ПРОБЛЕМЫ ПЕРЕДАЧИ ИНФОРМАЦИИ

Том 55 2019 Вып. 4

УДК 621.391.1 : 519.72

© 2019 г. **Б. Накибоглу**

АВГУСТИНОВСКАЯ ПРОПУСКНАЯ СПОСОБНОСТЬ И АВГУСТИНОВСКИЙ ЦЕНТР[1]

Для любого канала устанавливается существование и единственность августиновского среднего любого положительного порядка для любой функции вероятности на множестве входов. Показано, что августиновское среднее является единственной неподвижной точкой некоторого оператора, зависящего от порядка и распределения на входе. Показано, что информация Августина непрерывно дифференцируема по порядку. Для любого канала и любого выпуклого множества ограничений с конечной августиновской пропускной способностью установлено существование и единственность августиновского центра и получена соответствующая граница ван Эрвена – Харремоеса. Введены понятия информации, пропускной способности, центра и радиуса Августина – Лежандра (АЛ) и доказано, что последние три равны соответствующим величинам Реньи – Галлагера. Установлено равенство АЛ-пропускной способности и АЛ-радиуса для произвольных каналов, а также существование и единственность АЛ-центра для каналов с конечной АЛ-пропускной способностью. Для всех внутренних точек множества допустимых ограничений по стоимости получены выражения для августиновских пропускной способности и центра при наличии ограничений по стоимости через АЛ-пропускную способность и центр. В качестве примеров рассмотрены некоторые инвариантные относительно сдвига семейства вероятностей и некоторые гауссовские каналы.

Ключевые слова: расхождение Реньи, информация Реньи, информация Августина, августиновское среднее, августиновский центр, августиновская пропускная способность, пропускная способность и центр при наличии ограничений по стоимости, информационные величины Августина – Лежандра, информационные величины Реньи – Галлагера.

DOI: 10.1134/S0555292319040016

§ 1. Введение

Взаимная информация, иногда называемая информацией Шеннона, имеет первостепенную важность при анализе самых разных вопросов передачи информации. Хотя эта величина определяется без ссылки на задачу оптимизации, она удовлетворяет следующим двум тождествам, выраженным через расхождение Кульбака – Лейблера:

$$I(p;W) = \inf_{q\in\mathcal{P}(\mathcal{Y})} D(p \circledast W \,\|\, p \otimes q) = \tag{1}$$

$$= \inf_{q\in\mathcal{P}(\mathcal{Y})} \sum_x p(x) D(W(x) \,\|\, q), \tag{2}$$

[1] Работа была частично представлена на Международном симпозиуме IEEE по теории информации 2017 г. [1].

где $\mathcal{P}(\mathcal{Y})$ – множество всех вероятностных мер на выходном пространстве $(\mathscr{Y},\mathcal{Y})$, p – функция вероятности, положительная лишь на конечном подмножестве множества входов $\mathscr{X}$, а W – функция вида $W\colon \mathscr{X} \to \mathcal{P}(\mathcal{Y})$. За определение взаимной информации можно взять любое из этих двух выражений. Можно определить информацию Реньи порядка α, заменяя в этих выражениях расхождение Кульбака – Лейблера на расхождение Реньи порядка α. Поскольку расхождение Реньи порядка 1 совпадает с расхождением Кульбака – Лейблера, информация Реньи порядка 1 равна взаимной информации при обоих определениях, однако, как отмечено Чисаром в [2], для других порядков эти два определения не равносильны ни определению взаимной информации, ни друг другу. Обобщение, связанное с выражением (1), называется информацией Реньи порядка α и обозначается через $I_\alpha^{\mathrm{g}}(p;W)$, а обобщение, связанное с выражением (2), называется информацией Августина порядка α и обозначается через $I_\alpha(p;W)$. Аналогично шенноновской пропускной способности при входном ограничении августиновская пропускная способность порядка α для множества ограничений $\mathcal{A}$ определяется как $\sup\limits_{p\in\mathcal{A}} I_\alpha(p;W)$.

Как было недавно показано в [3, 4] соответственно, информация Августина порядков, меньших единицы, возникает в выражении для экспоненты сферической упаковки для кодов с постоянной композицией в классически-квантовых каналах без памяти, а информация Августина порядков, больших единицы, – в выражениях для экспоненты ошибки в сильном обращении теоремы кодирования. Для кодов с постоянной композицией в дискретных стационарных каналах-произведениях этот факт был неявно отмечен еще в [5, с. 172; 2].

Для кодов с ограничениями по стоимости в (возможно, нестационарных) каналах-произведениях с аддитивными функциями стоимости аналогичную роль в выражениях для экспоненты сферической упаковки и экспоненты ошибки в сильном обращении теоремы кодирования играет августиновская пропускная способность при наличии ограничений по стоимости. Наблюдение относительно экспоненты сферической упаковки для довольно общих моделей канала было сделано Августином в [6, замечание 36.7(i) и § 36]. Таким образом, августиновские информационные величины действительно имеют важное операциональное значение, по крайней мере, для задачи кодирования каналов. Однако основной целью настоящей статьи является изучение информации Августина и августиновской пропускной способности с точки зрения теории меры. В дальнейшем мы нигде не ссылаемся на задачу кодирования каналов или на операциональный смысл августиновских информационных величин, поскольку считаем, что информацию Августина и августиновскую пропускную способность нужно в первую очередь рассматривать как понятия теории меры. Уже после этого операциональное значение информации Августина и августиновской пропускной способности можно устанавливать с помощью теоретико-информационной техники и результатов анализа с позиций теории меры, как это сделано в [7].

Во всех предшествующих работах по информации Августина или августиновской пропускной способности, за исключением работы Августина [6], множество выходов $\mathscr{Y}$ канала W предполагалось конечным [2, 3, 8–11], что является серьезным недостатком, поскольку предположение конечности множества выходов не выполнено для некоторых моделей, интересных с аналитической точки зрения, а также исключительно важных с точки зрения инженерных приложений, таких как гауссовские или пуассоновские модели каналов. Мы проводим анализ для более общей модели и предполагаем[2], что выходное пространство $(\mathscr{Y},\mathcal{Y})$ является измеримым пространством, состоящим из множества выходов $\mathscr{Y}$ и σ-алгебры $\mathcal{Y}$ его подмножеств. В этой общей постановке наш анализ информации Августина и августиновской пропуск-

[2] В п. 5.4 вводятся дополнительные предположения, но и они, по существу, выполнены для практически всех имеющих интерес каналов.

ной способности строится вокруг двух фундаментальных понятий: августиновского среднего и августиновского центра.

Напомним, что взаимная информация $I(p;W)$ определяется следующим образом: $I(p;W) \triangleq \sum\limits_x p(x)D(W(x)\,\|\,q_{1,p})$, где $q_{1,p} = \sum\limits_x p(x)W(x)$. Следовательно, инфимум в (2) достигается на $q_{1,p}$. Кроме того, непосредственной подстановкой можно убедиться, что

$$\sum_x p(x)D(W(x)\,\|\,q) = I(p;W) + D(q_{1,p}\,\|\,q) \quad \forall q \in \mathcal{P}(\mathcal{Y}).$$

Таким образом, $q_{1,p}$ – единственная вероятностная мера, на которой достигается инфимум в (2), поскольку расхождение Кульбака – Лейблера для различных вероятностных мер положительно. То же самое верно и для других порядков: для любого $\alpha \in \mathbb{R}_+$ существует единственная вероятностная мера $q_{\alpha,p}$, такая что $I_\alpha(p;W) = \sum\limits_x p(x)D_\alpha(W(x)\,\|\,q_{\alpha,p})$. Мы называем эту вероятностную меру $q_{\alpha,p}$ *августиновским средним порядка* α. В [6, лемма 34.2] Августин доказал существование и единственность мер $q_{\alpha,p}$ для порядков $\alpha \in (0,1]$ и описал важные характеристики этих мер, являющиеся краеугольным камнем в анализе информации Августина и августиновской пропускной способности. Аналогичные соотношения для порядков, больших 1, выводятся в § 3 (см. лемму 13(d)).

В [12] доказано равенство шенноновской пропускной способности (при отсутствии ограничений) и шенноновского радиуса[3] для любого канала вида $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$, а также существование и единственность шенноновского центра для каналов с конечной шенноновской пропускной способностью. Используя идеи, уже содержащиеся в этом доказательстве, можно установить аналогичный результат для шенноновской пропускной способности при наличии ограничений при условии, что множество ограничений выпукло (см. [13, теорема 2]): для любого канала W вида $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ и выпуклого множества ограничений $\mathcal{A}$

$$\sup_{p\in\mathcal{A}} I(p;W) = \inf_{q\in\mathcal{P}(\mathcal{Y})} \sup_{p\in\mathcal{A}} \sum_x p(x)D(W(x)\,\|\,q). \tag{3}$$

Ввиду (2) можно интерпретировать (3) как теорему о минимаксе. Кроме того, если шенноновская пропускная способность для множества ограничений $\mathcal{A}$ конечна, то существует единственная вероятностная мера $q_{1,W,\mathcal{A}}$, называемая шенноновским центром для множества ограничений $\mathcal{A}$, такая что

$$\sup_{p\in\mathcal{A}} I(p;W) = \sup_{p\in\mathcal{A}} \sum_x p(x)D(W(x)\,\|\,q_{1,W,\mathcal{A}}).$$

Термин "центр" происходит из названия соответствующей величины для случая отсутствия ограничений, рассматривавшегося в [12]. Августин доказал аналогичный результат для $I_\alpha(p;W)$ в предположении, что порядок α принадлежит $(0,1]$, а $\mathcal{A}$ – множество ограничений, определяемое ограничениями по стоимости (см. [6, лемма 34.7]). Аналогичное предложение для $I_\alpha(p;W)$ при любом $\alpha \in \mathbb{R}_+$ и любом выпуклом множестве ограничений $\mathcal{A}$ доказывается в § 4 (см. теорему 1). Соответствующую вероятностную меру $q_{\alpha,W,\mathcal{A}}$ мы называем *августиновским центром порядка* α *для множества ограничений* $\mathcal{A}$.

Множества ограничений, определяемые ограничениями по стоимости, довольно часто встречаются при применении августиновской пропускной способности к анализу задач кодирования каналов. Применение техники выпуклого сопряжения

[3] Шенноновский радиус определяется как $\inf\limits_{q\in\mathcal{P}(\mathcal{Y})} \sup\limits_{x\in\mathcal{X}} D(W(x)\,\|\,q)$.

позволяет получить альтернативную характеризацию августиновских пропускной способности и центра при наличии ограничений по стоимости. Августин проделал это в [6, § 35], опираясь на величину, которая ранее применялась для дискретных каналов Галлагером [14, с. 13–15; 15, § 7.3], а также для различных моделей гауссовских каналов[4] в [14, с. 15, 16; 15, §§ 7.4, 7.5; 16; 17]. Мы называем эту величину информацией Реньи – Галлагера и изучаем ее в п. 5.3. По сравнению с применением техники выпуклого сопряжения к шенноновской пропускной способности при наличии ограничений по стоимости, проделанным Чисаром и Кёрнером в [5, гл. 8], анализ Августина в [6, § 35], основанный на информации Реньи – Галлагера, довольно замысловат. В п. 5.2 мы придерживаемся более стандартного подхода, применяя некоторое обобщение метода из [5, гл. 8], основанное на новой величине, которую мы называем информацией Августина – Лежандра. Мы покажем эквивалентность этих двух подходов, используя теоремы о минимаксе, аналогичные вышеупомянутой теореме для августиновской пропускной способности при наличии ограничений.

