной и условной форм [8], и в таких случаях решение задачи немного упрощается. В настоящей работе будут рассматриваться классические задачи стохастической оптимизации, критериальной функцией в которых выступает либо функция вероятности, либо функция квантили.

В настоящее время разработано большое количество алгоритмов решения конкретных задач стохастического программирования. Большая часть исследований направлена на сведение таких задач к более простым. Например, в работе [9] разработаны алгоритмы сведения задачи квантильной оптимизации с дискретным распределением к задаче смешанного целочисленного программирования; к такому же типу задач возможно свести и двухшаговую задачу стохастического программирования при определенных условиях [10]. Также были разработаны алгоритмы решения двухэтапной линейной задачи стохастического программирования путем выборочных аппроксимаций [11] и более общие алгоритмы решения, основанные на аппроксимациях, например в задаче хеджирования [12].

В перечисленных выше и других работах рассматриваются различные постановки прикладных задач стохастической оптимизации в разнообразных областях, например задача оптимального инвестирования, задача оптимального управления спутниками Земли, задача оптимизации площади взлетно-посадочной полосы, задача об оптимальном проектировании систем водоснабжения и другие, что подчеркивает актуальность данной тематики.

Одним из важных вопросов в области решения задач стохастического программирования является нахождение значения градиента критериальной вероятностной функции. Решение такого вопроса дает возможность использовать для любой задачи с вероятностным критерием классические методы оптимизации, хоть и с некоторыми оговорками. Впервые выражение градиента функции вероятности в форме поверхностного интеграла в смысле Римана было получено в [13]. Далее, в [14] приводится выражение градиента функции вероятности в форме интеграла Лебега по поверхности. Эта поверхность является границей области реализаций случайного вектора и потому ее нахождение в явном виде представляет некоторую трудность. В [15] было показано, что в некоторых случаях поверхностный интеграл в формуле для градиента функции вероятности может быть преобразован в поверхностный интеграл для векторного поля. Применяя формулу Гаусса-Остроградского, можно свести этот интеграл к объемному, но только для некоторых частных случаев. В задачах, где такое сведение невозможно, автором [16] было доказано, что градиент функции вероятности возможно представить в виде суммы интегралов по поверхности и по объему. В любом случае полученные представления градиента функции вероятности представляют собой поверхностные интегралы, вычисление которых связано с различного вида сложностями, а представления или хотя бы оценки градиента функции вероятности в форме объемных интегралов до настоящего времени получены не были.

Более поздние работы, сопряженные с этой темой, предлагают некоторую аппроксимацию градиента функции вероятности или оценку его значения. Например, в [17] для задачи линейного программирования с вероятностным ограничением получены нижние оценки нормы градиента функции вероятности и проведен анализ зависимости оптимального значения критерия относительно распределения вектора случайных параметров, а в [18] получены верхние оценки для субдифференциалов Кларка и Мордуховича. Стоит отметить, что обе эти работы посвящены случаю гауссовского распределения параметров. Несколько другой подход, а именно использование слабых производных и сэмплирования, позволили оценить производную функции вероятности [19]. В [20] получены аппроксимации квантили и производных плотности распределения при помощи дополнения рядов Тейлора для нормального, экспоненциального и хи-квадрат распределений. Механизмы приближенного решения дифференциальных уравнений и нахождения значений градиента для распределения Вейбулла и их применение в сфере оценки радиационных данных было приведено в [21]. Наконец, в случае нормального распределения параметров с малой дисперсией авторами [22] предложены асимптотические формулы для первой и второй производных функции вероятности. Значения производных при этом оцениваются при помощи методов Монте-Карло. Результаты, полученные в данном блоке работ, решают задачу аппроксимации градиента вероятностной функции, но ставят дополнительные ограничения на распределение случайной величины либо привязывают качество аппроксимации к другим случайным механизмам.

Также необходимо отметить, что в последнее время интерес представляет некоторое обобщение классической функции вероятности, а именно буферная функция вероятности (англ. *buffered probability of exceedance, bPOE*). Эта конструкция, описанная вместе со своими свойствами в [23], по сути, представляет собой модификацию функции вероятности с некоторым так называемым «буфером безопасности» как по уровню предполагаемых потерь, так и по значению вероятности. Для вышеуказанного обобщения в [24] были найдены формулы для дифференциалов и субдифференциалов, но эти формулы неприменимы к классической функции вероятности.

В настоящей статье исследуется другой подход к аппроксимации вероятностных критериальных функций и их производных в случае случайной величины с абсолютно непрерывным распределением и непрерывной функции потерь. Основной идеей является получение аппроксимации критериальной функции вместе с ее производными через аппроксимацию индикаторной функции. Рассматриваемая аппроксимация универсальна для критериальной функции любого вида и в общем случае не представляет вычислительной сложности, так как производные критериальных функций представляются в виде объемных интегралов. Таким образом, представленный алгоритм аппроксимации позволит унифицировать процесс решения задачи, а также использовать широко известные методы оптимизации первого порядка.

2. Постановка задачи и проблема вычисления градиента. Пусть $(\Omega, \mathscr{F}, \mathbb{P})$ — полное вероятностное пространство, где $\Omega \subset \mathbb{R}$ — некоторое замкнутое множество. Пусть X — абсолютно непрерывная случайная величина, заданная на этом пространстве и имеющая плотность распределения $f(x)$. Пусть также задана функция потерь $\Phi(u, X) : U \times \Omega \to (-\infty; +\infty)$, где $U \subset \mathbb{R}^m$ — непустое множество допустимых стратегий u. Также пусть задан произвольный максимально допустимый уровень потерь φ.

Критериальной функцией является функция вероятности

$$P_\varphi(u) \stackrel{\Delta}{=} \mathbb{P}\left\{x : \Phi(u, x) \leqslant \varphi\right\},$$

а сама задача представляет собой максимизацию этой функции:

$$P_\varphi(u) \to \max_u . \qquad (1)$$

Существует также альтернативная постановка, когда в качестве критериальной выбрана функция квантили

$$\varphi_\alpha(u) \stackrel{\Delta}{=} \min\left\{\varphi : P_\varphi(u) \geqslant \alpha\right\},$$

а задача представляет собой минимизацию функции квантили для любого фиксированного уровня $\alpha \in (0; 1)$:

$$\varphi_\alpha(u) \to \min_u . \qquad (2)$$

Такие постановки описывают одношаговую модель принятия решения в условиях неопределенности, моделируемой случайной величиной X. При этом само принятие решения, то есть выбор стратегии управления u, производится априори до реализации случайной величины, опираясь исключительно на знание закона распределения случайного вектора. Стратегия при этом представляется в виде вектора, компоненты которого являются вещественными числами. Стратегия выбирается исходя из условия минимума (или максимума, в зависимости от формулировки) критерия оптимизации.

Также отметим, что в любой из постановок задач могут быть дополнительные ограничения, но так как к вероятностным ограничениям также применима аппроксимация, описанная далее, а нестохастические ограничения сужают лишь рассматриваемые области управления, и в обоих случаях структура целевой функции от ограничений не зависит, на данном шаге рассмотрения задачи ограничения мы опускаем.

Перейдем к сути проблемы вычисления градиента функции вероятности. Согласно [25], функция вероятности может быть представлена в виде:

$$P_\varphi(u) = \mathbb{P}\left\{X : \Phi(u, X) \leqslant \varphi\right\} = \mathbb{M}\left[I_\varphi(u, X)\right] = \int\limits_{-\infty}^{+\infty} I_\varphi(u, X) f(x) dx,$$

где

$$I_\varphi(u, X) = \begin{cases} 1, & \Phi(u, X) \leqslant \varphi, \\ 0, & \Phi(u, X) > \varphi. \end{cases}$$

Известно, что аналитическим представлением индикаторной функции является функция Хевисайда:

$$\Theta(y) = \begin{cases} 1, & y \geqslant 0, \\ 0, & y < 0. \end{cases}$$

Таким образом, функция вероятности может быть записана как:

$$P_\varphi(u) = \int\limits_{-\infty}^{+\infty} \Theta\left(\varphi - \Phi(u, x)\right) f(x) dx. \tag{3}$$

Правая часть выражения (3) при принятых условиях может быть дифференцируема. В результате дифференцирования подынтегрального выражения появится дельта-функция Дирака. Эта обобщенная функция ввиду своей природы не может быть использована в дальнейших аналитических преобразованиях, поэтому логичным видится замена функции Хевисайда в исходном определении функции вероятности на аппроксимирующую функцию, обладающую некоторыми свойствами:

– близость (сходимость) аппроксимирующей функции к функции Хевисайда, что значит стремление к нулю ошибки аппроксимации;

– непрерывность аппроксимирующей функции вместе с ее производными;

– близость производной аппроксимирующей функции к производной функции Хевисайда;

– близость аппроксимации функции вероятности к оригинальной функции, что значит стремление к нулю ошибки аппроксимации функции вероятности при стремлении к нулю ошибки аппроксимации функции Хевисайда;

– близость аппроксимации производной функции вероятности к производной оригинальной функции;

– выполнение этих же свойств для аппроксимации квантильной функции как критериальной функции.

3. Сигмоидальная функция и функция вероятности. Погрешность аппроксимации. В основе аппроксимации лежит замена функции Хевисайда в уравнении (3) на сигмоидальную функцию:

$$\Theta(\varphi - \Phi(u,x)) \approx S_\theta(\varphi - \Phi(u,x)) = \frac{1}{1 + e^{-\theta(\varphi - \Phi(u,x))}}, \qquad (4)$$

где θ — большое положительное число.

График сигмоидальной функции с параметром $\theta = 5$ представлен на рисунке 1. Сигмоидальная функция обладает свойствами непрерывности и гладкости, а параметр θ влияет на степень кривизны функции в окрестности нуля: чем выше параметр, тем меньше отрезок перехода от нуля к единице и тем круче график функции.

