© 2021 г. **А.Г. ЧХАРТИШВИЛИ, д-р физ.-мат. наук (sandro_ch@mail.ru)**
**(Институт проблем управления им. В.А. Трапезникова РАН, Москва;**
**Институт прикладной математики им. М.В. Келдыша РАН, Москва)**

# ЗАДАЧА НАХОЖДЕНИЯ МЕДИАННОГО ПРЕДПОЧТЕНИЯ ИНДИВИДОВ В СТОХАСТИЧЕСКОЙ МОДЕЛИ[1]

Рассматривается модель предпочтений индивидов, в которой предпочтение каждого индивида характеризуется стохастическим вектором (стохастическая модель предпочтений). При помощи социологического опроса либо анализа действий пользователей в онлайновой социальной сети можно получить вероятностное распределение предпочтений индивидов. Поставлена и решена задача нахождения медианного предпочтения: вектора, минимизирующего ожидаемое расстояние до предпочтений индивидов. Показано, что для нахождения медианного предпочтения достаточно знания частных (маргинальных) распределений. Приведены иллюстративные примеры нахождения медианных предпочтений для трехмерного случая.

*Ключевые слова*: предпочтения индивидов, стохастическая модель предпочтений, ДЛС-модель, расстояние общей вариации, задача оптимизации, медианное предпочтение.

**DOI:** 10.31857/S0005231021050093

## 1. Введение

Выявление предпочтений широких масс населения было актуальной задачей во все времена. Эти предпочтения интересуют политиков и политические партии (стремящиеся снискать популярность у избирателей), корпорации и фирмы (стремящиеся предлагать наиболее востребованные продукты и услуги) и т.д. Примерно сто лет назад стали активно развиваться математические методы, позволяющие на основании опросов делать обоснованные выводы о предпочтениях индивидов (см., например, [1, 2]).

На протяжении последнего десятилетия заметно возросла роль онлайновых социальных сетей (Facebook, Twitter, ВКонтакте и др.) в жизни общества. Они стали ареной активного обмена мнениями, информационного управления и противоборства [3]. Поэтому неудивительно, что теоретики и практики исследуют различные аспекты динамики состояния сетей (например, достижение консенсуса [3, 4]) и во многих случаях делают выводы о предпочтениях пользователей, их взаимном влиянии и т.п. на основании не результатов опроса (т.е. ответов на специально составленные вопросы), а на основании поведения в социальных сетях (см. [5–12] и др.).

[1]Исследование выполнено за счет гранта Российского научного фонда (проект № 20-11-20059) в части решения задачи нахождения медианного предпочтения.

В [11] были предложены две модели машинного обучения для автоматического определения политических взглядов российских пользователей ВКонтакте. В результате апробации на выборке, состоящей из 22 млн цифровых отпечатков аккаунтов совершеннолетних пользователей, были построены две оценки распределения симпатий соответствующих пользователей в преддверии выборов Президента РФ 2018 г. При использовании этих оценок для построения ретроспективного прогноза результатов выборов средние абсолютные ошибки составили 12 и 19,4 % соответственно.

В [13] была предложена математическая модель, позволяющая оценивать идейно-политические предпочтения пользователей и использовать полученную оценку для выбора политическим деятелем той стратегии идеологического позиционирования, которая потенциально позволит получить наибольшую поддержку у заданного конечного множества пользователей. В данной статье исходными данными для нахождения стратегии оптимального позиционирования является плотность вероятностного распределения на множестве предпочтений индивидов. Эта плотность может быть определена как методом традиционных социологических опросов, так и при помощи анализа действий в виртуальном пространстве (онлайновых социальных сетях), либо при помощи сочетания этих двух подходов.

Далее во втором разделе описана стохастическая модель предпочтений, в рамках которой в третьем разделе исследуется задача нахождения медианного предпочтения индивидов. В четвертом разделе приводятся примеры нахождения медианного предпочтения для заданных вероятностных распределений. В заключении приведены основные выводы и намечены направления дальнейших исследований.

## 2. Стохастическая модель предпочтений

Разнообразие внутреннего мира и поведения людей является предметом исследования различных отраслей науки. Это разнообразие учитывалось при создании теорий и моделей путем введения в рассмотрение типизации индивидов в соответствии с их свойствами или отношением к тем или иным объектам. Например, психологами был разработан ряд подходов, позволяющих определять как темперамент, так и другие психологические характеристики индивида (см. [14–16] и др.). Отметим, что «чистые» типы индивидов встречаются относительно редко, поэтому для описания индивидов используют различные методы определения промежуточных, «смешанных» типов.