Некоторые из наиболее важных наблюдений, представленных в настоящей статье, были уже получены ранее в [6, §§ 33–35; 10; 18; 19]. Наши основные достижения в контексте этих работ перечислены в п. 1.3. Перед этим мы вводим обозначения и соглашения в п. 1.1 и описываем рассматриваемую модель в п. 1.2.

1.1. Обозначения и соглашения. Скалярное произведение двух векторов μ и q в $\mathbb{R}^\ell$, т.е. $\sum\limits_{i=1}^{\ell} \mu^i q^i$, обозначается через $\mu \cdot q$. Вектор из всех единиц размерности ℓ будем обозначать через $\mathbb{1}$ для любого $\ell \in \mathbb{Z}_+$, размерность ℓ будет ясна из контекста. Замыкание, внутреннюю область и выпуклую оболочку множества $\mathcal{S}$ будем обозначать через $\operatorname{cl}\mathcal{S}$, $\operatorname{int}\mathcal{S}$ и $\operatorname{ch}\mathcal{S}$ соответственно; соответствующая топология или структура векторного пространства будет ясна из контекста.

Для любого множества $\mathcal{Y}$ обозначим множество всех его подмножеств через $2^{\mathcal{Y}}$, множество всех вероятностных мер на конечных подмножествах $\mathcal{Y}$ – через $\mathcal{P}(\mathcal{Y})$, а множество всех ненулевых конечных мер с тем же свойством – через $\mathcal{M}^+(\mathcal{Y})$. Для любой меры $p \in \mathcal{M}^+(\mathcal{Y})$ множество всех y, таких что $p(y) > 0$, будем называть носителем p и обозначать через $\operatorname{supp} p$.

Для измеримого пространства $(\mathcal{Y}, \mathscr{Y})$ будем обозначать множество всех конечных обобщенных мер через $\mathcal{M}(\mathscr{Y})$, множество всех конечных мер – через $\mathcal{M}_0^+(\mathscr{Y})$, множество всех ненулевых конечных мер – через $\mathcal{M}^+(\mathscr{Y})$, а множество всех вероятностных мер – через $\mathcal{P}(\mathscr{Y})$. Пусть μ и q – две меры на измеримом пространстве $(\mathcal{Y}, \mathscr{Y})$. Тогда μ абсолютно непрерывна относительно q, т.е. $\mu \prec q$, если $\mu(\mathcal{E}) = 0$ для любого $\mathcal{E} \in \mathscr{Y}$, такого что $q(\mathcal{E}) = 0$; μ и q эквивалентны, т.е. $\mu \sim q$, если $\mu \prec q$ и $q \prec \mu$; μ и q взаимно сингулярны, т.е. $\mu \perp q$, если $\exists \mathcal{E} \in \mathscr{Y}$, такое что $\mu(\mathcal{E}) = q(\mathcal{Y} \setminus \mathcal{E}) = 0$. Кроме того, множество мер $\mathcal{W}$ на $(\mathcal{Y}, \mathscr{Y})$ абсолютно непрерывно относительно q, т.е. $\mathcal{W} \prec q$, если $w \prec q$ для любой $w \in \mathcal{W}$, и равномерно абсолютно непрерывно относительно q, т.е. $\mathcal{W} \prec^{\mathrm{uni}} q$, если для любого $\varepsilon > 0$ существует $\delta > 0$, такое что $w(\mathcal{E}) < \varepsilon$ для любой $w \in \mathcal{W}$ при условии, что $q(\mathcal{E}) < \delta$.

Интеграл от измеримой функции f по мере μ будем обозначать через $\int f\,\mu(dy)$ или $\int f(y)\,\mu(dy)$. Интегралы по мере Лебега на вещественной прямой обозначаем через $\int f\,dy$ или $\int f(y)\,dy$. Если μ – вероятностная мера, то интеграл от f по мере μ будем называть математическим ожиданием f и обозначать через $\mathbf{E}_\mu[f]$ или $\mathbf{E}_\mu[f(y)]$.

Некоторые символы будут использоваться в различных смыслах, однако конкретный смысл будет всегда ясен из контекста. Через $\hbar(\cdot)$ будем обозначать как

[4] Августин не предполагал ни конкретной модели шума, ни конечности множества выходов. Тем не менее, гауссовские каналы не охватывались моделью Августина [6, § 35], так как функция стоимости в этой модели предполагалась конечной.

шенноновскую энтропию, так и двоичную: $\hbar(p) \triangleq \sum_y p(y) \ln \frac{1}{p(y)}$ для всех $p \in \mathcal{P}(\mathcal{Y})$ и $\hbar(z) \triangleq z \ln \frac{1}{z} + (1-z) \ln \frac{1}{1-z}$ для всех $z \in [0,1]$. Произведение топологий [20, с. 38], σ-алгебр [20, с. 118] и мер [20, теорема 4.4.4] будем обозначать через $\otimes$, а декартово произведение множеств [20, с. 38] – через $\times$. Будем использовать краткие обозначения $\mathcal{X}_1^n$ для декартова произведения множеств $\mathcal{X}_1, \ldots, \mathcal{X}_n$ и $\mathcal{Y}_1^n$ для произведения σ-алгебр $\mathcal{Y}_1, \ldots, \mathcal{Y}_n$. Через $|\cdot|$ обозначаем абсолютную величину вещественных чисел и мощность множеств. Знаком $\leqslant$ обозначается как обычное отношение для вещественных чисел, так и соответствующее поточечное неравенство для функций и векторов. Для двух мер μ и q на измеримом пространстве $(\mathcal{Y}, \mathcal{Y})$ пишем $\mu \leqslant q$, если $\mu(\mathcal{E}) \leqslant q(\mathcal{E})$ для всех $\mathcal{E} \in \mathcal{Y}$.

Для $a, b \in \mathbb{R}$ через $a \wedge b$ будем обозначать минимум a и b. Для $f\colon \mathcal{Y} \to \mathbb{R}$ и $g\colon \mathcal{Y} \to \mathbb{R}$ функция $f \wedge g$ – поточечный минимум f и g. Для $\mu, q \in \mathcal{M}(\mathcal{Y})$ мера $\mu \wedge q$ – единственная мера, удовлетворяющая ν-п.в. условию $\frac{d\mu \wedge q}{d\nu} = \frac{d\mu}{d\nu} \wedge \frac{dq}{d\nu}$ для любой ν, такой что $\mu \prec \nu$ и $q \prec \nu$. Для набора $\mathcal{F}$ вещественнозначных функций через $\bigwedge_{f \in \mathcal{F}} f$ обозначаем поточечный инфимум функций f в $\mathcal{F}$ – расширенную вещественнозначную функцию. Для набора мер $\mathcal{U} \subset \mathcal{M}(\mathcal{Y})$, такого что $w \leqslant u$ для всех $u \in \mathcal{U}$ для некоторой $w \in \mathcal{P}(\mathcal{Y})$, через $\bigwedge_{u \in \mathcal{U}} u$ обозначим инфимум $\mathcal{U}$ относительно частичного порядка $\leqslant$; такая инфимум-мера существует и единственна согласно [21, теорема 4.7.5]. Символ $\vee$ используется для максимумов и супремумов аналогично символу $\wedge$ для минимумов и инфимумов.

1.2. Модель канала. *Канал* W – это функция из *множества входов* $\mathcal{X}$ в множество всех вероятностных мер на *выходном пространстве* $(\mathcal{Y}, \mathcal{Y})$:

$$W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y}). \tag{4}$$

Множество $\mathcal{Y}$ называется *множеством выходов*, а $\mathcal{Y}$ – *σ-алгеброй выходных событий*. Множество всех каналов из множества входов $\mathcal{X}$ в выходное пространство $(\mathcal{Y}, \mathcal{Y})$ обозначим через $\mathcal{P}(\mathcal{Y}\,|\,\mathcal{X})$. Для любых $p \in \mathcal{P}(\mathcal{X})$ и $W \in \mathcal{P}(\mathcal{Y}\,|\,\mathcal{X})$ вероятностная мера с маргинальным распределением на $\mathcal{X}$, равным p, и условным распределением при условии x, равным $W(x)$, обозначается через $p \circledast W$. Вплоть до п. 5.4 мы ограничиваемся рассмотрением входных распределений, принадлежащих $\mathcal{P}(\mathcal{X})$, избегая тонкостей, связанных с измеримостью. Более общий случай входных распределений из $\mathcal{P}(\mathcal{X})$ рассмотрен[5] в п. 5.4.

Канал W называется *дискретным каналом*, если и $\mathcal{X}$, и $\mathcal{Y}$ – конечные множества. Для любого $n \in \mathbb{Z}_+$ и любых каналов $W_t\colon \mathcal{X}_t \to \mathcal{P}(\mathcal{Y}_t)$, $t \in \{1, \ldots, n\}$, *канал-произведение* $W_{[1,n]}\colon \mathcal{X}_1^n \to \mathcal{P}(\mathcal{Y}_1^n)$ *длины* n определяется следующим образом:

$$W_{[1,n]}(x_1^n) = \bigotimes_{t=1}^{n} W_t(x_t) \quad \forall x_1^n \in \mathcal{X}_1^n.$$

Канал-произведение является *стационарным*, если $W_t = W$ при всех $t \in \{1, \ldots, n\}$ для некоторого $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$.

Для любого $\ell \in \mathbb{Z}_+$ ℓ-мерная *функция стоимости* ρ – это функция из множества входов в $\mathbb{R}^\ell$, ограниченная снизу, т.е. функция вида $\rho\colon \mathcal{X} \to \mathbb{R}^\ell_{\geqslant z}$ для некоторого

[5] Структуры, описанной в (4), самой по себе недостаточно для существования и единственности меры $p \circledast W$ с требуемыми свойствами для всех $p \in \mathcal{P}(\mathcal{X})$. Существование и единственность такой меры $p \circledast W$ гарантировано для всех $p \in \mathcal{P}(\mathcal{X})$, если W – функция вероятностей переходов из $(\mathcal{X}, \mathcal{X})$ в $(\mathcal{Y}, \mathcal{Y})$, т.е. элемент $\mathcal{P}(\mathcal{Y}\,|\,\mathcal{X})$, а не $\mathcal{P}(\mathcal{Y}\,|\,\mathcal{X})$.

$z \in \mathbb{R}$. Без ограничения общности будем предполагать[6], что

$$\inf_{x \in \mathcal{X}} \rho^i(x) \geqslant 0 \quad \forall i \in \{1, \ldots, \ell\}.$$

Множество всех ограничений по стоимости, которые выполняются хотя бы для одного элемента из $\mathcal{X}$, обозначим через $\Gamma_\rho^{\mathrm{ex}}$, а множество всех ограничений по стоимости, выполненных для хотя бы одного элемента из $\mathcal{P}(\mathcal{X})$, – через Γ_ρ:

$$\Gamma_\rho^{\mathrm{ex}} \triangleq \left\{ \varrho \in \mathbb{R}_{\geqslant 0}^{\ell} : \exists x \in \mathcal{X}, \text{ такой что } \rho(x) \leqslant \varrho \right\}, \tag{5}$$

$$\Gamma_\rho \triangleq \left\{ \varrho \in \mathbb{R}_{\geqslant 0}^{\ell} : \exists p \in \mathcal{P}(\mathcal{X}), \text{ такой что } \sum_x p(x)\rho(x) \leqslant \varrho \right\}. \tag{6}$$

Тогда оба множества $\Gamma_\rho^{\mathrm{ex}}$ и Γ_ρ имеют непустые внутренние области, причем Γ_ρ является выпуклой оболочкой $\Gamma_\rho^{\mathrm{ex}}$, т.е. $\Gamma_\rho = \operatorname{ch} \Gamma_\rho^{\mathrm{ex}}$.