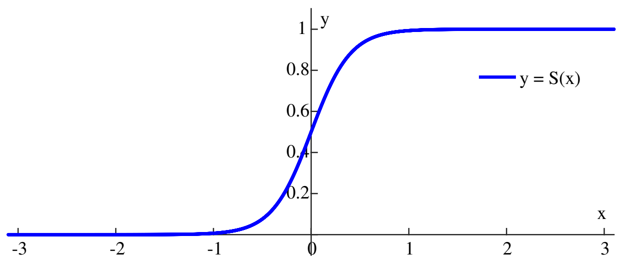

Рис. 1. Сигмоидальная функция

Рассмотрим функции Хевисайда и сигмоиды при различных значениях x случайной величины X.

1) Пусть $x^* \in \Omega^+, \Omega^+ = \{x \mid \Phi(u,x) > \varphi\}$. Тогда

$$\Theta(\varphi - \Phi(u,x^*)) = 1,$$

$$S_\theta(\varphi - \Phi(u, x^*)) = \frac{1}{1 + e^{-\theta(\varphi - \Phi(u, x^*))}} \to 1 \text{ при } \theta \to +\infty.$$

Отсюда следует поточечная сходимость функции сигмоиды к функции Хевисайда при $\theta \to +\infty$, что влечет сходимость почти наверное на данном промежутке.

2) Пусть $x^* \in \Omega^-, \Omega^- = \{x \mid \Phi(u, x) < \varphi\}$. Тогда по аналогии

$$\Theta(\varphi - \Phi(u, x^*)) = 0,$$

$$S_\theta(\varphi - \Phi(u, x^*)) = \frac{1}{1 + e^{-\theta(\varphi - \Phi(u, x^*))}} \to 0 \text{ при } \theta \to +\infty,$$

что влечет сходимость почти наверное на данном промежутке.

3) Пусть $x^0 \in \Omega^0, \Omega^0 = \{x \mid \Phi(u, x) = \varphi\}$. Тогда

$$\Theta(\varphi - \Phi(u, x^0)) = 1,\ S_\theta(\varphi - \Phi(u, x^0)) = \frac{1}{2}. \tag{5}$$

Наложим дополнительное ограничение на функцию $\Phi(u, x)$: пусть теперь $\Phi(u, x)$ является строго кусочно-монотонной функцией по x (в дополнение к ее непрерывности). Это означает, что не существует непустого интервала, на котором функция $\Phi(u, x)$ постоянна, то есть заданный уровень φ пересекается по множеству отделенных друг от друга точек. Тогда множество Ω^0 — дискретное множество, состоящее из не более чем счетного числа точек. Учитывая непрерывность исходной пространственной меры $\mathbb{P}$ и пространства Ω, получим, что при условии на строгую кусочную монотонность $\Phi(u, x)$ по x верно:

$$\mathbb{P}\{\Omega^0\} = 0.$$

Это означает, что при дополнительном условии равенства (5) выполняются на множестве Ω^0 меры нуль, то есть имеет место сходимость почти наверное на всей области Ω.

Таким образом, справедливо следующее утверждение.

Утверждение 1. Пусть X — абсолютно непрерывная случайная величина на полном вероятностном пространстве $(\Omega, \mathscr{F}, \mathbb{P})$. Пусть задана непрерывная строго кусочно-монотонная функция потерь $\Phi(u, X) : U \times \Omega \to \mathbb{R}$, где $U \subset \mathbb{R}^m$ — непустое множество допустимых стратегий u. Также пусть задан уровень потерь φ. Тогда верно

$$S_\theta(\varphi - \Phi(u, X)) \overset{\text{п.н.}}{\to} \Theta(\varphi - \Phi(u, X)) \quad \text{при } \theta \to +\infty, \tag{6}$$

где $S_\theta(\cdot)$ — сигмоидальная функция параметра θ, $\Theta(\cdot)$ — функция Хевисайда.

Теперь рассмотрим функцию вероятности

$$P_\varphi(u) = \mathbb{M}\left[\Theta\left(\varphi - \Phi(u,x)\right)\right] = \int\limits_{-\infty}^{+\infty} \Theta\left(\varphi - \Phi(u,x)\right) f(x)dx \qquad (7)$$

и ее аппроксимацию

$$P_\varphi^\theta(u) = \mathbb{M}\left[S_\theta(\varphi - \Phi(u,x))\right] = \int\limits_{-\infty}^{+\infty} S_\theta(\varphi - \Phi(u,x)) f(x)dx. \qquad (8)$$

Для краткости дальнейших выкладок сделаем замену обозначений, пусть

$$y = \varphi - \Phi(u,x). \qquad (9)$$

Эта замена не подразумевает перехода к новой переменной, а служит только для уменьшения записей. Все отличные от данного аргументы функций Хевисайда и сигмоиды будут указываться непосредственно.

Принимая во внимание то, что $S_\theta(y) \leqslant 1$ при любых значения θ и аргументов, а также учитывая (6), по теореме Лебега о мажорируемой сходимости получим

$$\mathbb{M}\left[\left|S_\theta(y) - \Theta\left(y\right)\right|\right] \to 0 \text{ при } \theta \to +\infty,$$

что эквивалентно

$$\int\limits_{-\infty}^{+\infty} \left|\Theta\left(y\right) - S_\theta(y)\right| f(x)dx \to 0 \text{ при } \theta \to +\infty. \qquad (10)$$

Так как модуль разности интегралов не превосходит интеграла модуля разности, то

$$\left|P_\varphi(u) - P_\varphi^\theta(u)\right| \leqslant \int\limits_{-\infty}^{+\infty} \left|\Theta\left(y\right) - S_\theta(y)\right| f(x)dx. \qquad (11)$$

Далее, из (10) и (11) получаем, что

$$\left|P_\varphi(u) - P_\varphi^\theta(u)\right| \to 0 \text{ при } \theta \to +\infty.$$

Таким образом, доказано следующее утверждение.

Утверждение 2. Пусть X — абсолютно непрерывная случайная величина на полном вероятностном пространстве $(\Omega, \mathscr{F}, \mathbb{P})$. Пусть задана непрерывная строго кусочно-монотонная функция потерь $\Phi(u, X) : U \times \Omega \to \mathbb{R}$, где $U \subset \mathbb{R}^m$ — непустое множество допустимых стратегий u и задан уровень потерь φ. Пусть также выражениями (7) и (8) заданы функции $P_\varphi^\theta(u)$ и $P_\varphi(u)$. Тогда верно

$$P_\varphi^\theta(u) \to P_\varphi(u) \quad \text{при } \theta \to +\infty.$$

Теперь получим выражение для верхней оценки погрешности аппроксимации. Для этого рассмотрим подробнее подынтегральную функцию (11). Множитель $f(x)$ — непрерывная положительная ограниченная функция, не влияющая на знаки неравенств или значения другого сомножителя. Поэтому основная содержательная часть состава погрешности зависит от модуля разности функций Хевисайда и сигмоиды. Таким образом, далее будет рассматриваться функция $g(y) = |\Theta\,(y) - S_\theta(y)|$.

Необходимо сразу уточнить, что значение функции Хевисайда в нуле доопределяется единицей, то есть $\Theta\,(0) = 1$. Это сделано специально для непрерывности и симметричности функции $g(y)$.

Разобьем пространство $\mathbb{R}$ аргумента x на четыре непересекающися области:

$$\chi_1^+ = \{x : 0 \leqslant \varphi - \Phi(u, x) < 1\}, \tag{12}$$

$$\chi_1^- = \{x : -1 < \varphi - \Phi(u, x) < 0\}, \tag{13}$$

$$\chi_2^+ = \{x : \varphi - \Phi(u, x) \geqslant 1\}, \tag{14}$$

$$\chi_2^- = \{x : \varphi - \Phi(u, x) \leqslant -1\}. \tag{15}$$

Учитывая замену (9), эти области отвечают разбиению пространства $\mathbb{R}$ аргумента y функции $g(y)$, то есть например, область χ_1^+ эквивалентна области

$$\chi_1^+ = \{y : 0 \leqslant y < 1\}.$$

Аналогичные размышления относительно разбиения справедливы и для областей из (13), (14), (15).

Отметим, что ввиду непрерывности $\Phi(u, x)$ данные области представляют собой объединение непересекающихся интервалов и не более чем счетного числа точек. С учетом этого факта введем обозначение:

$$\int\limits_{\chi} (\cdot)\, dx = \sum_{j=1}^{n} \int\limits_{x_j}^{x_{j+1}} (\cdot)\, dx, \tag{16}$$

где

$$\chi = \bigcup_{j=1}^{n} (x_j, x_{j+1}),$$

$$\forall j = \overline{1,n}, \forall k = \overline{1,n}, j \neq k : (x_j, x_{j+1}) \cap (x_k, x_{k+1}) = \emptyset.$$

Это означает, что в обозначении (16) «интеграл по области» χ представляет собой сумму интегралов по конечным не пересекающимся интервалам, составляющим эту область. Поскольку интеграл аддитивен, а все выкладки далее применимы сразу к любому элементу из любого интервала фиксированной области, то выкладки справедливы для суммы всех интегралов по всем областям, для которой и введено обозначение (16).

Перейдем к рассмотрению областей.