Богатая традиция типизации идейно-политических взглядов существует в социолого-политологических исследованиях. Приведем несколько примеров.

Одна из первых классификаций была предложена М. Рокичем в книге «Природа человеческих ценностей» [17]. Рокич выдвинул идею о том, что содержание главных четырех идеологических течений XX века — социализма, коммунизма, фашизма и капитализма — можно представить в виде двумерной шкалы, образованной системами координат двух ценностей: свободы и равенства. Эта модель содержит четыре ячейки, образованные высоким и низким положением каждой из этих ценностей.

В [18] была предложена классификация, состоящая из шести политических направлений: социал-государственники, социал-демократы, радикальные националисты, державники, радикальные рыночники, западники. В [19] были предложены четыре типа политических ценностей: рыночные, демократические, социалистические, державнические.

В [13] идейно-политические взгляды современных россиян предлагается рассматривать как смесь трех ярко выраженных базисных установок: Державник-Либерал-Социалист. Компонента «Державник» описывает, насколько индивид привержен идее о едином, сильном и независимом государстве. Компонента «Либерал» выражает важность для индивида уважения со стороны государства его (индивида) личных прав и свобод (в первую очередь политических и экономических). Компонента «Социалист» отвечает за стремление индивида к социально-экономической справедливости. В рамках предлагаемой модели (*ДЛС-модели*, по первым буквам базисных установок) существует три чистых типа, и взгляды каждого индивида представимы в виде ДЛС-вектора (вектора-строки) из трех неотрицательных компонент, в сумме равных единице.

В общем случае можно рассматривать $k$ базисных установок, позволяющих типизировать предпочтения отдельного индивида в виде стохастического вектора (вектора неотрицательных вещественных чисел, в сумме равных единице) $x = (x_1, \ldots, x_k)$. Величину $x_j$ можно интерпретировать как вероятность совершения индивидом действия, характерного для данной базисной установки, в ситуации выбора из $k$ альтернатив.

Для решения ряда задач, связанных с предпочтениями индивидов, требуется находить расстояние между предпочтениями индивидов. В [12, 13] было применено следующее нормированное (принимающее значения от 0 до 1) расстояние между двумя стохастическими векторами $x^1 = \left(x_1^1, \ldots, x_k^1\right)$ и $x^2 = \left(x_1^2, \ldots, x_k^2\right)$:

$$d\left(x^1, x^2\right) = 1 - \sum_{j=1}^{k} \min\left(x_j^1, x_j^2\right) = \frac{1}{2} \sum_{j=1}^{k} \left|x_j^1 - x_j^2\right|. \tag{1}$$

Отметим, что справедливость второго равенства в (1) для стохастических векторов легко показать:

$$\frac{1}{2}\sum_{j=1}^{k}\left|x_j^1 - x_j^2\right| = \frac{1}{2}\sum_{j=1}^{k}\left(\max\left(x_j^1, x_j^2\right) - \min\left(x_j^1, x_j^2\right)\right) =$$

$$= \frac{1}{2}\left(2 - \sum_{j=1}^{k}\min\left(x_j^1, x_j^2\right) - \sum_{j=1}^{k}\min\left(x_j^1, x_j^2\right)\right) =$$

$$= 1 - \sum_{j=1}^{k}\min\left(x_j^1, x_j^2\right).$$

Метрика (1) является частным случаем хорошо известного в теории вероятностей расстояния по вариации (total variation distance) — см., например, [20, с. 463]; также (1) совпадает с манхеттенской метрикой (Manhattan metric) с точностью до множителя $1/2$. Рассмотрим на примере вопрос о том, в чем ее особенность в предлагаемой модели по сравнению с более традиционной для расчетов в конечномерных пространствах евклидовой метрикой