Функция стоимости на канале-произведении называется аддитивной, если ее можно представить в виде суммы функций стоимости на каналах-компонентах. Для заданных $W_t \colon \mathcal{X}_t \to \mathcal{P}(\mathcal{Y}_t)$ и $\rho_t \colon \mathcal{X}_t \to \mathbb{R}_{\geqslant 0}^{\ell}$, $t \in \{1, \ldots, n\}$, обозначим результирующую аддитивную функцию стоимости на $\mathcal{X}_1^n$ для канала $W_{[1,n]}$ через $\rho_{[1,n]}$, т.е.

$$\rho_{[1,n]}(x_1^n) = \sum_{t=1}^{n} \rho_t(x_t) \quad \forall x_1^n \in \mathcal{X}_1^n.$$

1.3. Предшествующие работы и основные результаты. Приведем подробный список наших результатов, важный для ясного понимания августиновских информационных величин и связанных с ними результатов, упомянутых выше.

I. Для всех $\alpha \in (0,1)$ в [6, лемма 34.2] утверждается существование и единственность вероятностной меры $q_{\alpha,p}$, такой что $I_\alpha(p;W) = D_\alpha(W \,\|\, q_{\alpha,p} \,|\, p)$, и дана следующая характеризация $q_{\alpha,p}$ в терминах оператора[7] $\mathrm{T}_{\alpha,p}(\cdot)$:

- $\mathrm{T}_{\alpha,p}(q_{\alpha,p}) = q_{\alpha,p}$ и $q_{\alpha,p} \sim q_{1,p}$;
- Если $q_{1,p} \prec q$ и $\mathrm{T}_{\alpha,p}(q) = q$, то $q_{\alpha,p} = q$;
- $\lim\limits_{j \to \infty} \|q_{\alpha,p} - \mathrm{T}_{\alpha,p}^{j}(q_{1,p})\| = 0$;
- $D_\alpha(W \,\|\, q \,|\, p) \geqslant I_\alpha(p;W) + D_\alpha(q_{\alpha,p} \,\|\, q)$ для[8] всех $q \in \mathcal{P}(\mathcal{Y})$.

Мы не смогли проверить корректность доказательства леммы 34.2 из [6]; это обсуждается в [22, Приложение C]. Лемма 13(c) настоящей статьи доказывается[9] с использованием идей Августина из доказательства леммы 34.2 в [6]. Из нашей леммы 13(c) вытекают все утверждения леммы 34.2 из [6], кроме $\lim\limits_{j \to \infty} \|q_{\alpha,p} - \mathrm{T}_{\alpha,p}^{j}(q_{1,p})\| = 0$; вместо этого в лемме 13(c) установлено равенство $\lim\limits_{j \to \infty} \|q_{\alpha,p} - \mathrm{T}_{\alpha,p}^{j}(q_{\alpha,p}^{\mathrm{g}})\| = 0$ (см. (26) и [22, замечание 6]). В отличие от леммы 34.2 из [6], лемма 13(c) также дает оценку сверху на $D_\alpha(W \,\|\, q \,|\, p)$. Насколько нам известно, это новая оценка. Верхняя и нижняя оценки на $D_\alpha(W \,\|\, q \,|\, p)$, полученные в лемме 13(c),(d), таковы:

$$D_{1 \vee \alpha}(q_{\alpha,p} \,\|\, q) \geqslant D_\alpha(W \,\|\, q \,|\, p) - I_\alpha(p;W) \geqslant D_{1 \wedge \alpha}(q_{\alpha,p} \,\|\, q) \quad \forall q \in \mathcal{P}(\mathcal{Y}). \tag{7}$$

---

[6] Августин в [6, § 33] накладывал дополнительное условие $\bigvee_{x \in \mathcal{X}} \rho(x) \leqslant \mathbb{1}$, однако это условие исключает из рассмотрения некоторые важные случаи, такие как гауссовские каналы.

[7] Оператор $\mathrm{T}_{\alpha,p}(\cdot)$, определенный в (20), однозначно задается порядком α и распределением p и корректно определен для всех q с конечным $D_\alpha(W \,\|\, q \,|\, p)$.

[8] Строго говоря, в [6, лемма 34.2] приведено неравенство $D_\alpha(W \,\|\, q \,|\, p) \geqslant I_\alpha(p;W) + \frac{\alpha}{2}\|q_{\alpha,p} - q\|^2$. Однако Августин вначале доказывает вышеуказанное неравенство, а затем с помощью неравенства Пинскера получает неравенство, приведенное в [6, лемма 34.2].

[9] Лемму 13(c) можно также доказать, используя идеи доказательства леммы 13(d).

Для случая конечного $\mathcal{Y}$ существование $q \in \mathcal{P}(\mathcal{Y})$, такой что $q \sim q_{1,p}$ и $\mathrm{T}_{\alpha,p}(q) = q$, обсуждалось и другими авторами. Сделаем краткое отступление и укажем на соответствующие обсуждения этого вопроса о существовании.

- При доказательстве границы сферической упаковки для кодов с постоянной композицией в дискретных стационарных каналах-произведениях Фано неявно утверждал существование неподвижной точки, эквивалентной $q_{1,p}$, для любого $\alpha \in (0,1)$ (см. [23, § 9.2, формула (9.24) и с. 292]. Однако в [23, § 9.2] Фано не объяснял, почему такая неподвижная точка должна существовать, и не изучал ее единственность и связь с $q_{\alpha,p}$;
- При доказательстве эквивалентности своей экспоненты сферической упаковки в случае конечного $\mathcal{Y}$ и соответствующей величины, полученной Фано в [23], Арутюнян доказывал существование неподвижной точки, эквивалентной $q_{1,p}$, для любого $\alpha \in (0,1)$ (см. [18, формулы (16)–(19)]);
- При обсуждении границ случайного кодирования для дискретных стационарных каналов-произведений Полтырев делал утверждение, равносильное утверждению о существовании для любого $\alpha \in [1/2, 1)$ неподвижной точки, эквивалентной $q_{1,p}$ (см. [19, формулы (3.15), (3.16) и теорема 3.2]. Однако Полтырев не формулировал это в виде утверждения о неподвижной точке.

На наш взгляд, основное концептуальное достижение в [6, лемма 34.2] – это характеризация августиновского среднего как неподвижной точки оператора $\mathrm{T}_{\alpha,p}(\cdot)$, эквивалентной $q_{1,p}$. Границы, такие как (7), вытекают из этого наблюдения с помощью неравенства Йенсена.

II. Для $\alpha \in (1,\infty)$ в лемме 13(d) установлено существование и единственность августиновского среднего $q_{\alpha,p}$ и доказано, что оно удовлетворяет границе (7) и обладает следующими свойствами:

- $\mathrm{T}_{\alpha,p}(q_{\alpha,p}) = q_{\alpha,p}$ и $q_{\alpha,p} \sim q_{1,p}$;
- Если $\mathrm{T}_{\alpha,p}(q) = q$, то $q_{\alpha,p} = q$.

Насколько нам известно, лемма 13(d) является новой. Для случая $\alpha \in (1,\infty)$ ни характеризация $q_{\alpha,p}$ через $\mathrm{T}_{\alpha,p}(\cdot)$, ни неравенства (7) раньше не были известны, даже для случая конечного $\mathcal{Y}$.

III. Функция $I_\alpha(p;W)$ из $\mathbb{R}_+$ в $[0, \hbar(p)]$ непрерывно дифференцируема по α согласно лемме 17(e).

IV. В теореме 1 получено следующее минимаксное тождество для любого выпуклого множества ограничений $\mathcal{A}$:

$$\sup_{p\in\mathcal{A}} \inf_{q\in\mathcal{P}(\mathcal{Y})} D_\alpha(W \,\|\, q \,|\, p) = \inf_{q\in\mathcal{P}(\mathcal{Y})} \sup_{p\in\mathcal{A}} D_\alpha(W \,\|\, q \,|\, p).$$

Теорема 1 устанавливает существование и единственность августиновского центра $q_{\alpha,W,\mathcal{A}}$ для любого выпуклого множества $\mathcal{A}$ с конечной августиновской пропускной способностью, а также сходимость $\{q_{\alpha,p^{(i)}}\}_{i\in\mathbb{Z}_+}$ к $q_{\alpha,W,\mathcal{A}}$ в топологии полной вариации для любой последовательности $\{p^{(i)}\}_{i\in\mathbb{Z}_+} \subset \mathcal{A}$, такой что $\lim_{i\to\infty} I_\alpha(p^{(i)};W) = C_{\alpha,W,\mathcal{A}}$. Августином этот результат был доказан только для $\alpha \in (0,1]$ и множеств ограничений, определяемых ограничениями по стоимости (см. [6, лемма 34.7]). Для случая $\mathcal{A} = \mathcal{P}(\mathcal{X})$ аналогичные результаты были доказаны в [2, предложение 1] в предположении, что $\mathcal{X}$ и $\mathcal{Y}$ – конечные множества, а также в [8, теорема 34] в предположении, что множество $\mathcal{Y}$ конечно.

V. Следующая граница в терминах августиновских пропускной способности и центра, установленная в лемме 21, насколько нам известно, является новой:

$$\sup_{p\in\mathcal{A}} D_\alpha(W \,\|\, q \,|\, p) \geqslant C_{\alpha,W,\mathcal{A}} + D_{\alpha\wedge 1}(q_{\alpha,W,\mathcal{A}} \,\|\, q) \quad \forall q \in \mathcal{P}(\mathcal{Y}).$$

Аналогичная граница была выдвинута в качестве гипотезы ван Эрвеном и Харремоесом в [8]. Для пропускной способности и центра Реньи эта гипотеза доказана и обобщена на случай наличия ограничений в [13, леммы 19, 25].