1) Рассмотрим область χ_2^+, эквивалентную области, задаваемой условием $y \geqslant 1$ для аргумента y. Функция Хевисайда на этой области равна единице. Тогда

$$g(y) = |1 - S_\theta(y)| = 1 - \frac{1}{1 + e^{-\theta \cdot y}}. \tag{17}$$

Из условия задания области следует цепочка неравенств:

$$y \geqslant 1,$$
$$1 + e^{-\theta y} \leqslant 1 + e^{-\theta},$$
$$1 - \frac{1}{1 + e^{-\theta y}} \leqslant 1 - \frac{1}{1 + e^{-\theta}}.$$

Таким образом, на области χ_2^+ выполняется неравенство:

$$g(y) \leqslant 1 - \frac{1}{1 + e^{-\theta}}.$$

Из свойств интеграла и положительности подынтегральной функции следует:

$$\int\limits_{\chi_2^+} g(y)f(x)dx \leqslant \int\limits_{\chi_2^+} \left(1 - \frac{1}{1+e^{-\theta}}\right) f(x)dx.$$

2) Рассмотрим область χ_2^-, эквивалентную области, которая задается условием $y \leqslant -1$ для аргумента y. Функция Хэвисайда на этой области равна нулю. Тогда

$$g(y) = |0 - S_\theta(y)| = S_\theta(y) = \frac{1}{1+e^{-\theta \cdot y}}$$

Аналогично цепочке неравенств для предыдущей области разбиения, из неравенства условия:

$$y \leqslant -1$$

получаем неравенство:

$$\frac{1}{1+e^{-\theta y}} \leqslant \frac{1}{1+e^{\theta}}.$$

Аналогично предыдущей области далее получаем:

$$\int\limits_{\chi_2^-} g(y)f(x)dx \leqslant \int\limits_{\chi_2^-} \left(\frac{1}{1+e^{\theta}}\right) f(x)dx.$$

3) Рассмотрим область χ_1^+, эквивалентную области, которая задается условием $0 \leqslant y < 1$ для аргумента y. Нетрудно убедиться, что функция $g(y)$ является убывающей по y.

Возьмем произвольную точку $y^* \in (0,1)$. Исходя из непрерывности и убывания $g(y)$ следуют неравенства:

$$\forall y \in (0, y^*) \ : \ g(y) \leqslant \frac{1}{2}, \tag{18}$$

$$\forall y \in (y^*, 1) \ : \ g(y) \leqslant g(y^*). \tag{19}$$

Графическая иллюстрация полученных неравенств представлена на рисунке 2.

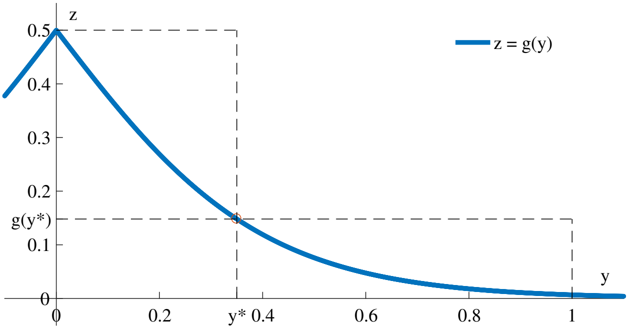

Рис. 2. График функции $g(y)$ для $0 \leqslant y < 1$

Разобьем интеграл функции вероятности на сумму двух интегралов:

$$\int\limits_{\chi_1^{+}} g(y)f(x)dx = \int\limits_{\chi_1^{++}} g(y)f(x)dx + \int\limits_{\chi_1^{+-}} g(y)f(x)dx, \tag{20}$$

где

$$\chi_1^{++} = \{x : y^* < \varphi - \Phi(u,x) < 1\}, \tag{21}$$

$$\chi_1^{+-} = \{x : 0 < \varphi - \Phi(u,x) < y^*\}. \tag{22}$$

Таким образом, области (21) и (22) сопоставимы с ограничениями на y из (18) и (19). Следовательно,

$$\int\limits_{\chi_1^{++}} g(y)f(x)dx \leqslant \int\limits_{\chi_1^{++}} g(y^*)f(x)dx, \tag{23}$$

$$\int\limits_{\chi_1^{++}} g(y^*)f(x)dx = \left(1 - \frac{1}{1+e^{-\theta\cdot y^*}}\right) \int\limits_{\chi_1^{++}} f(x)dx,$$

$$\int\limits_{\chi_1^{+-}} g(y)f(x)dx \leqslant \frac{1}{2} \int\limits_{\chi_1^{+-}} f(x)dx. \tag{24}$$

Так как неравенства выше действуют для любых $y^* \in (0,1)$, возьмем в качестве y^* определенные значения, представляющие собой последовательность, зависящую от θ, причем

$$y^*_\theta = \frac{1}{\sqrt{\theta}}. \tag{25}$$

Тогда из (23) следует:

$$\int\limits_{\chi_1^{++}} g(y) f(x) dx \leqslant \left(1 - \frac{1}{1+e^{-\sqrt{\theta}}}\right) \int\limits_{\chi_1^{++}} f(x) dx.$$

4) Рассмотрим область χ_1^-, эквивалентную области, которая задается условием $-1 < y \leqslant 0$ для аргумента y. Ввиду симметричности функции $g(y)$, все рассуждения для области χ_1^+ переносятся на текущую область симметричным образом с некоторыми отличиями.

Так же как и в предыдущем случае оценивается каждый из интегралов, получаемый при разбиении области χ_1^- на две подобласти χ_1^{-+} и χ_1^{--} точкой y^*, то есть

$$\chi_1^{-+} = \left\{x : y^* < \varphi - \Phi(u, x) < 0\right\},$$

$$\chi_1^{--} = \left\{x : -1 < \varphi - \Phi(u, x) < y^*\right\},$$

но в качестве y^* берется точка

$$y^*_\theta = -\frac{1}{\sqrt{\theta}}.$$

Аналогично предыдущему случаю ищем верхнюю оценку для интегралов, получаем

$$\int\limits_{\chi_1^{--}} g(y) f(x) dx \leqslant \left(\frac{1}{1+e^{\sqrt{\theta}}}\right) \int\limits_{\chi_1^{--}} f(x) dx, \tag{26}$$

$$\int\limits_{\chi_1^{-+}} g(y) f(x) dx \leqslant \frac{1}{2} \int\limits_{\chi_1^{-+}} f(x) dx. \tag{27}$$

В итоге после рассмотрения всех четырех областей получаем следующее утверждение.

Утверждение 3. Пусть условия аналогичны таковым в Утверждении 2. Тогда верхняя оценка погрешности аппроксимации функции вероятности может быть найдена по формуле:

$$
\begin{aligned}
|P_\varphi(u) - P_\varphi^\theta(u)| \leqslant \left(1 - \frac{1}{1+e^{-\theta}}\right) \int\limits_{\chi_2^+} f(x)dx + \left(\frac{1}{1+e^{\theta}}\right) \int\limits_{\chi_2^-} f(x)dx + \\
+ \frac{1}{2} \int\limits_{\chi_1^{+-}} f(x)dx + \left(1 - \frac{1}{1+e^{-\sqrt{\theta}}}\right) \int\limits_{\chi_1^{++}} f(x)dx + \quad (28) \\
+ \frac{1}{2} \int\limits_{\chi_1^{-+}} f(x)dx + \left(\frac{1}{1+e^{\sqrt{\theta}}}\right) \int\limits_{\chi_1^{--}} f(x)dx,
\end{aligned}
$$

где области $\chi_2^+, \chi_2^-, \chi_1^{+-}, \chi_1^{++}, \chi_1^{-+}, \chi_1^{--}$ задаются соответствующими уравнениями.

4. Производные функции вероятности и ее аппроксимации. Далее для удобства и краткости выкладок переобозначим u как одну из компонент u_i многомерного управления $u = [u_1, u_2, ..., u_m]^T$. Таким образом, полученные далее формулы будут справедливы для каждой из компонент вектора управления.

Для дальнейших выкладок необходимо показать некоторые свойства дельта-функции Дирака и производной сигмоидальной функции. Дельта-функция Дирака — обобщенная функция, определяемая следующим образом:

$$
\delta(x) = 0,\ x \neq 0,
$$

$$
\int\limits_{-\infty}^{+\infty} \delta(x)dx = 1.
$$

Известно, что дельта-функция Дирака может быть представлена как предел бесконечной последовательности обычных функций, площадь под кривыми которых равна единице, а значения функций вне точки $x = 0$ стремятся к нулю. Также для дельта-функции выполняется свойство:

$$
\int\limits_{-\infty}^{+\infty} \delta(y - y_0)g(y)dy = g(y_0), \tag{29}
$$

где y_0 есть ноль аргумента дельта-функции.

Производная сигмоидальной функции:

$$\frac{\partial S_\theta(y)}{\partial y} = \left(\frac{1}{1+e^{-\theta y}}\right)' = \theta \cdot S_\theta(y) \cdot (1 - S_\theta(y)) \, .$$

График функции $g(y) = \theta \cdot S_\theta(y) \cdot (1 - S_\theta(y))$ представлен на рисунке 3 (параметр $\theta = 5$).

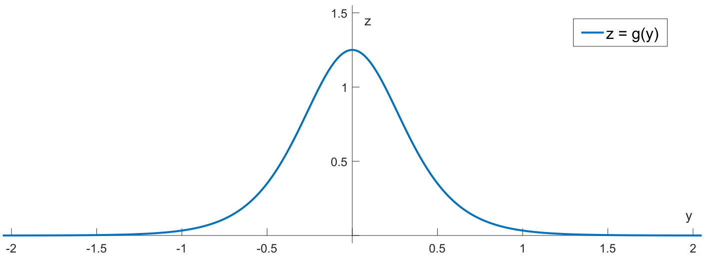

Рис. 3. Производная сигмоидальной функции

При рассмотрении производных сигмоиды при различных θ в качестве элементов последовательности, то есть рассмотрении $\{(S_\theta(y))'\}_{\theta=1}^{\infty}$, выполняется:

$$(S_\theta(y))'\Big|_{y=0} = \theta \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{\theta}{4} \to +\infty \quad \text{при } \theta \to +\infty.$$

Также справедливо

$$(S_\theta(y))'\Big|_{y\neq 0} = \theta \cdot S_\theta(y) \cdot (1 - S_\theta(y)) \to 0 \quad \text{при } \theta \to +\infty,$$

$$\int\limits_{-\infty}^{+\infty} (S_\theta(y))' dy = S_\theta(y)\Big|_{-\infty}^{+\infty} = 1 - 0 = 1.$$

Таким образом, справедливо следующее утверждение.