$$d_E(x^1, x^2) = \sqrt{\sum_{j=1}^{k} \left(x_j^1 - x_j^2\right)^2}.$$

Предположим, что $k = 3$ и имеется три вектора, характеризующих предпочтения первого, второго и третьего индивидов соответственно: $x^1 = (1,\ 0,\ 0)$, $x^2 = (0,\ 1,\ 0)$, $x^3 = (0,\ \frac{1}{2},\ \frac{1}{2})$. Первый индивид всегда выбирает действия, характерные для первой базисной установки, в то время как второй и третий индивиды такие действия не выбирают никогда. Можно положить, что второй и третий индивиды одинаково далеки от первого, что соответствует соотношению $d\left(x^1, x^2\right) = d\left(x^1, x^3\right)$, в справедливости которого легко убедиться. В этой логике метрика (1) представляется более адекватной, чем евклидова метрика, в которой расстояния между этими же векторами предпочтений различны: $d_E\left(x^1, x^2\right) \neq d_E\left(x^1, x^3\right)$.

Таким образом, рассматриваемая в данной работе стохастическая модель предпочтений основывается на представлении предпочтений индивидов в виде стохастических векторов с метрикой (1).

## 3. Постановка и решение задачи нахождения медианного предпочтения

Информацию о предпочтениях индивидов можно использовать различным образом. В данной работе остановимся на одной из возможных задач: найти вектор предпочтений, максимально близкий (в смысле минимизации математического ожидания расстояния (1)) к предпочтениям совокупности индивидов, для которых известна плотность вероятностного распределения.

Медианное предпочтение активно применяется в математическом моделировании политической конкуренции начиная с работ Даунса (см. [21]). Если стохастический вектор отражает идейно-политические предпочтения индивидов [13], то медианное предпочтение можно интерпретировать как оптимальную идейно-политическую позицию политика или политической партии.

Опишем задачу более формально. Пусть $k$-мерная случайная величина, принимающая значения в пределах стандартного $(k-1)$-мерного симплекса

$$S = \left\{ (x_1, \ldots, x_k) \;\middle|\; x_j \geq 0,\ j \in \{1, \ldots, k\},\ \sum_{j=1}^{k} x_j = 1 \right\},$$

задана при помощи плотности распределения $f$; соответствующую интегральную функцию распределения обозначим через $F$. Сформулируем задачу поиска оптимального вектора предпочтений (т.е. медианного предпочтения) $a = (a_1, \ldots, a_k)$ следующим образом:

$$L(a_1, \ldots, a_k) = \int\limits_S d(\sigma, a)\ dF(\sigma) \xrightarrow[(a_1,\ldots,a_k)\in S]{} \min. \tag{2}$$

Задача (2) является, по сути, задачей нахождения многомерной медианы (точнее говоря, одной из многомерных медиан — см. [22–24]). Отметим, что в случае евклидовой метрики задача нахождения медианы множества точек является довольно сложной и трудоемкой для численного решения [25].

С учетом (1) задача (2) записывается следующим образом (для сокращения записи опущен множитель $1/2$ с сохранением того же обозначения $L$):

$$L(a_1, \ldots, a_k) = \int\limits_S \left( \sum_{j=1}^{k} |x_j - a_j| \right) dF(\sigma) \xrightarrow[(a_1,\ldots,a_k)\in S]{} \min . \tag{3}$$

*Замечание.* Нетрудно видеть, что минимизируемая функция в (3) является суммой $k$ функций, каждая из которых зависит только от соответствующей $i$-й компоненты $a_i$. Поэтому при отсутствии требования принадлежности вектора $a$ симплексу $S$ задача (3) распадается на $k$ независимых оптимизационных задач, решением каждой из которых является медиана одномерного частного (маргинального) распределения переменной $x_i$. Однако составленный из этих медиан вектор не принадлежит, вообще говоря, симплексу $S$ (см. далее пример 1 в разделе 4), т.е. не является решением задачи (3). ■

Обозначим через $F_j(x)$, $x \in [0, 1]$, функцию одномерного частного распределения переменной $x_j$, являющегося компонентой исходного распределения. Справедливо следующее утверждение, характеризующее необходимое условие решения задачи (3).

*Утверждение 1. Если вектор* $(a_1, \ldots, a_k)$ *является решением задачи* (3), *то для него выполняется следующее условие:*

$$F_1(a_1) = \cdots = F_k(a_k) . \tag{4}$$

*Доказательство.* Предположим, что некоторый вектор $(a_1, \ldots, a_k) \in$ $\in S$ является решением задачи (3), но условие (4) для него не выполняется. Это означает, что существуют компоненты этого вектора с номерами $i$ и $j$, для которых справедливо неравенство

$$F_i(a_i) < F_j(a_j) . \tag{5}$$

Обозначим через $L_0 = L(a_1, \ldots, a_k)$ значение целевой функции задачи (3) на этом векторе.