VI. Информация Августина – Лежандра $I_\alpha^\lambda(p;W)$, определяемая как $I_\alpha(p;W) - \lambda \cdot \mathbf{E}_p[\rho]$, а также соответствующие пропускная способность, центр и радиус, – новые понятия, не изучавшиеся ранее, кроме случая $\alpha = 1$. Таким образом, формально говоря, все утверждения п. 5.2 являются новыми. Анализ, проведенный в п. 5.2, является стандартным применением техники выпуклого сопряжения для характеризации августиновских пропускной способности и центра при наличии ограничений по стоимости. Аналогичный анализ в случае $\alpha = 1$ проведен в [5, гл. 8] для дискретных каналов с одним ограничением по стоимости. Наиболее важные результаты проведенного в п. 5.2 анализа таковы:

- Величина $C_{\alpha,W}^\lambda$, определяемая как $\sup\limits_{p\in\mathcal{P}(\mathcal{X})} I_\alpha^\lambda(p;W)$, удовлетворяет равенству $C_{\alpha,W}^\lambda = \sup\limits_{\varrho\geqslant 0} C_{\alpha,W,\varrho} - \lambda\cdot\varrho$ для всех $\lambda\in\mathbb{R}_{\geqslant 0}^\ell$ – формула (55);
- $C_{\alpha,W,\varrho} = \inf\limits_{\lambda\geqslant 0} C_{\alpha,W}^\lambda + \lambda\cdot\varrho$ для всех $\varrho\in\operatorname{int}\Gamma_\rho$, причем множество всех λ, доставляющих этот инфимум, является непустым, выпуклым и компактным, если $C_{\alpha,W,\varrho}$ конечно, – лемма 29;
- $C_{\alpha,W}^\lambda = S_{\alpha,W}^\lambda$, где $S_{\alpha,W}^\lambda$ определяется как $\inf\limits_{q\in\mathcal{P}(\mathcal{Y})}\sup\limits_{x\in\mathcal{X}} D_\alpha(W(x)\,\|\,q) - \lambda\cdot\rho(x)$, – теорема 2;
- Если $C_{\alpha,W}^\lambda < \infty$, то существует единственный АЛ-центр $q_{\alpha,W}^\lambda$, такой что $C_{\alpha,W}^\lambda = \sup\limits_{x\in\mathcal{X}} D_\alpha\big(W(x)\,\|\,q_{\alpha,W}^\lambda\big) - \lambda\cdot\rho(x)$. Кроме того, $\lim\limits_{i\to\infty}\|q_{\alpha,p} - q_{\alpha,W}^\lambda\| = 0$ для всех $\{p^{(i)}\}_{i\in\mathbb{Z}_+}\subset\mathcal{P}(\mathcal{X})$, таких что $\lim\limits_{i\to\infty} I_\alpha^\lambda(p^{(i)};W) = C_{\alpha,W}^\lambda$, – теорема 2;
- Если $C_{\alpha,W,\varrho} = C_{\alpha,W}^\lambda + \lambda\cdot\varrho < \infty$ для некоторого $\lambda\in\mathbb{R}_{\geqslant 0}^\ell$, то $q_{\alpha,W,\varrho} = q_{\alpha,W}^\lambda$ – лемма 31;
- Для канала-произведения $W_{[1,n]}$ с аддитивной функцией стоимости $C_{\alpha,W_{[1,n]}}^\lambda = \sum\limits_{t=1}^{n} C_{\alpha,W_t}^\lambda$ для всех $\lambda\in\mathbb{R}_{\geqslant 0}^\ell$ и $\alpha\in\mathbb{R}_+$, и если какая-либо из этих величин существует, то $q_{\alpha,W_{[1,n]}}^\lambda$ равно $\bigotimes\limits_{t=1}^{n} q_{\alpha,W_t}^\lambda$ – лемма 32.

VII. Информация Реньи – Галлагера $I_\alpha^{\mathrm{g}\lambda}(p;W)$ является обобщением информации Реньи $I_\alpha^{\mathrm{g}}(p;W)$ с множителем Лагранжа, поскольку $I_\alpha^{\mathrm{g}0}(p;W) = I_\alpha^{\mathrm{g}}(p;W)$. Эта величина впервые была использована Галлагером в [14] при другой параметризации и другом масштабировании; в дальнейшем она рассматривалась в [24, § IV; 6; 16; 17; 25–27] также с другими параметризациями и масштабированиями под различными названиями. Мы выбираем такие же параметризацию и масштабирование, как и для информации Августина – Лежандра. Наиболее важные результаты анализа в п. 5.3 таковы:

- $C_{\alpha,W}^{\mathrm{g}\lambda} = S_{\alpha,W}^\lambda$ согласно теореме 3, где $C_{\alpha,W}^{\mathrm{g}\lambda}$ определяется как $\sup\limits_{p\in\mathcal{P}(\mathcal{X})} I_\alpha^{\mathrm{g}\lambda}(p;W)$;
- Если $C_{\alpha,W}^\lambda < \infty$ и $\lim\limits_{i\to\infty} I_\alpha^{\mathrm{g}}\big(p^{(i)};W\big) = C_{\alpha,W}^\lambda$, то $\lim\limits_{i\to\infty}\|q_{\alpha,p}^{\mathrm{g}\lambda} - q_{\alpha,W}^\lambda\| = 0$ – теорема 3;
- $\sup\limits_{x\in\mathcal{X}} D_\alpha(W(x)\,\|\,q) - \lambda\cdot\rho(x) \geqslant C_{\alpha,W}^\lambda + D_\alpha\big(q_{\alpha,W}^\lambda\,\|\,q\big)$ для всех $q\in\mathcal{P}(\mathcal{Y})$ – лемма 35.

Насколько нам известно, лемма 35 является новой. В случае, когда $\alpha\in(0,1)$ и $\bigvee\limits_{x\in\mathcal{X}}\rho(x)\leqslant\mathbb{1}$, теорема 3 вытекает из [6, лемма 35.2].

При проведении аналогичного анализа в [6, § 35] Августин предполагал функцию стоимости ограниченной. Однако такое предположение исключает из рассмотрения такие важные и интересные случаи как гауссовские каналы. Это не вопрос смены масштабирования: некоторые результаты проделанного Августином анализа, например, [6, лемма 35.3(a)], не верны, когда функция стоимости неограничена. Мы не будем предполагать функцию стоимости ограниченной. Таким образом, наша модель включает в себя не только модель Августина из [6, § 35], но и другие ранее изучавшиеся модели: как дискретные [14, с. 13–15; 15, § 7.3; 24, § IV; 26; 27], так и гауссовские [14, с. 15, 16; 15, §§ 7.4, 7.5; 16; 17; 25].

VIII. Для каналов с несчетным множеством входов шенноновская информация и пропускная способность часто определяются через вероятностные меры на входном пространстве $(\mathcal{X}, \mathscr{X})$, а не через функции вероятности на множестве входов $\mathcal{X}$. В п. 5.4 обсуждается, как и при каких условиях можно проделать такое обобщение для августиновских информационных величин. Наиболее важные результаты нашего анализа таковы:

- Если W – функция вероятностей переходов из $(\mathcal{X}, \mathscr{X})$ в $(\mathcal{Y}, \mathscr{Y})$, т.е. $W \in \mathcal{P}(\mathscr{Y} \,|\, \mathscr{X})$, и $\mathscr{Y}$ счетно порождена, то
  - величина $I_\alpha(p; W)$ корректно определена для всех $\alpha \in \mathbb{R}_+$ и $p \in \mathcal{P}(\mathscr{X})$ – формулы (75), (76) и лемма 37;
  - величина $I_\alpha^\lambda(p; W)$ корректно определена для всех $\alpha \in \mathbb{R}_+$, $p \in \mathcal{P}(\mathscr{X})$ и $\lambda \in \mathbb{R}_{\geqslant 0}^\ell$, если функция ρ является $\mathscr{X}$-измеримой – формула (77);
- Если $W \in \mathcal{P}(\mathscr{Y} \,|\, \mathscr{X})$, $\mathscr{X}$ счетно отделима, $\mathscr{Y}$ счетно порождена, а ρ является $\mathscr{X}$-измеримой, то
  - $C_{\alpha,W}^\lambda = \sup\limits_{p \in \mathcal{A}^\lambda} I_\alpha^\lambda(p; W)$ для всех $\lambda \in \mathbb{R}_{\geqslant 0}^\ell$, где $\mathcal{A}^\lambda$ определяется как $\{p \in \mathcal{P}(\mathscr{X}) :$ $\lambda \cdot \mathbf{E}_p[\rho] < \infty\}$, – теорема 4;
  - Если $C_{\alpha,W}^\lambda < \infty$ для некоторого $\lambda \in \mathbb{R}_{\geqslant 0}^\ell$, то $C_{\alpha,W}^\lambda = \sup\limits_{p \in \mathcal{A}^\lambda} D_\alpha\big(W \,\|\, q_{\alpha,W}^\lambda \,|\, p\big) - \lambda \cdot \mathbf{E}_p[\rho]$ – теорема 4;
  - $C_{\alpha,W,\varrho} = \sup\limits_{p \in \mathcal{A}(\varrho)} I_\alpha(p; W)$ для всех $\varrho \in \operatorname{int} \Gamma_\rho$, где $\mathcal{A}(\varrho)$ определяется как $\{p \in \mathcal{P}(\mathscr{X}) :\ \mathbf{E}_p[\rho] \leqslant \varrho\}$, – теорема 5;
  - Если $C_{\alpha,W,\varrho} < \infty$ для некоторого $\varrho \in \operatorname{int} \Gamma_\rho$, то $C_{\alpha,W,\varrho} = \sup\limits_{p \in \mathcal{A}(\varrho)} D_\alpha(W \,\|\, q_{\alpha,W,\varrho} \,|\, p)$ – теорема 5.

Таким образом, АЛ-пропускная способность и АЛ-центр, так же как и августиновские пропускная способность и центр при наличии ограничений по стоимости, определяемые через функции вероятности, равны соответствующим величинам, которые можно определить через вероятностные меры на $(\mathcal{X}, \mathscr{X})$ при условии, что $\mathscr{X}$ счетно отделима, а $\mathscr{Y}$ счетно порождена.

§ 2. Предварительные сведения

2.1. Расхождение Реньи.

Определение 1. Для любых $\alpha \in \mathbb{R}_+$ и $w, q \in \mathcal{M}^+(\mathscr{Y})$ *расхождением Реньи порядка α между w и q* называется

$$D_\alpha(w \,\|\, q) \triangleq \begin{cases} \dfrac{1}{\alpha - 1} \ln \displaystyle\int \left(\frac{dw}{d\nu}\right)^{\alpha} \left(\frac{dq}{d\nu}\right)^{1-\alpha} \nu(dy), & \alpha \neq 1, \\ \displaystyle\int \frac{dw}{d\nu} \left[\ln \frac{dw}{d\nu} - \ln \frac{dq}{d\nu}\right] \nu(dy), & \alpha = 1, \end{cases} \tag{8}$$

где ν – любая мера, такая что $w \prec \nu$ и $q \prec \nu$.

Обычно расхождение Реньи определяется для пар вероятностных мер (см., например, [8,28]), а не для пар ненулевых конечных мер. Мы пользуемся этим несколько более общим определением, поскольку оно позволяет выражать с помощью расхождения Реньи некоторые результаты в более кратком виде (см. следующую лемму 1 и п. 5.3). Для пар вероятностных мер определение 1 равносильно обычному определению, используемому в [8, теорема 5].

Лемма 1 [13, лемма 8]. *Пусть α – положительное вещественное число, а w, q, v – ненулевые конечные меры на $(\mathcal{Y}, \mathscr{Y})$.*

- *Если $v \leqslant q$, то $D_\alpha(w \,\|\, q) \leqslant D_\alpha(w \,\|\, v)$;*
- *Если $q = \gamma v$ для некоторого $\gamma \in \mathbb{R}_+$ и либо w – вероятностная мера, либо $\alpha \neq 1$, то $D_\alpha(w \,\|\, q) = D_\alpha(w \,\|\, v) - \ln \gamma$.*

Следующая лемма утверждает, что если оба аргумента расхождения Реньи – вероятностные меры, то оно положительно, кроме случая, когда аргументы равны.