Утверждение 4. Последовательность производных сигмоидальных функций сходится к дельта-функции Дирака при стремлении параметра сигмоидальной функции к бесконечности, или

$$\lim_{\theta \to +\infty} (S_\theta(y))' = \delta(y). \qquad (30)$$

Таким образом, на последовательность производных сигмоидальной функции в пределе распространяются свойства дельта-функции, в частности свойство (29). Это также фактически означает, что на любом отрезке $[a, b]$, не включающем в себя y_0, интеграл произведения производной сигмоидальной функции и другой функции стремится к нулю.

Перейдем к вопросу о сходимости производной аппроксимированной функции вероятности к производной исходной функции вероятности. Для начала найдем производную исходной функции вероятности по u. Для этого воспользуемся тем, что $\Phi(u, x)$ является гладкой кусочно-монотонной по x функцией. Это позволяет представить выражение для P_φ в виде суммы интегралов особого вида:

$$P_\varphi(u) = \sum_i \int\limits_{x_i^-(u)}^{x_i^+(u)} \Theta(\varphi - \Phi(u, x)) f(x) dx, \tag{31}$$

где $x_i^-(u)$ и $x_i^+(u)$ – границы i-го промежутка монотонности функции $\Phi(u, x)$ при заданном управлении u. Эти точки определяются как точки перегибов и локальных экстремумов функции $\Phi(u, x)$ при заданном u, то есть являются решениями уравнения:

$$\frac{\partial}{\partial x} \Phi(u, x) = 0. \tag{32}$$

В случае линейной по x функции потерь, промежуток монотонности совпадает со всей областью определения функции потерь по второму аргументу. Уравнение (32) задает неявную функцию, а значит, по теореме о неявной функции возможно найти производную границы интервала монотонности по u. Эта производная будет задаваться выражением:

$$\frac{\partial x^*}{\partial u} = -\frac{\Phi''_{ux}(u, x^*)}{\Phi''_{xx}(u, x^*)}, \tag{33}$$

где x^* – решение уравнения (32).

Воспользуемся представлением (31) и запишем производную $P_\varphi(u)$:

$$\frac{\partial}{\partial u} P_\varphi(u) = \sum_i \frac{\partial}{\partial u} \int\limits_{x_i^-(u)}^{x_i^+(u)} \Theta(\varphi - \Phi(u, x)) f(x) dx. \tag{34}$$

На каждом из промежутков монотонности возможен один из четырех вариантов:

1) $\min\limits_{x \in [x_i^-(u); x_i^+(u)]} \Phi(u,x) \geqslant \varphi$.

В этом случае интеграл

$$\int\limits_{x_i^-(u)}^{x_i^+(u)} \Theta(\varphi - \Phi(u,x)) f(x) dx = 0,$$

и его производная на отрезке:

$$\frac{\partial}{\partial u} \int\limits_{x_i^-(u)}^{x_i^+(u)} \Theta(\varphi - \Phi(u,x)) f(x) dx = 0. \qquad (35)$$

2) $\max\limits_{x \in [x_i^-(u); x_i^+(u)]} \Phi(u,x) \leqslant \varphi$.

В этом случае

$$\int\limits_{x_i^-(u)}^{x_i^+(u)} \Theta(\varphi - \Phi(u,x)) f(x) dx = \int\limits_{x_i^-(u)}^{x_i^+(u)} f(x) dx.$$

Производная интеграла по u, согласно формуле Лейбница, равна

$$\frac{\partial}{\partial u} \int\limits_{x_i^-(u)}^{x_i^+(u)} f(x) dx = f(x_i^+(u)) \frac{dx_i^+(u)}{du} - f(x_i^-(u)) \frac{dx_i^-(u)}{du} =$$

$$= f(x_i^-(u)) \frac{\Phi''_{ux}(u, x_i^-(u))}{\Phi''_{xx}(u, x_i^-(u))} - f(x_i^+(u)) \frac{\Phi''_{ux}(u, x_i^+(u))}{\Phi''_{xx}(u, x_i^+(u))}. \qquad (36)$$

3) $\Phi(u, x_i^-(u)) \leqslant \varphi \leqslant \Phi(u, x_i^+(u))$.

В этом случае

$$\int\limits_{x_i^-(u)}^{x_i^+(u)} \Theta(\varphi - \Phi(u,x)) f(x) dx = \int\limits_{x_i^-(u)}^{x_i^*(u)} f(x) dx,$$

где $x_i^*(u)$ – корень уравнения $\Phi(u,x) = \varphi$ на отрезке $[x_i^-(u); x_i^+(u)]$ при заданном u, причем единственный. Производная этого интеграла по u равна:

$$\frac{\partial}{\partial u}\int\limits_{x_i^-(u)}^{x_i^*(u)} f(x)dx = f(x_i^*(u))\frac{dx_i^*(u)}{du} - f(x_i^-(u))\frac{dx_i^-(u)}{du}. \qquad (37)$$

Производная $x_i^*(u)$ по u определяется также с помощью теоремы о неявной функции и равна:

$$\frac{dx_i^*(u)}{du} = -\frac{\Phi'_u(u,x_i^*(u))}{\Phi'_x(u,x_i^*(u))}. \qquad (38)$$

Следовательно,

$$\int\limits_{x_i^-(u)}^{x_i^+(u)} \Theta(\varphi - \Phi(u,x))f(x)dx = f(x_i^-(u))\frac{\Phi''_{ux}(u,x_i^-(u))}{\Phi''_{xx}(u,x_i^-(u))} - \\ - f(x_i^*(u))\frac{\Phi'_u(u,x_i^*(u))}{\Phi'_x(u,x_i^*(u))}. \qquad (39)$$

4) $\Phi(u,x_i^-(u)) \geqslant \varphi \geqslant \Phi(u,x_i^+(u))$.

В этом случае

$$\int\limits_{x_i^-(u)}^{x_i^+(u)} \Theta(\varphi - \Phi(u,x))f(x)dx = \int\limits_{x_i^*(u)}^{x_i^+(u)} f(x)dx, \qquad (40)$$

где $x_i^*(u)$ – также корень уравнения $\Phi(u,x) = \varphi$ на отрезке $[x_i^-(u); x_i^+(u)]$ при заданном u, причем единственный. Производная этого интеграла по

u равна:

$$\frac{\partial}{\partial u}\int\limits_{x_i^*(u)}^{x_i^+(u)} f(x)dx = f(x_i^+(u))\frac{dx_i^+(u)}{du} - f(x_i^*(u))\frac{dx_i^*(u)}{du} =$$

$$= f(x_i^*(u))\frac{\Phi'_u(u, x_i^*(u))}{\Phi'_x(u, x_i^*(u))} - f(x_i^+(u))\frac{\Phi''_{ux}(u, x_i^+(u))}{\Phi''_{xx}(u, x_i^+(u))}. \quad (41)$$

Таким образом, производная функции вероятности по u определена на всем множестве значений x.

Проведем аналогичные выкладки для аппроксимации функции вероятности. Рассматриваемая производная определяется аналогично выкладкам для исходной функции и в итоге получаем аналог формулы (34):

$$\frac{\partial}{\partial u}P_\varphi^\theta(u) = \sum_i \frac{\partial}{\partial u}\int\limits_{x_i^-(u)}^{x_i^+(u)} S_\theta(\varphi - \Phi(u,x))f(x)dx. \quad (42)$$

Рассмотрим аналогичные исходной функции вероятности промежутки монотонности.

1) $\min\limits_{x\in[x_i^-(u);x_i^+(u)]} \Phi(u,x) \geqslant \varphi$.

В этом случае при росте θ интеграл стремится к нулю, следовательно,

$$\frac{\partial}{\partial u}\int\limits_{x_i^-(u)}^{x_i^+(u)} S_\theta(\varphi - \Phi(u,x))f(x)dx \to 0 \quad \text{при } \theta \to +\infty, \quad (43)$$

что сопоставимо с результатом (35).

2) $\max\limits_{x\in[x_i^-(u);x_i^+(u)]} \Phi(u,x) \leqslant \varphi$.

В этом случае производная интеграла по u равна:

$$\frac{\partial}{\partial u}\int\limits_{x_i^-(u)}^{x_i^+(u)} S_\theta(\varphi-\Phi(u,x))f(x)dx =$$

$$= S_\theta(\varphi-\Phi(u,x_i^+))f(x_i^+(u))\frac{dx_i^+(u)}{du} - S_\theta(\varphi-\Phi(u,x_i^-))f(x_i^-(u))\frac{dx_i^-(u)}{du} -$$

$$- \int\limits_{x_i^-(u)}^{x_i^+(u)} f(x)S'_\theta(\varphi-\Phi(u,x))\Phi'_u(u,x)dx. \quad (44)$$

Значения множителей $S_\theta(\varphi-\Phi(u,x_i^+))$ и $S_\theta(\varphi-\Phi(u,x_i^-))$ стремятся к единице, а последнее слагаемое, согласно свойствам производной сигмоиды, стремится к нулю при $\theta\to+\infty$. Следовательно,

$$\frac{\partial}{\partial u}\int\limits_{x_i^-(u)}^{x_i^+(u)} S_\theta(\varphi-\Phi(u,x))f(x)dx \to f(x_i^-(u))\frac{\Phi''_{ux}(u,x_i^-(u))}{\Phi''_{xx}(u,x_i^-(u))} -$$

$$- f(x_i^+(u))\frac{\Phi''_{ux}(u,x_i^+(u))}{\Phi''_{xx}(u,x_i^+(u))} \quad \text{при } \theta\to+\infty,$$

или, сопоставляя полученное выражение и (36), получим:

$$\frac{\partial}{\partial u}\int\limits_{x_i^-(u)}^{x_i^+(u)} S_\theta(\varphi-\Phi(u,x))f(x)dx \to \frac{\partial}{\partial u}\int\limits_{x_i^-(u)}^{x_i^+(u)} \Theta(\varphi-\Phi(u,x))f(x)dx. \quad (45)$$

3) $\Phi(u,x_i^-(u))\leqslant\varphi\leqslant\Phi(u,x_i^+(u))$.