Увеличим $a_i$ на малую величину $\varepsilon > 0$, а $a_j$ уменьшим на ту же величину (в дальнейшем значение $\varepsilon$ будет уточнено). Это можно сделать, не выходя за границы отрезка [0,1], поскольку в силу (5) $F_i(a_i) < 1$, $a_i < 1$ и $F_j(a_j) > 0$, $a_j > 0$.

Тем самым получим новый вектор, принадлежащий симплексу $S$; значение целевой функции на нем обозначим через $L_\varepsilon$. Запишем разность между значениями $L_\varepsilon$ и $L_0$ (с учетом того, что в сумме под знаком интеграла в (3) изменились только слагаемые, содержащие компоненты $a_i$ и $a_j$):

$$
\begin{gathered}
L_\varepsilon - L_0 = \qquad (6)\\
= \int\limits_S (|x_i - (a_i+\varepsilon)| + |x_j - (a_j - \varepsilon)|)\, dF(\sigma) - \int\limits_S (|x_i - a_i| + |x_j - a_j|)\, dF(\sigma) = \\
= \int\limits_S (|x_i - a_i - \varepsilon| - |x_i - a_i|)\ dF(\sigma) + \int\limits_S (|x_j - a_j + \varepsilon)| - |x_j - a_j|)\ dF(\sigma).
\end{gathered}
$$

Представим $S$ как объединение трех непересекающихся областей $S_i^{\varepsilon-}$, $S_i^{\varepsilon}$ и $S_i^{\varepsilon+}$, на которые ее делят гиперплоскости $x_i = a_i$ и $x_i = a_i + \varepsilon$:

$$
\begin{aligned}
S_i^{\varepsilon-} &= \{(x_1,\ldots,x_k) \in S \mid x_i \le a_i\},\\
S_i^{\varepsilon} &= \{(x_1,\ldots,x_k) \in S \mid a_i < x_i \le a_i + \varepsilon\},\\
S_i^{\varepsilon+} &= \{(x_1,\ldots,x_k) \in S \mid x_i > a_i + \varepsilon\}.
\end{aligned}
$$

Также представим $S$ как объединение трех областей $S_j^{\varepsilon-}$, $S_j^{\varepsilon}$ и $S_j^{\varepsilon+}$, на которые ее делят гиперплоскости $x_j = a_j - \varepsilon$ и $x_j = a_j$:

$$
\begin{aligned}
S_j^{\varepsilon-} &= \{(x_1,\ldots,x_k) \in S \mid x_j \le a_j - \varepsilon\},\\
S_j^{\varepsilon} &= \{(x_1,\ldots,x_k) \in S \,\Big|\, a_j - \varepsilon < x_j \le a_j\},\\
S_j^{\varepsilon+} &= \{(x_1,\ldots,x_k) \in S \mid x_j > a_j\}.
\end{aligned}
$$

Эти обозначения позволяют записать $L_\varepsilon - L_0$ следующим образом (продолжая (6)):

$$
\begin{gathered}
(7)\qquad L_\varepsilon - L_0 = \int\limits_{S_i^{\varepsilon-}} \varepsilon\, dF(\sigma) + \int\limits_{S_i^{\varepsilon}} (-2x_i + 2a_i + \varepsilon)\ dF(\sigma) - \int\limits_{S_i^{\varepsilon+}} \varepsilon\, dF(\sigma) - \\
- \int\limits_{S_j^{\varepsilon-}} \varepsilon\, dF(\sigma) + \int\limits_{S_j^{\varepsilon}} (2x_j - 2a_j + \varepsilon)\ dF(\sigma) + \int\limits_{S_j^{\varepsilon+}} \varepsilon\, dF(\sigma) \le \\
\le \varepsilon \left( \int\limits_{S_i^{\varepsilon-}} dF(\sigma) + \int\limits_{S_i^{\varepsilon}} dF(\sigma) - \int\limits_{S_i^{\varepsilon+}} dF(\sigma) - \int\limits_{S_j^{\varepsilon-}} dF(\sigma) + \int\limits_{S_j^{\varepsilon}} dF(\sigma) + \int\limits_{S_j^{\varepsilon+}} dF(\sigma) \right).
\end{gathered}
$$