Лемма 2 [8, теоремы 3 и 31]. *Для любого $\alpha \in \mathbb{R}_+$ и любых вероятностных мер w и q на $(\mathcal{Y}, \mathscr{Y})$*

$$D_\alpha(w \,\|\, q) \geqslant \frac{1 \wedge \alpha}{2} \|w - q\|^2.$$

Для порядков, принадлежащих $(0, 1]$, это неравенство называется неравенством Пинскера [29,30]. Для порядков, принадлежащих $(0, 1)$, можно оценить расхождение Реньи сверху через расстояние полной вариации. Для случая $\alpha = 1/2$ формула (21) в [31, с. 364] утверждает, что

$$D_{1/2}(w \,\|\, q) \leqslant 2 \ln \frac{2}{2 - \|w - q\|}. \tag{9}$$

Расхождение Реньи порядка α непрерывно по своим аргументам в топологии полной вариации при условии, что $\alpha \in (0, 1)$. Для произвольных порядков имеет место лишь полунепрерывность снизу, однако она выполнена даже в топологии помножественной (setwise) сходимости.

Лемма 3 [8, теорема 15]. *Для любого $\alpha \in \mathbb{R}_+$ функция $D_\alpha(w \,\|\, q)$ полунепрерывна снизу по паре вероятностных мер (w, q) в топологии помножественной сходимости.*

Лемма 4 [8, теорема 17]. *Для любого $\alpha \in (0, 1)$ функция $D_\alpha(w \,\|\, q)$ равномерно непрерывна по паре вероятностных мер (w, q) в топологии полной вариации.*

Расхождение Реньи выпукло по второму аргументу для всех положительных порядков, совместно выпукло по двум аргументам для положительных порядков не выше единицы и совместно квазивыпукло по двум аргументам для всех положительных порядков.

Лемма 5 [8, теорема 12]. *Для любых $\alpha \in \mathbb{R}_+$, $w, q_0, q_1 \in \mathcal{P}(\mathscr{Y})$, $\beta \in (0, 1)$ и ν, таких что $(q_0 + q_1) \prec \nu$,*

$$D_\alpha(w \,\|\, \beta q_1 + (1 - \beta) q_0) \leqslant \beta D_\alpha(w \,\|\, q_1) + (1 - \beta) D_\alpha(w \,\|\, q_0).$$

Кроме того, равенство имеет место тогда и только тогда, когда $\frac{dq_1}{d\nu} = \frac{dq_0}{d\nu}$ w-почти наверное.

Лемма 6 [8, теорема 11]. *Для любых $\alpha \in (0, 1]$, $w_0, w_1, q_0, q_1 \in \mathcal{P}(\mathscr{Y})$, $\beta \in (0, 1)$ и ν, таких что $(w_0 + w_1 + q_0 + q_1) \prec \nu$,*

$$D_\alpha(\beta w_1 + (1 - \beta) w_0 \,\|\, \beta q_1 + (1 - \beta) q_0) \leqslant \beta D_\alpha(w_1 \,\|\, q_1) + (1 - \beta) D_\alpha(w_0 \,\|\, q_0). \tag{10}$$

Кроме того, при $\alpha = 1$ *равенство имеет место тогда и только тогда, когда* $\dfrac{dw_0}{d\nu}\dfrac{dq_1}{d\nu} = \dfrac{dw_1}{d\nu}\dfrac{dq_0}{d\nu}$, *а при* $\alpha \in (0,1)$ *равенство имеет место тогда и только тогда, когда* $\dfrac{dw_0}{d\nu}\dfrac{dq_1}{d\nu} = \dfrac{dw_1}{d\nu}\dfrac{dq_0}{d\nu}$ *и* $D_\alpha(w_1 \,\|\, q_1) = D_\alpha(w_0 \,\|\, q_0)$.

Лемма 7 [8, теорема 13]. *Для любых* $\alpha \in \mathbb{R}_+$, $w_0, w_1, q_0, q_1 \in \mathcal{P}(\mathcal{Y})$ *и* $\beta \in (0,1)$

$$D_\alpha(\beta w_1 + (1-\beta)w_0 \,\|\, \beta q_1 + (1-\beta)q_0) \leqslant D_\alpha(w_1 \,\|\, q_1) \vee D_\alpha(w_0 \,\|\, q_0).$$

Лемма 8 [8, теоремы 3 и 7]. *Для любых* $w, q \in \mathcal{P}(\mathcal{Y})$ *функция* $D_\alpha(w \,\|\, q)$ *не убывает и полунепрерывна снизу по* α *на* $\mathbb{R}_+$ *и непрерывна на* $(0, (1 \vee \chi_{w,q})]$, *где* $\chi_{w,q} \triangleq \sup\{\alpha : D_\alpha(w \,\|\, q) < \infty\}$.

Так как $D_\alpha(w \,\|\, q) = \dfrac{\alpha}{1-\alpha} D_{1-\alpha}(q \,\|\, w)$ для всех $\alpha \in (0,1)$, то из леммы 8 и формулы (9) следует

$$D_\alpha(w \,\|\, q) \leqslant \begin{cases} D_{1/2}(w \,\|\, q), & \text{если } \alpha \in (0, 1/2], \\ \dfrac{\alpha}{1-\alpha} D_{1/2}(w \,\|\, q), & \text{если } \alpha \in (1/2, 1), \end{cases} \leqslant$$
$$\leqslant \frac{2}{1-\alpha} \ln \frac{2}{2 - \|w - q\|} \quad \forall \alpha \in (0,1).$$

Чуть более точную границу см. в [31, формула (24), с. 365].

Если $\mathcal{G}$ – σ-подалгебра $\mathcal{Y}$, то для любых w и q из $\mathcal{P}(\mathcal{Y})$ равенства $w|_{\mathcal{G}}(\mathcal{E}) = w(\mathcal{E})$ для всех $\mathcal{E} \in \mathcal{G}$ и $q|_{\mathcal{G}}(\mathcal{E}) = q(\mathcal{E})$ для всех $\mathcal{E} \in \mathcal{G}$ однозначно определяют вероятностные меры $w|_{\mathcal{G}}$ и $q|_{\mathcal{G}}$ на $(\mathcal{Y}, \mathcal{G})$. Будем обозначать $D_\alpha\big(w|_{\mathcal{G}} \,\|\, q|_{\mathcal{G}}\big)$ через $D_\alpha^{\mathcal{G}}(w \,\|\, q)$.

Лемма 9 [8, теорема 21]. *Пусть* $\mathcal{Y}_1 \subset \mathcal{Y}_2 \subset \ldots \subset \mathcal{Y}$ – *возрастающее семейство* σ*-алгебр, и пусть* $\mathcal{Y}_\infty = \sigma\Big(\bigcup\limits_{i=1}^{\infty} \mathcal{Y}_i\Big)$ – *наименьшая* σ*-алгебра, содержащая их. Тогда для любого порядка* $\alpha \in \mathbb{R}_+$

$$\lim_{i\to\infty} D_\alpha^{\mathcal{Y}_i}(w \,\|\, q) = D_\alpha^{\mathcal{Y}_\infty}(w \,\|\, q).$$

2.2. Скошенная вероятностная мера.

Определение 2. Для любого $\alpha \in \mathbb{R}_+$ и любых $w, q \in \mathcal{P}(\mathcal{Y})$, таких что $D_\alpha(w \,\|\, q) < \infty$, *скошенной вероятностной мерой* w_α^q *порядка* α называется

$$\frac{dw_\alpha^q}{d\nu} \triangleq e^{(1-\alpha)D_\alpha(w\|q)} \left(\frac{dw}{d\nu}\right)^{\alpha} \left(\frac{dq}{d\nu}\right)^{1-\alpha}.$$

Отметим, что $w_1^q = w$ для любой q, такой что $D_1(w \,\|\, q) < \infty$. Для остальных порядков непосредственной подстановкой можно проверить следующее утверждение: если $D_\alpha(w \,\|\, q) < \infty$, то для любой $v \in \mathcal{P}(\mathcal{Y})$, такой что $D_1(v \,\|\, w) < \infty$ и $D_1(v \,\|\, q) < \infty$, справедливо

$$\frac{1}{1-\alpha} D_1\big(v \,\|\, w_\alpha^q\big) + D_\alpha(w \,\|\, q) = \frac{\alpha}{1-\alpha} D_1(v \,\|\, w) + D_1(v \,\|\, q).$$

Это тождество используется для вывода следующей вариационной характеризации расхождения Реньи порядка, отличного от единицы.

Л е м м а 10 [8, теорема 30]. *Для любых* $w, q \in \mathcal{P}(\mathcal{Y})$

$$D_\alpha(w \,\|\, q) = \begin{cases} \inf\limits_{v \in \mathcal{P}(\mathcal{Y})} \dfrac{\alpha}{1-\alpha} D_1(v \,\|\, w) + D_1(v \,\|\, q), & \alpha \in (0,1), \\ \sup\limits_{v \in \mathcal{P}(\mathcal{Y})} \dfrac{\alpha}{1-\alpha} D_1(v \,\|\, w) + D_1(v \,\|\, q), & \alpha \in (1,\infty), \end{cases}$$

где $\dfrac{\alpha}{1-\alpha} D_1(v \,\|\, w) + D_1(v \,\|\, q)$ *равно* $-\infty$ *при* $\alpha \in (1,\infty)$ *и* $D_1(v \,\|\, w) = D_1(v \,\|\, q) = \infty$. *Кроме того, если* $D_\alpha(w \,\|\, q)$ *конечно и либо* $\alpha \in (0,1)$, *либо* $D_1\big(w_\alpha^q \,\|\, w\big) < \infty$, *то*

$$D_\alpha(w \,\|\, q) = \frac{\alpha}{1-\alpha} D_1\big(w_\alpha^q \,\|\, w\big) + D_1\big(w_\alpha^q \,\|\, q\big). \tag{11}$$

В лемме 8 мы видели, что $D_\alpha(w \,\|\, q)$ непрерывна по α на замыкании того интервала, где она конечна. Далее в лемме 11 устанавливается аналитичность $D_\alpha(w \,\|\, q)$ по α на внутренности того интервала, где $D_\alpha(w \,\|\, q)$ конечна. Лемма 11 также устанавливает аналитичность, а следовательно, и конечность, функций $D_1\big(w_\alpha^q \,\|\, w\big)$ и $D_1\big(w_\alpha^q \,\|\, q\big)$ на том же интервале. Это позволяет утверждать, что и равенство (11) справедливо на том же интервале:

$$D_\alpha(w \,\|\, q) = \frac{\alpha}{1-\alpha} D_1\big(w_\alpha^q \,\|\, w\big) + D_1\big(w_\alpha^q \,\|\, q\big) \quad \forall \alpha \in (0, \chi_{w,q}).$$

Л е м м а 11. *Для любых* $w, q \in \mathcal{P}(\mathcal{Y})$, *таких что* $\chi_{w,q} > 0$, *где* $\chi_{w,q} \triangleq \sup\{\alpha : D_\alpha(w \,\|\, q) < \infty\}$, *функции* $D_\alpha(w \,\|\, q)$, $D_1\big(w_\alpha^q \,\|\, w\big)$ *и* $D_1\big(w_\alpha^q \,\|\, q\big)$ *аналитичны по* α *на* $(0, \chi_{w,q})$. *Кроме того,*

$$\left.\frac{\partial^\kappa D_\alpha(w \,\|\, q)}{\partial \alpha^\kappa}\right|_{\alpha=\phi} = \begin{cases} \kappa! \displaystyle\sum_{t=0}^{\kappa} \frac{(-1)^{\kappa-t}}{(\phi-1)^{\kappa-t+1}} G_{w,q}^t(\phi), & \phi \neq 1, \\ \kappa!\, G_{w,q}^{\kappa+1}(1), & \phi = 1, \end{cases} \tag{12}$$