В этом случае существует $x_i^*(u)$ – корень уравнения $\Phi(u,x)=\varphi$ на отрезке $[x_i^-(u);x_i^+(u)]$ при заданном u, причем единственный. Производная интеграла по u определяется согласно формуле (44), и на этом промежутке монотонности выполняются следующие соотношения:

$$S_\theta(\varphi-\Phi(u,x_i^+))\to 0 \quad \text{при } \theta\to+\infty, \quad (46)$$

$$S_\theta(\varphi-\Phi(u,x_i^-))\to 1 \quad \text{при } \theta\to+\infty. \quad (47)$$

В последнем слагаемом (44), а именно интеграле, сделаем замену переменных

$$t = \Phi(u, x), \quad dt = \Phi'_x(u, x)dx, \tag{48}$$

получаем:

$$\int\limits_{x_i^-(u)}^{x_i^+(u)} f(x)S'_\theta(\varphi - \Phi(u,x))\Phi'_u(u,x)dx = \int\limits_{t^-}^{t^+} f(x(t))S'_\theta(\varphi - t)\frac{\Phi'_u(u, x(t))}{\Phi'_x(u, x(t))}dt,$$

где $t^- = t(x_i^-(u)),\ t^+ = t(x_i^+(u))$. Исходя из свойств производной сигмоиды, значение этого интеграла стремится к значению подынтегральной функции в точке $t = \varphi$, которая по определению эквивалентна точке $x = x_i^*(u)$. Таким образом, верно

$$\int\limits_{x_i^-(u)}^{x_i^+(u)} f(x)S'_\theta(\varphi - \Phi(u,x))\Phi'_u(u,x)dx \to f(x_i^*(u))\frac{\Phi'_u(u, x_i^*(u))}{\Phi'_x(u, x_i^*(u))}. \tag{49}$$

Рассматривая поведение (44) с учетом (46), (47), (49), получим, что

$$\frac{\partial}{\partial u}\int\limits_{x_i^-(u)}^{x_i^+(u)} S_\theta(\varphi - \Phi(u,x))f(x)dx \to f(x_i^-(u))\frac{\Phi''_{ux}(u, x_i^-(u))}{\Phi''_{xx}(u, x_i^-(u))} -$$
$$- f(x_i^*(u))\frac{\Phi'_u(u, x_i^*(u))}{\Phi'_x(u, x_i^*(u))} \quad \text{при } \theta \to +\infty,$$

а отсюда следует выполнение соотношения (45) для этого промежутка монотонности.

4) $\Phi(u, x_i^-(u)) \geqslant \varphi \geqslant \Phi(u, x_i^+(u))$.

Действуем по аналогии с прошлым случаем; получаем, что существует единственный корень $x_i^*(u)$ уравнения $\Phi(u,x) = \varphi$ на данном отрезке монотонности, производная определяется по формуле (44) и что верны выражения:

$$S_\theta(\varphi - \Phi(u, x_i^+)) \to 1 \quad \text{при } \theta \to +\infty,$$

$$S_\theta(\varphi - \Phi(u, x_i^-)) \to 0 \quad \text{при } \theta \to +\infty.$$

В интегральном слагаемом предполагается замена (48), и по свойству производной сигмоиды получаем:

$$\int\limits_{x_i^-(u)}^{x_i^+(u)} f(x)S'_\theta(\varphi - \Phi(u,x))\Phi'_u(u,x)dx \to f(x_i^*(u))\frac{\Phi'_u(u,x_i^*(u))}{\Phi'_x(u,x_i^*(u))}. \quad (50)$$

В итоге, собрав все соотношения, получим:

$$\frac{\partial}{\partial u}\int\limits_{x_i^-(u)}^{x_i^+(u)} S_\theta(\varphi - \Phi(u,x))f(x)dx \to -f(x_i^+(u))\frac{\Phi''_{ux}(u,x_i^+(u))}{\Phi''_{xx}(u,x_i^+(u))}+$$

$$+ f(x_i^*(u))\frac{\Phi'_u(u,x_i^*(u))}{\Phi'_x(u,x_i^*(u))} \quad \text{при } \theta \to +\infty,$$

а отсюда следует выполнение соотношения (45) для этого промежутка монотонности.

Таким образом, рассмотрев все промежутки монотонности, было доказано следующее утверждение.

Утверждение 5. Пусть условия аналогичны таковым в Утверждении 2, тогда последовательность частных производных по управлению аппроксимированной вероятностной функции стремится к частной производной по управлению исходной вероятностной функции, т.е.

$$\frac{\partial P_\varphi^\theta(u)}{\partial u} \to \frac{\partial P_\varphi(u)}{\partial u} \quad \text{при } \theta \to +\infty. \quad (51)$$

Для получения удобной формулы производной аппроксимации вероятностной функции рассмотрим математическое ожидание:

$$\mathbb{M}\left[G(u_0,u_1,x)\right] = \mathbb{M}\left[\frac{S_\theta(\varphi - \Phi(u_1,x)) - S_\theta(\varphi - \Phi(u_0,x))}{u_1 - u_0}\right], \quad (52)$$

где $u_1 = u_0 + \Delta u, \Delta u > 0$, и представление сигмоидальной функции

$$S_\theta(\varphi - \Phi(u,x)) = \int\limits_{-\infty}^{u} h(\beta,x)d\beta,$$

где

$$h(\beta, x) = \theta S_\theta(\varphi - \Phi(\beta, x))\,(1 - S_\theta(\varphi - \Phi(\beta, x)))\,(-\Phi'_\beta(\beta, x)).$$

Тогда из (52) следует:

$$\mathbb{M}\left[G(u_0, u_1, x)\right] = \mathbb{M}\left[\frac{1}{u_1 - u_0}\int\limits_{u_0}^{u_1} h(\beta, x)d\beta\right]. \tag{53}$$

Используя свойства индикаторной функции, свойства интегралов и теорему Фубини, доказывается равномерная интегрируемость семейства функций $\{G(u_0, u_1, x) : u_1 > u_0, u_0 \in U, u_1 \in U\}$. Из этого следует, что

$$\frac{\partial}{\partial u}\mathbb{M}\int\limits_{-\infty}^{u} h(\beta, x)d\beta = \mathbb{M}h(\beta, x).$$

Переходя к принятым ранее обозначениям, получим:

$$\frac{\partial}{\partial u}P_\varphi^\theta(u) = \mathbb{M}[\theta S_\theta(\varphi - \Phi(u, x))\,(1 - S_\theta(\varphi - \Phi(u, x)))\,(-\Phi'_u(u, x))],$$

или, в конечном виде:

$$\frac{\partial P_\varphi^\theta(u)}{\partial u} = \int\limits_{-\infty}^{+\infty} \theta S_\theta(y(u, x))\,(S_\theta(y(u, x)) - 1)\,\Phi'_u(u, x)f(x)dx, \tag{54}$$

где $y(u, x) = \varphi - \Phi(u, x)$.

Данную формулу можно использовать при решении задачи (1) с помощью методов оптимизации первого порядка.

По аналогии с формулой (54) вычисляется производная аппроксимации функции вероятности по параметру φ:

$$\frac{\partial P_\varphi^\theta(u)}{\partial \varphi} = \int\limits_{-\infty}^{+\infty} \theta S_\theta(\varphi - \Phi(u, x))\,(1 - S_\theta(\varphi - \Phi(u, x)))\,f(x)dx, \tag{55}$$

Так как все свойства, соотношения дельта-функции Дирака и рассуждения переносятся на случай производной по параметру φ без изменений, то при помощи рассуждений, аналогичных таковым при получении

(51), справедливо следующее:

$$\frac{\partial P_\varphi^\theta(u)}{\partial\varphi} \to \frac{\partial P_\varphi(u)}{\partial\varphi} \quad \text{при } \theta \to +\infty. \tag{56}$$

Данное соотношение необходимо для получения выражения для производной функции квантили.

5. Квантильный критерий. Аппроксимация. Согласно одномерному аналогу формулы представления градиента функции квантили из [6], справедливо равенство

$$\frac{\partial}{\partial u}\varphi_\alpha(u) = -\frac{\left.\dfrac{\partial}{\partial u}P_\varphi(u)\right|_{\varphi=\varphi_\alpha(u)}}{\left.\dfrac{\partial}{\partial\varphi}P_\varphi(u)\right|_{\varphi=\varphi_\alpha(u)}}. \tag{57}$$

Подставив в (57) выражения (54) и (55), получим выражение для аппроксимации производной квантили по управлению:

$$\frac{\partial\varphi_\alpha^\theta(u)}{\partial u} \approx \frac{\int\limits_{-\infty}^{+\infty} S_\theta(y(u,x))\left(1 - S_\theta(y(u,x))\right)\Phi'_u(u,x)f(x)dx}{\int\limits_{-\infty}^{+\infty} S_\theta(y(u,x))\left(1 - S_\theta(y(u,x))\right)f(x)dx}, \tag{58}$$

где $y(u,x) = \varphi_\alpha(u) - \Phi(u,x)$.

Данное выражение может быть использовано при решении задачи (2) методами оптимизации первого порядка.