Рассмотрим поведение правой части (7) при стремлении $\varepsilon$ к нулю справа. Нетрудно видеть, что при $\varepsilon \to +0$ выполняются следующие соотношения:

$$
\begin{gathered}
\int\limits_{S_i^{\varepsilon-}} dF(\sigma) = F_i(a_i), \quad \int\limits_{S_i^{\varepsilon}} dF(\sigma) \to 0, \quad \int\limits_{S_i^{\varepsilon+}} dF(\sigma) \to 1 - F_i(a_i),\\
\int\limits_{S_j^{\varepsilon-}} dF(\sigma) \to F_j(a_j), \quad \int\limits_{S_j^{\varepsilon}} dF(\sigma) \to 0, \quad \int\limits_{S_j^{\varepsilon+}} dF(\sigma) = 1 - F_j(a_j).
\end{gathered}
$$

Поэтому сумма интегралов в скобках правой части (7) стремится при $\varepsilon \to +0$ к строго отрицательному (см. (5)) пределу $2\left(F_i\left(a_i\right) - F_j\left(a_j\right)\right)$.

Таким образом, существует (близкое к нулю) число $\varepsilon_0 > 0$, для которого выполняется неравенство $L_{\varepsilon_0} - L_0 < 0$. Это противоречит исходному предположению об оптимальности вектора $(a_1, \ldots, a_k)$. Утверждение 1 доказано. ■

Следующее утверждение показывает, что если для двух стохастических векторов выполняется условие (4), то значения целевой функции $L$ на этих векторах совпадают.

*Утверждение 2. Пусть для двух стохастических векторов* $(a_1, \ldots, a_k)$ *и* $(b_1, \ldots, b_k)$ *выполнены следующие условия:*

$$F_1\left(a_1\right) = \cdots = F_k\left(a_k\right) = t_1, \quad F_1\left(b_1\right) = \cdots = F_k\left(b_k\right) = t_2. \tag{8}$$

*Тогда*

$$t_1 = t_2 \quad \text{и} \quad L\left(a_1, \ldots, a_k\right) = L(b_1, \ldots, b_k).$$

*Доказательство.* Докажем сначала, что $t_1 = t_2$. Предположим, что $t_1 < t_2$; тогда для каждого $j$ выполняется неравенство $F_j\left(a_j\right) < F_j\left(b_j\right)$, откуда следует $a_j < b_j$. Суммируя по всем $j$, получаем неравенство

$$\sum_{j=1}^{k} a_j < \sum_{j=1}^{k} b_j,$$

которое противоречит стохастичности обоих векторов.

Итак, $t_1 = t_2 = t$. Заметим, что равенство двух значений одномерной функции распределения для данного $j$ в (8) свидетельствует о равенстве нулю интеграла на промежутке между $a_j$ и $b_j$, т.е. для любого $j$ от 1 до $k$ выполняется равенство

$$\int_{a_j}^{b_j} dF_i(x) = F_i(b_i) - F_i(a_i) = 0. \tag{9}$$

Введем следующие обозначения:

$$S_j^{a-} = \left\{(x_1, \ldots, x_k) \in S \mid x_j \le a_j\right\}, \quad S_j^{a+} = \left\{(x_1, \ldots, x_k) \in S \mid x_j > a_j\right\},$$
$$S_j^{b-} = \left\{(x_1, \ldots, x_k) \in S \mid x_j \le b_j\right\}, \quad S_j^{b+} = \left\{(x_1, \ldots, x_k) \in S \mid x_j > b_j\right\}.$$

Тогда справедливы равенства

$$\int_{S_j^{a-}} dF(\sigma) = \int_{S_j^{b-}} dF(\sigma) = t, \quad \int_{S_j^{a+}} dF(\sigma) = \int_{S_j^{b+}} dF(\sigma) = 1 - t.$$