где $G_{w,q}^t(\phi)$ *определяется через множество* $\mathcal{J}_t$ *следующим образом*:

$$\mathcal{J}_t \triangleq \{(j_1, j_2, \ldots, j_t) : \ j_i \in \mathbb{Z}_{\geqslant 0} \ \forall i \ \text{и} \ 1j_1 + 2j_2 + \ldots + tj_t = t\}, \tag{13}$$

$$G_{w,q}^t(\phi) \triangleq$$

$$\triangleq \begin{cases} (\phi-1) D_\phi(w \,\|\, q), & t = 0, \\ \displaystyle\sum_{\mathcal{J}_t} \frac{-(j_1+j_2+\ldots+j_t-1)!}{j_1!\, j_2! \ldots j_t!} \prod_{i=1}^{t} \left(\frac{(-1)}{i!} \mathbf{E}_{w_\phi^q}\left[\left(\ln \frac{dw}{d\nu} - \ln \frac{dq}{d\nu} \right)^i \right] \right)^{j_i}, & t \in \mathbb{Z}_+. \end{cases} \tag{14}$$

Лемма 11, насколько нам известно, является новой; она доказывается в [22, Приложение A] с помощью стандартных результатов о непрерывности и дифференцируемости параметрических интегралов и формулы Фаа-ди-Бруно для производной композиции гладких функций.

Заметим, что $\mathcal{J}_1 = \{(1)\}$, $\mathcal{J}_2 = \{(2,0), (0,1)\}$ и $\mathcal{J}_3 = \{(3,0,0), (1,1,0), (0,0,1)\}$. Таким образом, с помощью (14) непосредственной подстановкой можно убедиться, что

$$G_{w,q}^1(\phi) = \mathbf{E}_{w_\phi^q}[\xi],$$

$$G_{w,q}^2(\phi) = \frac{1}{2}\big(\mathbf{E}_{w_\phi^q}[\xi^2] - \mathbf{E}_{w_\phi^q}[\xi]^2\big) = \frac{1}{2} \mathbf{E}_{w_\phi^q}\left[\big(\xi - \mathbf{E}_{w_\phi^q}[\xi]\big)^2 \right],$$

$$G_{w,q}^3(\phi) = \frac{1}{3!} \mathbf{E}_{w_\phi^q}[\xi^3] - \frac{1}{2} \mathbf{E}_{w_\phi^q}[\xi^2]\, \mathbf{E}_{w_\phi^q}[\xi] + \frac{1}{3} \mathbf{E}_{w_\phi^q}[\xi]^3 = \frac{1}{3!} \mathbf{E}_{w_\phi^q}\left[\big(\xi - \mathbf{E}_{w_\phi^q}[\xi]\big)^3 \right],$$

где $\xi = \ln \frac{dw}{d\nu} - \ln \frac{dq}{d\nu}$. Если подставить эти выражения в (12) и использовать тождество $\xi = \frac{1}{\phi - 1}\Big(\ln \frac{dw_\phi^q}{dw} + G_{w,q}^0(\phi)\Big)$, которое справедливо w_ϕ^q-почти наверное для $\phi \in (0, \chi_{w,q}) \setminus \{1\}$, получим следующие более краткие выражения для первых двух производных $D_\alpha(w \,\|\, q)$ по α:

$$
\frac{\partial}{\partial \alpha} D_\alpha(w \,\|\, q)\Big|_{\alpha=\phi} = \begin{cases} \dfrac{1}{(\phi-1)^2} D_1\big(w_\phi^q \,\|\, w\big), & \phi \neq 1, \\ \dfrac{1}{2}\, \mathbf{E}_w \left[\Big(\ln \dfrac{dw}{dq} - D_1(w \,\|\, q)\Big)^2\right], & \phi = 1, \end{cases}
$$

$$
\frac{\partial^2}{\partial \alpha^2} D_\alpha(w \,\|\, q)\Big|_{\alpha=\phi} =
$$

$$
= \begin{cases} \dfrac{1}{(\phi-1)^3} \left(\mathbf{E}_{w_\phi^q}\left[\left(\ln \dfrac{dw_\phi^q}{dw}\right)^2\right] - 2D_1\big(w_\phi^q \,\|\, w\big) - \big[D_1\big(w_\phi^q \,\|\, w\big)\big]^2\right), & \phi \neq 1, \\ \dfrac{1}{3}\, \mathbf{E}_w \left[\Big(\ln \dfrac{dw}{dq} - D_1(w \,\|\, q)\Big)^3\right], & \phi = 1. \end{cases}
$$

Из аналитичности $D_\alpha(w \,\|\, q)$ на $(0, \chi_{w,q})$ следует, что для любого $\phi \in (0, \chi_{w,q})$ существует открытый интервал, содержащий ϕ, на котором $D_\alpha(w \,\|\, q)$ равна степенному ряду, задаваемому производными $D_\alpha(w \,\|\, q)$ в точке $\alpha = \phi$. Если имеется конечный набор пар вероятностных мер $\{(w_i, q_i)\}_{i \in \mathcal{I}}$, то для любого ϕ, принадлежащего $(0, \chi_{w_i,q_i})$ для всех $i \in \mathcal{I}$, существует открытый интервал, содержащий ϕ, на котором каждая $D_\alpha(w_i \,\|\, q_i)$ равна степенному ряду, задаваемому производными $D_\alpha(w_i \,\|\, q_i)$ в точке $\alpha = \phi$. Если набор пар вероятностных мер бесконечен, то открытого интервала, содержащего ϕ и принадлежащего всем $(0, \chi_{w_i,q_i})$, может не быть. Следующая лемма 12 утверждает существование такого интервала, когда $D_\beta(w_i \,\|\, q_i)$ равномерно ограничена при $\beta > \phi$ для всех $i \in \mathcal{I}$. Кроме того, лемма 12 дает равномерную по всем $i \in \mathcal{I}$ оценку остаточного члена такого степенного ряда на этом интервале.

Лемма 12. *Для любых $\gamma, \phi, \beta \in \mathbb{R}_+$, таких что $\phi \in (0, \beta)$, и любых $w, q \in \mathcal{P}(\mathcal{Y})$, таких что $D_\beta(w \,\|\, q) \leqslant \gamma$,*

$$
\left| \frac{\partial^\kappa D_\alpha(w \,\|\, q)}{\partial \alpha^\kappa}\Big|_{\alpha=\phi} \right| \leqslant \begin{cases} \kappa!\, \tau^{\kappa+1}\kappa, & \phi \neq 1, \\ \kappa!\, \tau^{\kappa+1}, & \phi = 1, \end{cases}
$$

$$
\left| D_\eta(w \,\|\, q) - \sum_{i=0}^{\kappa-1} \frac{(\eta-\phi)^i}{i!} \, \frac{\partial^i D_\alpha(w \,\|\, q)}{\partial \alpha^i}\Big|_{\alpha=\phi} \right| \leqslant
$$

$$
\leqslant \begin{cases} \dfrac{\tau^{\kappa+1}|\eta-\phi|^\kappa}{1-|\eta-\phi|\tau} \left[\kappa - 1 + \dfrac{1}{1-|\eta-\phi|\tau}\right], & \phi \neq 1, \\ \dfrac{\tau^{\kappa+1}|\eta-\phi|^\kappa}{1-|\eta-\phi|\tau}, & \phi = 1, \end{cases} \qquad \forall \eta : \ |\eta - \phi| \leqslant \frac{1}{\tau},
$$

где

$$
\tau \triangleq \begin{cases} \dfrac{1}{|\phi-1|} \vee \left[\dfrac{1+e^{(1\vee\beta)\gamma}}{\phi \wedge (\beta-\phi)} + \gamma\right], & \phi \neq 1, \\ \dfrac{1+e^{\beta\gamma}}{1 \wedge (\beta-1)}, & \phi = 1. \end{cases}
$$

Лемма 12, насколько нам известно, является новой; доказательство, использующее формулу (12) и простейшие свойства вещественных аналитических функций и степенных рядов, приведено в [22, Приложение A].

2.3. Условное расхождение Реньи и скошенный канал. Понятия условного расхождения Реньи и скошенного канала позволяют записывать многие часто используемые выражения в более кратком виде.

Определение 3. Для любых $\alpha \in \mathbb{R}_+$, $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$, $Q\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ и $p \in \mathcal{P}(\mathcal{X})$ *условным расхождением Реньи порядка α для распределения на входе p* называется

$$D_\alpha(W \,\|\, Q \,|\, p) \triangleq \sum_{x \in \mathcal{X}} p(x) D_\alpha(W(x) \,\|\, Q(x)). \tag{15}$$

Если $\exists q \in \mathcal{P}(\mathcal{Y})$, такая что $Q(x) = q$ для всех $x \in \mathcal{X}$, то $D_\alpha(W \,\|\, Q \,|\, p)$ будем обозначать через $D_\alpha(W \,\|\, q \,|\, p)$.

Замечание 1. В работах [11,32] через $D_\alpha(W \,\|\, Q \,|\, p)$ обозначено $D_\alpha(p \circledast W \,\|\, p \circledast Q)$. При $\alpha = 1$ это равносильно нашим соглашениям, но в случае $\alpha \neq 1$ это уже не так. Если либо $\alpha = 1$, либо $D_\alpha(W(x) \,\|\, Q(x))$ принимает одно и то же значение для всех x с положительной $p(x)$, то $D_\alpha(p \circledast W \,\|\, p \circledast Q) = \sum\limits_x p(x) D_\alpha(W(x) \,\|\, Q(x))$, в противном случае $D_\alpha(p \circledast W \,\|\, p \circledast Q) < \sum\limits_x p(x) D_\alpha(W(x) \,\|\, Q(x))$ при $\alpha \in (0,1)$ и $D_\alpha(p \circledast W \,\|\, p \circledast Q) > \sum\limits_x p(x) D_\alpha(W(x) \,\|\, Q(x))$ при $\alpha \in (1,\infty)$. Эти неравенства следуют из неравенства Йенсена и строгой вогнутости функции натурального логарифма.

Определение 4. Для любых $\alpha \in \mathbb{R}_+$, $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ и $Q\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ *скошенным каналом W_α^Q порядка α* называется функция из $\{x : D_\alpha(W(x) \,\|\, Q(x)) < \infty\}$ в $\mathcal{P}(\mathcal{Y})$ следующего вида:

$$\frac{dW_\alpha^Q(x)}{d\nu} \triangleq e^{(1-\alpha)D_\alpha(W(x) \,\|\, Q(x))} \left[\frac{dW(x)}{d\nu}\right]^\alpha \left[\frac{dQ(x)}{d\nu}\right]^{1-\alpha}. \tag{16}$$

Если $\exists q \in \mathcal{P}(\mathcal{Y})$, такая что $Q(x) = q$ для всех $x \in \mathcal{X}$, то канал W_α^Q будем обозначать через W_α^q.