Согласно определению квантили и функции распределения, справедливо равенство:

$$F_X(\varphi_\alpha) = \alpha,$$

где φ_α — квантиль уровня α случайной величины X с функцией распределения $F_X(x)$ и плотностью $f_X(x)$. Продифференцируем равенство по α в обеих частях, получим:

$$f_X(\varphi_\alpha)\cdot\frac{d\varphi_\alpha}{d\alpha} = 1,$$

или

$$\frac{d\varphi_\alpha}{d\alpha} = \frac{1}{f_X(\varphi_\alpha)}.$$

Учитывая выражение (55), получим уравнение апроксимации производной квантили по уровню для нашего случая

$$\frac{\partial \varphi_\alpha^\theta(u)}{\partial \alpha} = \frac{1}{\int\limits_{-\infty}^{+\infty} \theta S_\theta(y(u,x))\left(1 - S_\theta(y(u,x))\right) f(x)dx}, \tag{59}$$

где $y(u,x) = \varphi_\alpha(u) - \Phi(u,x)$.

Таким образом, оценив или определив значение квантили для заданных u и α, можно вычислить квантиль определенного уровня и использовать градиентные методы оптимизации для решения исходной задачи в квантильной постановке.

Доказательство сходимости аппроксимации квантили и ее производных к исходным функциям выходит за рамки настоящей статьи и будет исследовано в дальнейших работах.

6. Примеры применения. В этом разделе приведены 3 примера применения полученных оценок градиента к решению задач стохастического программирования. Первый пример показывает, что полученное с помощью градиентного спуска решение соответствует точному, которое может быть найдено аналитически. Второй пример показывает возможность применения оценок градиента для решения задачи с двумерным случайным вектором, хотя вопрос точности и сходимости оценок в случае многомерного распределения остается предметом дальнейшего исследования. Третий пример показывает возможность применения оценок в прикладной задаче: задаче минимизации площади взлетно-посадочной полосы при ограничении на вероятность успешной посадки.

Пример 1. Пусть функция потерь задана следующим образом:

$$\Phi(u, X) = u_0^2 + (u_1 - 1)^2 \cdot X + X,$$

где $[u_0, u_1]^T \in [-a, a] \times [-a, a],\ a > 1$, а случайная величина X имеет равномерное распределение на $[1, 3]$:

$$X \sim \mathbb{U}[1, 3].$$

Зафиксируем уровень $\alpha = 0.75$ и рассмотрим задачу оптимизации квантили:

$$\varphi_{0.75}(u) \to \min_u .$$

Нетрудно заметить, что при таких ограничениях на управление и соответствующей целевой функции минимум квантили достигается при управлении

$$u_{opt} = [0, 1]^T,$$

а само значение квантили в этой точке равно:

$$\varphi_{0.75}(u_{opt}) = 2.5.$$

Решать задачу будем методом градиентного спуска с фиксированным шагом. Для начала зададим начальную точку u_0 и найдем значение квантили в этой точке. В этой задаче значение можно вычислить непосредственно; в более сложных задачах подразумевается оценка начального приближения, например методом Монте-Карло. Итак, пусть

$$u_0 = [1, 0]^T, \quad \varphi_{0.75}(u_0) = 6.$$

Градиент квантили, являющийся основой алгоритма, вычисляется на каждом шаге в соответствии с формулой (58). Интегралы вычисляются приближенно при помощи известных численных методов. Зададим величину шага алгоритма $h = 0.001$, параметр остановки алгоритма $\varepsilon = 0.00001$ и параметр сигмоидальной функции $\theta = 20$. Алгоритм был реализован при помощи пакетов $scipy$ и $numpy$ языка Python.

В результате работы алгоритма было получено следующее решение:

$$u^{\theta}_{opt} = [0.0019, 1.0000]^T, \quad \varphi_{0.75}(u^{\theta}_{opt}) = 2.4982.$$

Количество итераций при этом составило около 2600, что является неплохим показателем скорости сходимости, учитывая шаг алгоритма. Точность аппроксимации может быть увеличена за счет изменения параметров h и ε алгоритма, а также за счет увеличения параметра θ.

Пример 2. Обобщение на случайный вектор произвольной размерности X. Пусть функция потерь задана следующим образом:

$$\Phi(u, X) = u_0^2 + (u_1 - 1)^2 \cdot X_1 + (u_2 - 1)^2 \cdot X_2 + X_1 + X_2,$$

где $[u_0, u_1, u_2]^T \in [-a, a] \times [-a, a] \times [-a, a],\ a > 1$, а случайные величины X_1 и X_2 независимы и имеют распределения:

$$X_1 \sim \mathbb{U}[1, 3], \quad X_2 \sim \mathbb{U}[1, 3].$$

То есть по сути задана случайная величина $X = [X_1, X_2]$. Рассматривать будем задачу, аналогичную таковой в предыдущем примере — оптимизация по квантильному критерию с уровнем $\alpha = 0.75$. По аналогии точка оптимума и значение в этой точке находятся непосредственно:

$$u_{opt} = [0, 1, 1]^T, \quad \varphi_{0.75}(u_{opt}) = 4.5.$$

Алгоритм решения остается неизменным, с тем лишь отличием, что в аналоге (58) для текущего случая будут вычисляться двойные интегралы по каждому из возможных значений X_1 и X_2, а совместная плотность распределения $f_X(x_1, x_2)$, входящая в вычисляемые выражения, ввиду независимости величин будет равна:

$$f_X(x_1, x_2) = \begin{cases} \frac{1}{4}, & (x_1, x_2) \in [1, 3] \times [1, 3], \\ 0, & (x_1, x_2) \notin [1, 3] \times [1, 3]. \end{cases}$$

Для выбранной начальной точки u_0 квантиль оценим непосредственно:

$$u_0 = [1, 0, 0]^T, \quad \varphi_{0.75}(u_0) = 10.$$

Параметры алгоритма и сигмоиды оставим аналогичные предыдущему примеру.

В результате работы алгоритма было получено следующее решение:

$$u^{\theta}_{opt} = [0.009, 0.998, 0.998]^T, \quad \varphi_{0.75}(u^{\theta}_{opt}) = 4.509.$$

Количество итераций составляло порядка 12000. Как и в предыдущем случае, лучшей точности можно добиться, управляя параметрами алгоритма и сигмоидальной функции.

Пример 3. Задача оптимизации площади взлетно-посадочной полосы. Проблема оптимизации площади взлетно-посадочной полосы (ВПП) связана с двумя факторами:

– действием порывов ветра при посадке, что влияет на точку касания летательным аппаратом (ЛА) полотна, поэтому с точки зрения безопасности посадки полотно ВПП должно быть максимально обширным;

– дороговизной покрытия квадратного метра полотна ВПП, поэтому с этой точки зрения площадь ВПП должна быть минимально возможной.

В результате учета этих двух факторов и возникает задача оптимизации площади взлетно-посадочной полосы: найти минимально возможную площадь ВПП, при которой гарантируется некий установленный уровень безопасности при посадке ЛА. Суть задачи сводится в поиске оптимальных длины и ширины полотна взлетно-посадочной полосы, при которых выполняются вероятностные ограничения по безопасности, то есть превышается заданная вероятность успешной посадки.

Пусть O — расчетная точка касания самолетом ВПП; l_0 — номинальная длина пробега ЛА до полной остановки при отсутствии внешних возмущений; z_0 — половина расстояния между посадочными шасси. Учет требований надежности осуществляется путем введения величин: l_1 — запас по длине ВПП на случай недолета до точки O; l_2 — запас по длине ВПП на случай перелета; z_1 — половина ширины ВПП. Все описанные величины представлены на рисунке 4.

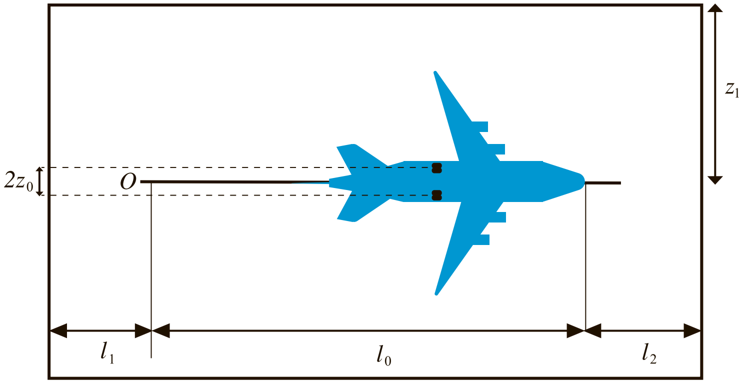

Рис. 4. Параметры ВПП и ЛА при посадке

Площадь ВПП определяется выражением:

$$S(l_1, l_2, z_1) = 2z_1(l_0 + l_1 + l_2).$$

На параметры из физического смысла задачи накладываются ограничения:

$$l_1 \geqslant 0, \quad l_2 \geqslant 0, \quad z_1 \geqslant z_0. \tag{60}$$

В целях упрощения модели предполагаем, что $z_0 = 0$.

Пусть X и Z — случайные отклонения от точки касания O вдоль осей Ox и Oy соответственно. Эти отклонения связаны с компонентами W_x и W_y соотношениями:

$$X = a_{11}W_x + a_{12}|W_z|, \quad Z = a_{22}W_z,$$

где a_{11}, a_{12}, a_{22} — известные коэффициенты. Предположим, что W_x и W_z — независимые нормально распределенные случайные величины с известными параметрами:

$$W_x \sim \mathbb{N}(m_x, \sigma_x^2), \quad W_z \sim \mathbb{N}(m_z, \sigma_z^2).$$

Посадка считается успешной, если

$$-l_1 \leqslant X \leqslant l_2, \quad |Z| \leqslant z_1.$$

Таким образом, требование по безопасности посадки можно сформулировать как

$$\mathbb{P}\{-l_1 \leqslant X \leqslant l_2, |Z| \leqslant z_1\} \geqslant \alpha, \tag{61}$$

где α — доверительная вероятность, отражающая заданный уровень надежности.

Таким образом, задача оптимизации описывается как:

$$S(l_1, l_2, z_1) \to \min_{l_1, l_2, z_1} \tag{62}$$

при ограничениях (60), (61).