Воспользовавшись этими равенствами, найдем с учетом (9) разность между значениями $L(a_1, \ldots, a_k)$ и $L(b_1, \ldots, b_k)$:

$$L(a_1, \ldots, a_k) - L(b_1, \ldots, b_k) =$$

$$= \sum_{j=1}^{k} \left( \int\limits_{S_j^{a-}} (a_j - x_j)\, dF(\sigma) + \int\limits_{S_j^{a+}} (x_j - a_j)\, dF(\sigma) - \right.$$

$$\left. - \int\limits_{S_j^{b-}} (b_j - x_j)\, dF(\sigma) - \int\limits_{S_j^{b+}} (x_j - b_j)\, dF(\sigma) \right) = (2t - 1) \sum_{j=1}^{k} (a_j - b_j) = 0.$$

Утверждение 2 доказано. ■

Утверждения 1 и 2 позволяют полностью описать решение оптимизационной задачи (3). Сформулируем соответствующее утверждение.

*Утверждение 3. Для того чтобы стохастический вектор* $(a_1, \ldots, a_k)$ *был решением оптимизационной задачи* (3), *необходимо и достаточно выполнение условия* (4).

*Доказательство.* Обозначим через $S^{=}$ множество стохастических векторов, для которых выполняется условие (4). Поскольку множество $S$ стохастических векторов является компактным, а функция $L$ непрерывной, задача (3) имеет решение. В соответствии с утверждением 1 все решения задачи (3) содержатся в $S^{=}$. В соответствии с утверждением 2 на всех векторах из $S^{=}$ целевая функция принимает одно и то же значение, поэтому множество решений задачи (3) совпадает с $S^{=}$. Утверждение 3 доказано. ■

В заключение данного раздела отметим, что определение стохастических представлений индивидов (т.е. позиционирования предпочтений индивида в симплексе $S$) является тем более сложным, чем большее значение принимает размерность вектора предпочтений $k$. Однако утверждение 3 показывает, что для определения медианного предпочтения не требуется знания общего многомерного распределения, достаточно знания одномерных частных распределений. Поэтому для оценки медианного предпочтения достаточно провести $k$ более простых («одномерных») социологических опросов либо исследований активности в социальных сетях.

## 4. Примеры

В данном разделе рассмотрим два иллюстративных примера нахождения медианного предпочтения $(a, b, c) \in S$ для случая трехкомпонентного вектора предпочтений,

$$S = \{(x, y, z) \mid x, y, z \geq 0,\ x + y + z = 1\},$$

и вероятностного распределения, заданного в виде плотности $f(x, y, z)$. В этом случае симплекс $S$ является ограниченной частью плоскости (треугольником). Поэтому симплекс $S$ или его часть можно представить как график функции от двух переменных, а нахождение поверхностного интеграла сводится к нахождению повторного интеграла.

Рассмотрим, например, область

$$S_{\alpha\beta} = \{(x, y, z) \in S \mid \alpha \leq x \leq \beta\}.$$

Интеграл по плотности $f$ можно вычислить, рассматривая $S_{\alpha\beta}$ как часть графика функции $z = 1 - x - y$, следующим образом:

$$\int\limits_{S_{\alpha\beta}} f(\sigma) d\sigma = \int\limits_{\alpha}^{\beta} dx \int\limits_{0}^{1-x} f(x, y, 1 - x - y) \sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2} dy.$$

Из последнего соотношения видно, что одномерное частное распределение переменной $x$ задается плотностью $f_1(x)$ следующего вида:

$$f_1(x) = \sqrt{3} \int\limits_{0}^{1-x} f(x, y, 1 - x - y) dy.$$

Аналогично, частные распределения переменных $y$ и $z$ задаются соответственно формулами

$$f_2(y) = \sqrt{3} \int\limits_{0}^{1-y} f(1 - y - z, y, z) dz,$$

$$f_3(z) = \sqrt{3} \int\limits_{0}^{1-z} f(x, 1 - x - z, z) dx.$$

*Пример 1.* Пусть плотность распределения является константой: $f(x, y, z) = 2/\sqrt{3}$. Тогда плотность частного распределения $f_1(x)$ имеет вид

$$f_1(x) = \sqrt{3} \int\limits_{0}^{1-x} \frac{2}{\sqrt{3}} dy = 2(1 - x),$$

а функция частного распределения $F_1(x)$ — вид

$$F_1(x) = 2x - x^2.$$

Аналогичный вид имеют остальные два частных распределения. Поэтому условие медианного предпочтения (4) вместе с условием стохастичности приводит к системе уравнений

$$\begin{cases} 2a - a^2 = 2b - b^2 = 2c - c^2, \\ a + b + c = 1, \end{cases}$$

решение которой является медианным предпочтением ($a = b = c = 1/3$).