§ 3. Информация Августина

Основная цель этого параграфа – ввести понятия информации Августина и августиновского среднего. В п. 3.1 определяется информация Августина порядка α для распределения на входе p и устанавливается существование и единственность августиновского среднего для любого распределения на входе p и любого положительного конечного порядка α. Затем информация Августина анализируется сперва как функция распределения на входе при заданном порядке (п. 3.2), а затем как функция порядка при заданном распределении на входе (п. 3.3). В заключение в п. 3.4 мы сравниваем информацию Августина с информацией Реньи, характеризуя каждую из этих двух величин через другую. Некоторые из важнейших свойств информации Августина и августиновского среднего были впервые описаны Августином в [6, § 34] для порядков не выше единицы, и именно поэтому мы предлагаем такие названия для этих величин. Доказательства лемм из этого параграфа можно найти в [22, Приложение B].

3.1. Существование и единственность августиновского среднего.

Определение 5. Для любых $\alpha \in \mathbb{R}_+$, $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ и $p \in \mathcal{P}(\mathcal{X})$ *информацией Августина порядка α для распределения на входе p* называется

$$I_\alpha(p; W) \triangleq \inf_{q \in \mathcal{P}(\mathcal{Y})} D_\alpha(W \,\|\, q \,|\, p). \tag{17}$$

Непосредственной подстановкой нетрудно убедиться, что

$$D_1(W \,\|\, q \,|\, p) = D_1(W \,\|\, q_{1,p} \,|\, p) + D_1(q_{1,p} \,\|\, q) \quad \forall q \in \mathcal{P}(\mathcal{Y}),$$

где

$$q_{1,p} \triangleq \sum_x p(x) W(x). \tag{18}$$

Тогда из леммы 2 и формулы (17) вытекает

$$I_1(p; W) = D_1(W \,\|\, q_{1,p} \,|\, p).$$

Таким образом, для информации Августина порядка 1 имеется выражение в замкнутом виде, равное взаимной информации. Однако для других порядков информация Августина не выражается в замкнутом виде. Тем не менее приведенная ниже лемма 13 устанавливает существование и единственность вероятностной меры $q_{\alpha,p}$, такой что $I_\alpha(p; W) = D_\alpha(W \,\|\, q_{\alpha,p} \,|\, p)$ для[10] любого положительного порядка α и любого распределения на входе p. Кроме того, утверждения (c) и (d) леммы 13 дают альтернативную характеризацию $q_{\alpha,p}$ как единственной неподвижной точки оператора $\mathrm{T}_{\alpha,p}(\cdot)$, удовлетворяющей $q_{1,p} \prec q_{\alpha,p}$. Лемма 13(e) дает альтернативную характеризацию информации Августина для порядков, отличных от единицы.[11]

О п р е д е л е н и е 6. Пусть α – положительное вещественное число, и пусть W – канал вида $W \colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$.

- Для любого $p \in \mathcal{M}^+(\mathcal{X})$ *средняя мера порядка α для распределения на входе p* определяется как

  $$\frac{d\mu_{\alpha,p}}{d\nu} \triangleq \left[\sum_x p(x) \left(\frac{dW(x)}{d\nu} \right)^{\alpha} \right]^{\frac{1}{\alpha}},$$

  где ν – любая мера, такая что $\left(\sum_x p(x) W(x) \right) \prec \nu$;

- Для любого $p \in \mathcal{P}(\mathcal{X})$ *среднее Реньи порядка α для распределения на входе p* определяется как

  $$q^{\mathrm{g}}_{\alpha,p} \triangleq \frac{\mu_{\alpha,p}}{\|\mu_{\alpha,p}\|}; \tag{19}$$

- Для любого $p \in \mathcal{P}(\mathcal{X})$ *оператор Агустина* $\mathrm{T}_{\alpha,p}(\cdot) \colon \mathcal{Q}_{\alpha,p} \to \mathcal{P}(\mathcal{Y})$ *порядка α для распределения на входе p* определяется как

  $$\mathrm{T}_{\alpha,p}(q) \triangleq \sum_x p(x) W^q_\alpha(x) \quad \forall q \in \mathcal{Q}_{\alpha,p}, \tag{20}$$

  где $\mathcal{Q}_{\alpha,p} \triangleq \{q \in \mathcal{P}(\mathcal{Y}) : \ D_\alpha(W \,\|\, q \,|\, p) < \infty\}$, а скошенный канал W^q_α определен в (16). Кроме того, $\mathrm{T}^0_{\alpha,p}(q) = q$ и $\mathrm{T}^{i+1}_{\alpha,p}(q) \triangleq \mathrm{T}_{\alpha,p}\big(\mathrm{T}^i_{\alpha,p}(q)\big)$ для любого неотрицательного целого i.

[10] Это довольно легко доказать, когда $\mathcal{Y}$ – конечное множество. Единственность $q_{\alpha,p}$ следует из строгой выпуклости расхождения Реньи по второму аргументу, описанной в лемме 5. Если $\mathcal{Y}$ конечно, то $\mathcal{P}(\mathcal{Y})$ компактно и существование $q_{\alpha,p}$ следует из полунепрерывности снизу расхождения Реньи по второму аргументу, вытекающей из леммы 3, и теоремы об экстремуме для полунепрерывных снизу функций [33, гл. 3, п. 12.2]. Однако для каналов с произвольными выходными пространствами $\mathcal{P}(\mathcal{Y})$ не компактно, поэтому теорему об экстремуме для доказательства существования $q_{\alpha,p}$ применить нельзя.

[11] Эта альтернативная характеризация используется для доказательства равносильности двух определений экспоненты сферической упаковки и экспоненты сильного обращения теоремы кодирования.

Лемма 13. *Пусть W – канал вида $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$, а $p \in \mathcal{P}(\mathcal{X})$ – распределение на входе. Тогда*

(a) $I_\alpha(p;W) \leqslant D_\alpha(W \,\|\, q_{1,p} \,|\, p) \leqslant \hbar(p) < \infty$ *для всех* $\alpha \in \mathbb{R}_+$, *где* $q_{1,p}$ *определена в* (18);

(b) $I_1(p;W) = D_1(W \,\|\, q_{1,p} \,|\, p)$. *Кроме того,*

$$D_1(W \,\|\, q \,|\, p) - I_1(p;W) = D_1(q_{1,p} \,\|\, q) \quad \forall q \in \mathcal{P}(\mathcal{Y}); \tag{21}$$

(c) *Если* $\alpha \in (0,1)$, *то* $\exists!\, q_{\alpha,p}$, *такая что* $I_\alpha(p;W) = D_\alpha(W \,\|\, q_{\alpha,p} \,|\, p)$. *При этом*

$$\mathrm{T}_{\alpha,p}(q_{\alpha,p}) = q_{\alpha,p},$$
$$\lim_{j\to\infty} \|q_{\alpha,p} - \mathrm{T}^j_{\alpha,p}(q^{\mathrm{g}}_{\alpha,p})\| = 0, \tag{22}$$
$$D_1(q_{\alpha,p} \,\|\, q) \geqslant D_\alpha(W \,\|\, q \,|\, p) - I_\alpha(p;W) \geqslant D_\alpha(q_{\alpha,p} \,\|\, q) \quad \forall q \in \mathcal{P}(\mathcal{Y}) \tag{23}$$

и $q_{\alpha,p} \sim q_{1,p}$. *Кроме того,*[12] *если* $q_{1,p} \prec q$ *и* $\mathrm{T}_{\alpha,p}(q) = q$, *то* $q_{\alpha,p} = q$;

(d) *Если* $\alpha \in (1,\infty)$, *то* $\exists!\, q_{\alpha,p}$, *такая что* $I_\alpha(p;W) = D_\alpha(W \,\|\, q_{\alpha,p} \,|\, p)$. *При этом*

$$\mathrm{T}_{\alpha,p}(q_{\alpha,p}) = q_{\alpha,p},$$
$$D_\alpha(q_{\alpha,p} \,\|\, q) \geqslant D_\alpha(W \,\|\, q \,|\, p) - I_\alpha(p;W) \geqslant D_1(q_{\alpha,p} \,\|\, q) \quad \forall q \in \mathcal{P}(\mathcal{Y}) \tag{24}$$

и $q_{\alpha,p} \sim q_{1,p}$. *Кроме того, если* $\mathrm{T}_{\alpha,p}(q) = q$, *то* $q_{\alpha,p} = q$;

(e) *Если* $\alpha \in \mathbb{R}_+ \setminus \{1\}$, *то*

$$\begin{aligned}
I_\alpha(p;W) &= \frac{\alpha}{1-\alpha} D_1\big(W_\alpha^{q_{\alpha,p}} \,\|\, W \,|\, p\big) + I_1\big(p; W_\alpha^{q_{\alpha,p}}\big) = \\
&= \begin{cases} \inf\limits_{V \in \mathcal{P}(\mathcal{Y}|\mathcal{X})} \dfrac{\alpha}{1-\alpha} D_1(V \,\|\, W \,|\, p) + I_1(p;V), & \alpha \in (0,1), \\ \sup\limits_{V \in \mathcal{P}(\mathcal{Y}|\mathcal{X})} \dfrac{\alpha}{1-\alpha} D_1(V \,\|\, W \,|\, p) + I_1(p;V), & \alpha \in (1,\infty), \end{cases} = \\
&= \frac{\alpha}{1-\alpha} \inf_{V \in \mathcal{P}(\mathcal{Y}|\mathcal{X})} \left(D_1(V \,\|\, W \,|\, p) + \frac{1-\alpha}{\alpha} I_1(p;V) \right).
\end{aligned} \tag{25}$$

Сходимость, описанная в (22), имеет место не только для среднего Реньи $q^{\mathrm{g}}_{\alpha,p}$, но также и для некоторых других вероятностных мер. В [22, Приложение B, замечание 6] описано, как можно доказать следующий, более общий результат о сходимости для любых $\alpha \in (0,1)$ и $p \in \mathcal{P}(\mathcal{X})$:

$$\lim_{j\to\infty} \big\| q_{\alpha,p} - \mathrm{T}^j_{\alpha,p}(q) \big\| = 0, \quad \text{если} \quad q \sim q_{1,p} \quad \text{и} \quad \operatorname{ess\,sup}_{q_{1,p}} \left| \ln \frac{dq}{dq_{1,p}} \right| < \infty. \tag{26}$$

Утверждение (a) доказывается с помощью леммы 1; неравенство $I_\alpha(p;W) \leqslant \hbar(p)$ было доказано Чисаром в [2, формула (24)] с помощью других рассуждений. Хорошо известное утверждение (b) доказывается прямой подстановкой. Утверждение (c) принадлежит[13] Августину [6, лемма 34.2]. Утверждение (d), насколько нам известно, является новым. Утверждение (e) для случая конечного $\mathcal{Y}$ было доказано Чисаром [2, формулы (A24), (A27)].

[12] Заметим, что из одного равенства $\mathrm{T}_{\alpha,p}(q) = q$ на следует, что $q_{\alpha,p} = q$ для всех $\alpha \in (0,1)$. Рассмотрим, например, двоичный симметричный канал, и пусть q – вероятностная мера, сосредоточенная на одном из выходных символов. Тогда $\mathrm{T}_{\alpha,p}(q) = q$, но $q_{\alpha,p} \neq q$ при всех $p \in \mathcal{P}(\mathcal{X})$ и $\alpha \in (0,1)$.