В книге [6] авторами приведена задача, эквивалентная задаче (62), но в постановке квантильной оптимизации, и при помощи аналитического алгоритма на основе доверительного метода определено оптимальное решение для различных значений параметра a_{12}. В частности, при $a_{12} = 0$ приближенное оптимальное решение задачи выглядит как:

$$l_1 = l_2 = k_2, \quad z_1 = \frac{\gamma}{2}, \tag{63}$$

где

$$k_2 = |a_{11}|\sigma_x r_\alpha, \quad \gamma = 2|a_{22}|\sigma_z r_\alpha, \quad r_\alpha = \sqrt{-2\ln(1-\alpha)}.$$

Авторами [6] также отмечено, что использование метода Монте-Карло или прямого метода оценки уровня квантили связано с огромными вычислительными затратами, а получение точного решения невозможно в силу зависимости выпуклости целевой функции от ее случайных компонент.

Исходные параметры задачи были заданы следующим образом:

$$\alpha = 0.95, \quad l_0 = 150, \quad a_{11} = -2, \quad a_{12} = 0, \quad a_{22} = 1,$$

$$m_x = m_z = 0, \quad \sigma_x = \sigma_z = 3.$$

При таких значениях параметров, согласно (63), приближенное оптимальное решение равно:

$$l_1^* = l_2^* = 14.687, \quad z_1^* = 7.343, \quad S(l_1^*, l_2^*, z_1^*) = 1900.034.$$

Решение задачи (62) при ограничениях (60), (61) определяется при помощи метода доверительных областей (*trust-region*), где градиент вероятностного ограничения вычисляется в соответствии с формулой (54). Ввиду независимости полученных величин X и Z, их совместная плотность является произведением плотностей каждой из компонент с поправкой на коэффициенты a_{11} и a_{22}. Параметры алгоритма были заданы следующим образом: минимально допускаемые значения для градиента и радиуса доверительной области равны значению параметра остановки алгоритма $\varepsilon = 0.0001$, параметр сигмоидальной функции $\theta = 10$. Значение объемных интегралов определяется численно с точностью $\varepsilon_{int} = 0.0001$. Первоначальной точкой алгоритма была выбрана точка, эквивалентная значению переменных

$$l_1^0 = 19.163, \quad l_2^0 = 9.581 \quad z_1^0 = 32.186, \quad S(l_1^0, l_2^0, z_1^0) = 8287.491.$$

В результате работы алгоритма было получено следующее решение:

$$l_1^\theta = l_2^\theta = 16.896, \quad z_1^\theta = 7.946, \quad S(l_1^\theta, l_2^\theta, z_1^\theta) = 2126.235.$$

Число итераций алгоритма составило порядка 900. Лучшей точности можно добиться, управляя параметрами алгоритма и сигмоидальной функции.

На основе примеров 2 и 3 можно высказать предположение, что верны обобщения полученных формул на случай многомерности исходной случайной величины X. Однако эта тема требует более основательного исследования и выходит за рамки данной статьи.

Также стоит отметить, что многие схожие с рассмотренной в примере 3 задачи могут решаться с помощью описанных алгоритмов аппроксимации. В первую очередь это задачи, сводимые к задачам квантильной или вероятностной оптимизации или заданные с вероятностными ограничениями, такие как проектирование системы водоснабжения, формирование портфеля ценных бумаг, планирование бюджета госпиталя и другие, для которых проблематично найти аналитическое решение или методы нахождения приближенного решения связаны с вычислительными трудностями.

7. Заключение. Основным объектом исследования является алгоритм аппроксимации вероятностных критериальных функций в задачах стохастического программирования. Научная новизна результатов заключается в получении новых выражений для оценки функции вероятности, ее производных и производных функции квантили, а также получение выражения для оценки погрешности функции вероятности в форме объемных интегралов.

Полученные в статье оценки могут быть использованы при решении задач стохастического программирования при помощи численных методов оптимизации первого порядка, особенно в случаях, когда другие методы могут быть не применимы. Основным преимуществом такого способа решения является относительная вычислительная простота по сравнению, например, с доверительным методом, широко применяющимся при решении задач стохастического программирования. Вычисление градиента функции вероятности указанным способом также является менее трудоемким по сравнению с вычислением градиента в форме поверхностного интеграла.

Перспективным видится дальнейшее развитие идей, изложенных в настоящей статье. Предполагается доказательство сходимости аппроксимации функции квантили и ее производной. Также предполагается улучшение верхней оценки погрешности аппроксимации функции вероятности и оценка погрешности аппроксимации квантили. Дополнительно планируется рассмотрение условий, накладываемых на целевую функцию, и разработка алгоритма решения задач с вероятностными ограничениями. Наконец, наиболее масштабным и актуальным видится исследование, посвященное обобщению данного вида аппроксимации на случай многомерных случайных величин.

Литература

1. *Marchisio M.A.* Stochastic Modeling // Introduction to Synthetic Biology. 2018. pp. 39–52.
2. *Алешкин Н.А., Алешкин А.П., Карпова И.Р.* Модель оптимизации ресурсов в малосерийном производственном цикле при использовании нечеткой логики // Вопросы радиоэлектроники. 2018. № 9. С. 84–90.
3. *Мазелис Л.С., Солодухин К.С., Чен А.Я.* Нечеткие модели оптимизации портфеля проектов университета с учетом влияния на характеристики отношений со стейкхолдерами // Университетское управление: практика и анализ. 2017. Т. 21. № 5(111). С. 51–63.
4. *Бабуров В.И., Иванцевич Н.В., Олянюк П.В., Саута О.И.* Нечеткие критерии и понятия в задаче выбора приемника ГНСС для навигационно-пилотажного комплекса // Труды СПИИРАН. 2012. Вып. 23. С. 357–368.
5. *Фильченков А.А.* Меры истинности и вероятностные графические модели для представления знаний с неопределенностью // Труды СПИИРАН. 2012. Вып. 23. С. 254–295.

6. *Кибзун А.И., Кан Ю.С.* Задачи стохастического программирования с вероятностными критериями // М.: Физматлит. 2009. 372 с.
7. *Dentcheva D., Penev S., Ruszczyński. A.* Statistical estimation of composite risk functionals and risk optimization problems // Annals of the Institute of Statistical Mathematics. 2017. vol. 69. no 4. pp. 737–760.
8. *Dentcheva D., Ruszczyński. A.* Risk forms: representation, disintegration, and application to partially observable two-stage systems // Mathematical Programming. 2019. pp. 1–21.
9. *Кибзун А.И., Наумов А.В., Норкин В.И.* О сведении задачи квантильной оптимизации с дискретным распределением к задаче смешанного целочисленного программирования // Автоматика и телемеханика. 2013. № 6. С. 66–86.
10. *Кибзун А.И., Игнатов А.Н.* Сведение двухшаговой задачи стохастического оптимального управления с билинейной моделью к задаче смешанного целочисленного линейного программирования // Автоматика и телемеханика. 2016. № 12. С. 89–111.
11. *Иванов С.В., Кибзун А.И.* Выборочная аппроксимация двухэтапной задачи стохастического линейного программирования с квантильным критерием // Труды Института математики и механики УрО РАН. 2017. Т. 23. № 3. С. 134–143.
12. *Кибзун А.И., Соболь В.Р.* Двухшаговая задача хеджирования европейского колл-опциона при случайной длительности транзакций // Труды Института математики и механики УрО РАН. 2015. Т. 21. № 3. С. 164–174.
13. *Райк Э.* Дифференцируемость по параметру функции вероятности и стохастический псевдоградиентный метод для ее оптимизации // Известия Академии наук Эстонской ССР: физика, математика. 1975. № 1. С. 3–8.
14. *Кибзун А.И., Третьяков Г.Л.* Дифференцируемость функции вероятности. // Доклады Академии наук. 1997. № 2. С. 159–161.
15. *Uryasev S.*. Derivatives of probability functions and some applications. // Annals of Operations Research. 1995, Vol. 56. № 1. pp. 287–311.
16. *Uryasev S.*. Derivatives of probability functions and integrals over sets given by inequalities. // Journal of Computational and Applied Mathematics. 1994. vol. 56. no. 1. pp. 197–223.
17. *Henrion R.* Gradient estimates for Gaussian distribution functions: application to probabilistically constrained optimization problems // Numerical Algebra, Control and Optimization. 2012. vol. 2. no. 4. pp. 655–668.
18. *van Ackooij W., Henrion R.* (Sub-)Gradient Formulae for Probability Functions of Random Inequality Systems under Gaussian Distribution // SIAM/ASA Journal on Uncertainty Quantification. 2017. vol. 5. no. 1. pp. 63–87.
19. *Pflug G.C., Weisshaupt H.* Probability gradient estimation by set-valued calculus and applications in network design // SIAM Journal on Optimization. 2005. vol. 15. pp. 898–914.
20. *Yu C., Zelterman D.* A general approximation to quantiles. // Communications in Statistics-Theory and Methods. 2017. vol. 46(19). pp. 9834–9841.
21. *Okagbue H.I., Adamu M.O., Anake T.A.* Ordinary Differential Equations of the Probability Functions of the Weibull Distribution and their Application in Ecology. // International Journal of Engineering and Future Technology. 2018. vol. 15. pp. 57–78.
22. *Garniera J., Omraneb A., Rouchdyc Y.* Asymptotic formulas for the derivatives of probability functions and their Monte Carlo estimations // European Journal of Operational Research. 2009. vol. 198. no. 3. pp. 848–858.
23. *Mafusalov A., Uryasev S.* Buffered Probability of Exceedance: Mathematical Properties and Optimization // SIAM Journal on Optimization. 2018. vol. 28. no. 2. pp. 1077–1103.
24. *Zhang T., Uryasev S., Guan Y.* Derivatives and subderivatives of buffered probability of exceedance // Operations Research Letters. 2019. vol. 47. no. 2. pp 130–132.