Заметим, что медианой частного распределения $F_1(x)$ является решение уравнения $F_1(x) = \frac{1}{2}$, т.е. $x = 1 - \frac{\sqrt{2}}{2}$. Это же значение принимают медианы остальных двух частных распределений. Легко видеть, что вектор $\left(1 - \frac{\sqrt{2}}{2}; 1 - \frac{\sqrt{2}}{2}; 1 - \frac{\sqrt{2}}{2}\right) \approx (0{,}293;\ 0{,}293;\ 0{,}293)$ медиан частных распределений не принадлежит поверхности $S$ и, следовательно, не может быть решением задачи (3).

*Пример 2.* Пусть плотность распределения имеет вид $f(x,y,z) = 2\sqrt{3}x$. Тогда плотность частного распределения $f_1(x)$ имеет вид

$$f_1(x) = \sqrt{3}\int_0^{1-x} 2\sqrt{3}x\,dy = 6(x - x^2),$$

а функция частного распределения $F_1(x)$ — вид

$$F_1(x) = 3x^2 - 2x^3.$$

Далее, плотность частного распределения $f_2(y)$ имеет вид

$$f_2(y) = \sqrt{3}\int_0^{1-y} 2\sqrt{3}(1-y-z)\,dz = 3(1-y)^2,$$

а функция частного распределения $F_2(y)$ — вид

$$F_2(y) = 3y - 3y^2 + y^3.$$

Аналогичный вид имеет функция частного распределения $F_3(z)$.

Условие медианного предпочтения (4) вместе с условием стохастичности приводит к системе уравнений

$$\begin{cases} 3a^2 - 2a^3 = 3b - 3b^2 + b^3 = 3c - 3c^2 + c^3, \\ a + b + c = 1, \end{cases}$$

решение которой является медианным предпочтением ($a \approx 0{,}533$; $b = c \approx \approx 0{,}233$).

## 5. Заключение

В данной статье рассмотрена стохастическая модель предпочтений индивидов, в которой предпочтение каждого индивида характеризуется стохастическим вектором. Решена задача нахождения медианного предпочтения: вектора, минимизирующего ожидаемое расстояние до предпочтений индивидов. Показано, что для нахождения медианного предпочтения достаточно знания частных (маргинальных) распределений. Приведены примеры расчета медианного предпочтения при заданном вероятностном распределении.

Отметим, что стохастическая модель предпочтений является не только теоретической конструкцией, но и основой методики практического исследования идейно-политических [13, 26], электоральных [11] и других предпочтений.

Перспективным направлением дальнейших исследований представляется как исследование других свойств стохастической модели предпочтений, так и ее практическое применение для анализа общественного сознания.

## СПИСОК ЛИТЕРАТУРЫ


1. *Девятко И.Ф.* Диагностическая процедура в социологии. Очерк истории и теории. М.: Наука, 1993.
2. *Толстова Ю.Н.* Тестовая традиция в социологии. М.: КДУ, 2007.
3. *Chkhartishvili A.G., Gubanov D.A., Novikov D.A.* Social Networks: Models of information influence, control and confrontation. Springer, 2018.
4. *Проскурников А.В.* Консенсус в нелинейных стационарных сетях с идентичными агентами // АиТ. 2015. № 9. С. 44–63.

   *Proskurnikov A.V.* Consensus in Nonlinear Stationary Networks with Identical Agents // Autom. Remote Control. 2015. V. 76. No. 9. P. 1551–1565.
5. *Козякин В.С., Кузнецов Н.А., Чеботарев П.Ю.* Консенсус в асинхронных мультиагентных системах. I. Асинхронные модели консенсуса // АиТ. 2019. № 4. С. 3–40.

   *Kozyakin V.S., Kuznetsov N.A., Chebotarev P.Y.* Consensus in Asynchronous Multiagent Systems. I. Asynchronous Consensus Models // Autom. Remote Control. 2019. V. 80. No. 4. P. 593–623.
6. *Gayo-Avello D.* A meta-analysis of state-of-the-art electoral prediction from Twitter data // Social Science Computer Review. 2013. V. 3. No. 6. P. 649–679.
7. *Petrov A.P., Proncheva O.G.* Modeling Position Selection by Individuals during Informational Warfare with a Two-Component Agenda // Mathematical Models and Computer Simulations. 2020. V. 12. No. 2. P. 154–163.
8. *Barberá P.* Birds of the same feather tweet together: Bayesian ideal point estimation using Twitter data // Political analysis. 2015. V. 23. No. 1. P. 76–91.
9. *Makazhanov A., Rafiei D., Waqar M.* Predicting political preference of Twitter users // Social Network Analysis and Mining. 2014. V. 4. No. 1. P. 193.
10. *Губанов Д.А., Чхартишвили А.Г.* Влиятельность пользователей и метапользователей социальной сети // Проблемы управления. 2016. № 6. С. 12–17.

    *Gubanov D.A., Chkhartishvili A.G.* Influence Levels of Users and Meta-users of a Social Network // Autom. Remote Control. 2018. V. 79. No. 3. P. 545–553.
11. *Kozitsin I.V., Chkhartishvili A.G., Marchenko A.M. et al.* Modeling Political Preferences of Russian Users Exemplified by the Social Network Vkontakte // Mathematical Models and Computer Simulations. 2020. V. 12. No. 2. P. 185–194.
12. *Chkhartishvili A.G., Kozitsin I.V., Goiko V.L., Saifulin E.R.* On an Approach to Measure the Level of Polarization of Individuals' Opinions // Twelfth International Conference "Management of large-scale system development" (MLSD). Moscow. 2019. P. 1–5.
13. *Бызов Л.Г., Губанов Д.А., Козицин И.В., Чхартишвили А.Г.* Идеальный политик для социальной сети: подход к анализу идеологических предпочтений пользователей // Проблемы управления. 2020. № 4. С. 15–26.

14. *Goldberg L.R.* The development of markers for the Big-Five factor structure // Psychological Assessment. 1992. No. 4. P. 26–42.
15. *Eysenck S.B.G., Eysenck H.J., Barrett P.* A revised version of the psychoticism scale // Personality and Individual Difference. 1985. Vol. 6. No. 1. P. 21–29.
16. *Keirsey D., Bate M.* Please Understand Me: Character and Temperament Types. Del Mar, CA. Prometheus Nemesis books. 1984.
17. *Rokeach M.* The nature of human values. N.Y.: Free Press, 1973.
18. *Бызов Л.Г.* Динамика идейно-политических предпочтений за 25 лет. Три этапа трансформации общественного сознания // Россия XXI. 2019. № 2. С. 6–21.
19. *Блинов В.В.* Типы идеологических сторонников в современной России // Мониторинг общественного мнения: экономические и социальные перемены. 2010. № 6 (100).
20. *Ширяев А.Н.* Вероятность. В 2-х кн. 3-е изд., перераб. и доп. М.: МЦНМО, 2004.
21. *Downs A.* An Economic Theory of Democracy. N.Y.: Harper, 1957.
22. *Small C.G.* A survey of multidimensional medians // International Statistical Review. № 58. 1990. P. 263–277.
23. *Vardi Y., Zhang C.-H.* The multivariate $L1$-median and associated data depth // Proceedings of the National Academy of Sciences Feb 2000, No. 97 (4). P. 1423–1426. https://doi.org/10.1073/pnas.97.4.1423
24. *Chen F., Nason G.* A new method for computing the projection median, its influence curve and techniques for the production of projected quantile plots // PLoS ONE. 2020. No. 15(5): e0229845. https://doi.org/10.1371/journal.pone.0229845
25. *Croux C., Filzmoser P., Fritz H.* A comparison of algorithms for the multivariate L1-median // Computational Statistics. 2010. V. 27. No. 3. P. 1–18.
26. *Губанов Д.А., Козицин И.В., Чхартишвили А.Г.* О выявлении идейно-политических предпочтений пользователей городских сообществ в онлайновой социальной сети // Труды 10-й Международной социологической Грушинской конференции «Жить в России. Жить в мире. Социология повседневности», 20 мая–14 ноября 2020 г. М.: ВЦИОМ, 2020. С. 291–297.

*Статья представлена к публикации членом редколлегии Ф.Т. Алескеровым.*

Поступила в редакцию 22.10.2020

После доработки 09.12.2020

Принята к публикации 15.01.2021