25. *Ширяев А.Н.* Вероятность // М.:Наука. 2007. 552 с.

Соболь Виталий Романович — канд. физ.-мат. наук, ассистент, кафедра теории вероятностей и компьютерного моделирования, Московский авиационный институт (национальный исследовательский университет) (МАИ). Область научных интересов: математическая статистика, теория вероятности, математическое моделирование экономических процессов, стохастическое программирование. Число научных публикаций — 12. vitsobol@mail.ru; Волоколамское ш., 4, 125993, Москва, Российская Федерация; р.т.: +7 499 158-48-74.

Торишный Роман Олегович — аспирант, кафедра теории вероятностей и компьютерного моделирования, Московский авиационный институт (национальный исследовательский университет) (МАИ). Область научных интересов: математическая статистика, теория вероятности, математическое моделирование, искусственный интеллект. Число научных публикаций — 5. Arenas-26@yandex.ru; Волоколамское ш., 4, 125993, Москва, Российская Федерация; р.т.: +7 968 622-32-39.

Поддержка исследований. Работа выполнена в соответствии с государственным заданием № 9.7555.2017/БЧ.

DOI 10.15622/sp.2020.19.1.7

V.R. Sobol, R.O. Torishnyi

APPLICATION OF SMOOTH APPROXIMATION OF PROBABILITY FUNCTION AND QUANTILE FUNCTION IN SOLVING STOCHASTIC PROGRAMMING PROBLEMS

Sobol V.R., Torishnyi R.O. **Application of smooth approximation of probability function and quantile function in solving stochastic programming problems.**

Abstract. One of the possible variants of smooth approximation of probability criteria in stochastic programming problems allowing to obtain estimates of the probability function gradient and the quantile function gradient in the form of a volume integral is considered. The research is applied to problems of probability function maximization and quantile function minimization for the loss functional depending on the control vector and one-dimensional absolutely continuous random variable.

The main idea of the approximation is to replace the discontinuous Heaviside function in the integral representation of the probability function with a smooth function having such properties as continuity, smoothness, and easily computable derivatives. An example of such function is the distribution function of a random variable distributed according to the logistic law with zero mean and finite dispersion, which is a sigmoid. The value inversely proportional to the root of the variance is a parameter that provides the proximity of the original function and its approximation. This replacement allows to obtain a smooth approximation of the probability function, and for this approximation derivatives by the control vector and by other parameters of the problem can be easily found.

The main result of the article is the obtained expressions for approximation of the probability function derivatives by the control vector and by the acceptable loss level, as well as expressions for approximation of the quantile function gradient in the form of volume integrals. The article proves the convergence of the probability function approximation obtained by replacing the Heaviside function with the sigmoidal function to the original probability function, and the error estimate of such approximation is obtained. The convergence of the approximation of probability function derivatives to the true derivatives under a number of conditions on the loss functional is also proved.

Examples are considered to demonstrate the possibility of applying the proposed estimates to the solution of stochastic programming problems with criteria in the form of a probability function and a quantile function, including the case of a multidimensional random variable.

Keywords: stochastic programming, probability criteria, quantile criteria, approximation, numerical methods, sigmoidal function.

Sobol Vitaliy — Ph.D., Research Officer, Department of Probability Theory and Computer Modelling, Moscow Aviation Institute (National Research University). Research interests: mathematical statistics, probability theory, mathematical modeling of economic processes, stochastic programming. The number of publications — 12. vitsobol@mail.ru; 4, Volokolamskoe sh., 125993, Moscow, Russia; office phone: +7 499 158-48-74

Torishnyi Roman — Ph.D. Student, Department of Probability Theory and Computer Modelling, Moscow Aviation Institute (National Research University). Research interests: mathematical statistics, probability theory, mathematical modeling, artificial intelligence. The number of publications — 5. Arenas-26@yandex.ru; 4, Volokolamskoe sh., 125993, Moscow, Russia; office phone: +7 968 622-32-39.

Acknowledgements. This work was supported by the state task no. 9.7555.2017/BCh.

References

1. Marchisio M.A. Stochastic Modeling // Introduction to Synthetic Biology. 2018. pp. 39–52.
2. Aleshkin N.A., Aleshkin A.P., Karpova I.R. [The model of optimizing resources in a limited-edition production cycle based on the application of fuzzy logic]. *Voprosy radiojelektroniki — Issues of radio electronics*. 2018. № 9. pp. 84–90. (In Russ.)
3. Mazelis L.S., Soloduhin K.S., Chen A.Ja. [Fuzzy models for optimizing a university's project portfolio inclusive of characteristics of relationships with stakeholders]. *Universitetskoe upravlenie: praktika i analiz. — University Management: Practice and Analysis*. 2017. Issue 21. vol. 5(111). pp. 51–63. (In Russ.)
4. Baburov V.I., Ivancevich N.V., Oljanjuk P.V., Sauta O.I. [Fuzzy criteria and concepts in the problem of GNSS receiver selection for navigation and flight complex]. *Trudy SPIIRAN — SPIIRAS Proceedings*. 2012. vol. 23. pp. 357–368. (In Russ.)
5. Fil'chenkov A.A. [Truth measures and probabilistic graphical models of representation of data with uncertainty]. *Trudy SPIIRAN — SPIIRAS Proceedings*. 2012. vol. 23. pp. 254–295. (In Russ.)
6. Kibzun A.I., Kan Ju.S. Zadachi stohasticheskogo programmirovanija s verojatnostnymi kriterijami [Stochastic programming problems with probabilistic criteria]. M.: Fizmatlit. 2009. 372 p. (In Russ.)
7. Dentcheva D., Penev S., Ruszczyński A. Statistical estimation of composite risk functionals and risk optimization problems. *Annals of the Institute of Statistical Mathematics*. 2017. vol. 69. no. 4. pp. 737–760.
8. Dentcheva D., Ruszczyński A. Risk forms: representation, disintegration, and application to partially observable two-stage systems. *Mathematical Programming*. 2019. pp. 1–21.
9. Kibzun A.I., Naumov A.V., Norkin V.I. [On reducing a quantile optimization problem with discrete distribution to a mixed integer programming problem]. *Avtomatika i telemekhanika — Automation and Remote Control*. 2013. vol. 6. pp. 66—86. (In Russ.)
10. Kibzun A.I., Ignatov A.N. [Reduction of the two-step problem of stochastic optimal control with bilinear model to the problem of mixed integer linear programming]. *Avtomatika i telemekhanika — Automation and Remote Control*. 2016. vol. 12. pp. 89–111. (In Russ.)
11. Ivanov S.V., Kibzun A.I. [Sample Average Approximation in a Two-Stage Stochastic Linear Program with Quantile Criterion]. *Trudy Instituta matematiki i mekhaniki UrO RAN — Proceedings of the Steklov Institute of Mathematics*. 2017. Issue 23. no. 3. pp. 134—143. (In Russ.)
12. Kibzun A.I., Sobol' V.R. [A two-step problem of hedging a European call option under a random duration of transactions]. *Trudy Instituta matematiki i mekhaniki UrO RAN — Proceedings of the Steklov Institute of Mathematics*. 2015. Issue 21. no. 3. pp. 164-–174. (In Russ.)
13. Rajk Je. [Differentiability by parameter of the probability function and stochastic pseudo-gradient method for its optimization]. *Izvestiya Akademii nauk Estonskoj SSR: fizika, matematika — Proceedings of the Estonian Academy of Sciences. Physics, Mathematics*. 1975. vol. 1. pp. 3–8. (In Russ.)
14. Kibzun A.I., Tret'jakov G.L. [Differentiability of the probability function]. *Doklady Akademii nauk — Doklady Mathematics*. 1997. vol. 2. pp. 159–161. (In Russ.)
15. Uryasev S. Derivatives of probability functions and some applications. *Annals of Operations Research*. 1995. vol. 56. no. 1. pp. 287–311.
16. Uryasev S. Derivatives of probability functions and integrals over sets given by inequalities. *Journal of Computational and Applied Mathematics*. 1994. vol. 56. no. 1. pp. 197–223.

17. Henrion R. Gradient estimates for Gaussian distribution functions: application to probabilistically constrained optimization problems. *Numerical Algebra, Control and Optimization*. 2012. vol. 2. no. 4. pp. 655–668.
18. van Ackooij W., Henrion R. (Sub-)Gradient Formulae for Probability Functions of Random Inequality Systems under Gaussian Distribution. *SIAM/ASA Journal on Uncertainty Quantification*. 2017. vol. 5. no. 3. pp. 63–87.
19. Pflug G.C., Weisshaupt H. Probability gradient estimation by set-valued calculus and applications in network design. *SIAM Journal on Optimization*. 2005. vol. 15. no. 3. pp. 898–914.
20. Yu C., Zelterman D. A general approximation to quantiles. *Communications in Statistics-Theory and Methods*. 2017. vol. 46(19). pp. 9834–9841.
21. Okagbue H.I., Adamu M.O., Anake T.A. Ordinary Differential Equations of the Probability Functions of the Weibull Distribution and their Application in Ecology. *International Journal of Engineering and Future Technology*. 2018. vol. 15. pp. 57–78.
22. Garniera J., Omraneb A., Rouchdyc Y.. Asymptotic formulas for the derivatives of probability functions and their Monte Carlo estimations *European Journal of Operational Research*. 2009. vol. 198. no. 3. pp. 848–858.
23. Mafusalov A., Uryasev S. Buffered Probability of Exceedance: Mathematical Properties and Optimization *SIAM Journal on Optimization*. 2018. vol. 28. no. 2. pp. 1077–1103.
24. Zhang T., Uryasev S., Guan Y. Derivatives and subderivatives of buffered probability of exceedance. *Operations Research Letters*. 2019. vol. 47. no. 2. pp. 130–132.
25. Shiryaev A. N. *Verojatnost'* [Probability]. M.:Nauka. 2007. 552 p. (In Russ.)