[13] Точнее говоря, лемма 34.2 из [6] не содержит неравенства $D_1(q_{\alpha,p} \,\|\, q) \geqslant D_\alpha(W \,\|\, q \,|\, p) - I_\alpha(p;W)$, а равенство (22) в ней сформулировано для $q_{1,p}$, а не для $q^{\mathrm{g}}_{\alpha,p}$. Мы не смогли проверить правильность доказательства Августина этой леммы в [6], подробнее об этом см. в [22, Приложение C].

Определение 7. Для любых $\alpha \in \mathbb{R}_+$, $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ и $p \in \mathcal{P}(\mathcal{X})$ единственная вероятностная мера $q_{\alpha,p}$ на $(\mathcal{Y}, \mathcal{Y})$, для которой $I_\alpha(p; W) = D_\alpha(W \,\|\, q_{\alpha,p} \,|\, p)$, называется *августиновским средним порядка α для распределения на входе p.*

Из леммы 2 и соотношений (21), (23), (24) вытекает следующая граница, аналогичная результату Чисара [34, теорема 3.1]:

$$\sqrt{2\frac{D_\alpha(W \,\|\, q \,|\, p) - I_\alpha(p; W)}{\alpha \wedge 1}} \geqslant \|q_{\alpha,p} - q\| \quad \forall q \in \mathcal{P}(\mathcal{Y}) \quad \forall \alpha \in \mathbb{R}_+.$$

Информация Августина и августиновское среднее имеют выражения в замкнутом виде лишь для $\alpha = 1$; для других порядков таких выражений нет. Однако из свойства неподвижности точки $\mathrm{T}_{\alpha,p}(q_{\alpha,p}) = q_{\alpha,p}$, установленного в лемме 13(c), (d), и из определения оператора $\mathrm{T}_{\alpha,p}(\cdot)$ в (20) вытекает следующее тождество для августиновского среднего:

$$\frac{dq_{\alpha,p}}{d\nu} = \left[\sum_x p(x) \left(\frac{dW(x)}{d\nu}\right)^\alpha e^{(1-\alpha) D_\alpha(W(x) \,\|\, q_{\alpha,p})}\right]^{\frac{1}{\alpha}} \qquad \forall \nu:\ q_{1,p} \prec \nu. \tag{27}$$

В п. 3.3 при анализе $I_\alpha(p; W)$ и $q_{\alpha,p}$ как функций от α мы используем это тождество вместо выражения в замкнутом виде.

Лемма 14. *Для любого канала-произведения $W_{[1,n]}\colon \mathcal{X}_1^n \to \mathcal{P}(\mathcal{Y}_1^n)$ длины n и любого распределения на входе $p \in \mathcal{P}(\mathcal{X}_1^n)$ неравенство*

$$I_\alpha\bigl(p; W_{[1,n]}\bigr) \leqslant \sum_{t=1}^n I_\alpha(p_t; W_t) \tag{28}$$

справедливо для всех $\alpha \in \mathbb{R}_+$, где $p_t \in \mathcal{P}(\mathcal{X}_t)$ – маргинальное распределение для p на $\mathcal{X}_t$. Кроме того, неравенство в (28) обращается в равенство для некоторого $\alpha \in \mathbb{R}_+$ тогда и только тогда, когда $q_{\alpha,p}$ имеет вид

$$q_{\alpha,p} = \bigotimes_{t=1}^n q_{\alpha,p_t}. \tag{29}$$

Если $p = \bigotimes_{t=1}^n p_t$, то (29) имеет место для всех $\alpha \in \mathbb{R}_+$, и следовательно, (28) обращается в равенство для всех $\alpha \in \mathbb{R}_+$.

3.2. Информация Августина как функция входного распределения. Информация Августина порядка α для распределения на входе p определяется как точная нижняя грань семейства условных расхождений Реньи, являющихся линейными по p. Тогда информация Августина вогнута по p как поточечный инфимум семейства вогнутых функций. В лемме 15 этот результат уточняется с использованием леммы 13.

Лемма 15. *Для любых $\alpha \in \mathbb{R}_+$ и $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ функция $I_\alpha(p; W)$ вогнута по p, удовлетворяя неравенствам*

$$\begin{aligned} &I_\alpha(p_\beta; W) \geqslant \beta I_\alpha(p_1; W) + (1-\beta) I_\alpha(p_0; W) + \beta D_{\alpha\wedge 1}\bigl(q_{\alpha,p_1} \,\|\, q_{\alpha,p_\beta}\bigr) + \\ &\quad + (1-\beta) D_{\alpha\wedge 1}\bigl(q_{\alpha,p_1} \,\|\, q_{\alpha,p_\beta}\bigr), \end{aligned} \tag{30}$$

$$\begin{aligned} &I_\alpha(p_\beta; W) \leqslant \beta I_\alpha(p_1; W) + (1-\beta) I_\alpha(p_0; W) + \beta D_{\alpha\vee 1}\bigl(q_{\alpha,p_1} \,\|\, q_{\alpha,p_\beta}\bigr) + \\ &\quad + (1-\beta) D_{\alpha\vee 1}\bigl(q_{\alpha,p_0} \,\|\, q_{\alpha,p_\beta}\bigr), \end{aligned} \tag{31}$$

$$\begin{aligned} &I_\alpha(p_\beta; W) \leqslant \beta I_\alpha(p_1; W) + (1-\beta) I_\alpha(p_0; W) + \hbar(\beta) - \\ &\quad - D_{\alpha\wedge 1}\bigl(q_{\alpha,p_\beta} \,\|\, \beta q_{\alpha,p_1} + (1-\beta) q_{\alpha,p_0}\bigr), \end{aligned} \tag{32}$$

где $p_\beta = \beta p_1 + (1-\beta)p_0$ *для любых* $p_0, p_1 \in \mathcal{P}(\mathcal{X})$ *и* $\beta \in [0,1]$.

Из леммы 15 следует, что для любого положительного порядка α и любого канала W информация Августина $I_\alpha(p;W)$ порядка α непрерывна по распределению на входе p тогда и только тогда, когда $\sup\limits_{p\in\mathcal{P}(\mathcal{X})} I_\alpha(p;W)$ конечен.[14] Кроме того, если $\sup\limits_{p\in\mathcal{P}(\mathcal{X})} I_\eta(p;W)$ конечен для некоторого $\eta \in \mathbb{R}_+$, то семейство $\{I_\alpha(p;W)\}_{\alpha\in(0,\eta]}$ равномерно равностепенно непрерывно по p на $\mathcal{P}(\mathcal{X})$.

Чтобы увидеть, почему конечность $\sup\limits_{p\in\mathcal{P}(\mathcal{X})} I_\alpha(p;W)$ необходима для непрерывности, заметим, что из неотрицательности расхождения Реньи для вероятностных мер и неравенства (30) следует, что

$$I_\alpha(p_\beta;W) - I_\alpha(p_0;W) \geqslant \beta(I_\alpha(p_1;W) - I_\alpha(p_0;W)) + \beta D_{\alpha\wedge 1}\big(q_{\alpha,p_1} \,\|\, q_{\alpha,p_\beta}\big) + \\ + (1-\beta)D_{\alpha\wedge 1}\big(q_{\alpha,p_1} \,\|\, q_{\alpha,p_\beta}\big) \geqslant \beta(I_\alpha(p_1;W) - I_\alpha(p_0;W)).$$

С другой стороны, $\|p_\beta - p_0\| \leqslant 2\beta$. Таким образом, если существует последовательность $\{p_i\}_{i\in\mathbb{Z}_+} \subset \mathcal{P}(\mathcal{X})$, такая что $\lim\limits_{i\uparrow\infty} I_\alpha(p_i;W) = \infty$, то $I_\alpha(p;W)$ разрывна в любой точке $p \in \mathcal{P}(\mathcal{X})$.

Обратное утверждение, т.е. достаточность, можно установить вместе с равностепенной непрерывностью. Для любых $p_1, p_2 \in \mathcal{P}(\mathcal{X})$, таких что $p_1 \neq p_2$, определим $s_\wedge$, s_1 и s_0 как

$$s_\wedge = \frac{p_1 \wedge p_0}{\|p_1 \wedge p_0\|}, \quad s_1 = \frac{p_1 - p_1 \wedge p_0}{1 - \|p_1 \wedge p_0\|}, \quad s_0 = \frac{p_0 - p_1 \wedge p_0}{1 - \|p_1 \wedge p_0\|}.$$

Тогда $s_\wedge, s_1, s_0 \in \mathcal{P}(\mathcal{X})$ и $s_1 \perp s_0$. С другой стороны, $\|p_1 - p_0\| = 2 - 2\|p_1 \wedge p_0\|$. Поэтому

$$p_1 = \left(\frac{2 - \|p_1 - p_0\|}{2}\right) s_\wedge + \frac{\|p_1 - p_0\|}{2} s_1,$$
$$p_0 = \left(\frac{2 - \|p_1 - p_0\|}{2}\right) s_\wedge + \frac{\|p_1 - p_0\|}{2} s_0.$$

Следовательно, по леммам 2, 15 получаем

$$I_\alpha(p_0;W) - I_\alpha(p_1;W) \leqslant \hbar\left(\frac{\|p_1 - p_0\|}{2}\right) + \frac{\|p_1 - p_0\|}{2}\big(I_\alpha(s_0;W) - I_\alpha(s_1;W)\big) \leqslant \\ \leqslant \hbar\left(\frac{\|p_1 - p_0\|}{2}\right) + \frac{\|p_1 - p_0\|}{2} I_\alpha(s_0;W) \quad \forall p_1, p_0 \in \mathcal{P}(\mathcal{X}),\ \alpha \in \mathbb{R}_+.$$

Таким образом,

$$|I_\alpha(p_0;W) - I_\alpha(p_1;W)| \leqslant \hbar\left(\frac{\|p_1 - p_0\|}{2}\right) + \frac{\|p_1 - p_0\|}{2} \sup_{p\in\mathcal{P}(\mathcal{X})} I_\eta(p;W) \\ \forall p_1, p_0 \in \mathcal{P}(\mathcal{X}),\ \alpha \in (0,\eta].$$

3.3. Информация Августина как функция порядка. Основная цель этого пункта – охарактеризовать поведение информации Августина как функции порядка при

[14] Для информации Реньи, обсуждаемой в п. 3.4, аналогичные соотношения уже доказаны в [13, лемма 16(d), (e)]. Единственной существенной тонкостью является то, что для порядков из $(0,1)$ информация Реньи непрерывна по p, даже если выражение для соответствующей пропускной способности бесконечно, так как информация Реньи квазивогнута, а не вогнута по p для порядков из $(0,1)$ (см. [13, лемма 6(a)]).

заданном распределении на входе. В лемме 16 даны предварительные результаты, облегчающие дальнейший анализ информации Августина как функции порядка; результаты этого анализа представлены в лемме 17.

Лемма 16. *Для любого канала W вида $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ и любого распределения на входе $p \in \mathcal{P}(\mathcal{X})$*

(a) $D_\alpha(W(x) \,\|\, q_{\alpha,p}) \leqslant \ln \dfrac{1}{p(x)}$;

(b) $[p(x)]^{\frac{1}{\alpha\wedge 1}} W(x) \leqslant q_{\alpha,p}$;

(c) $\left|\ln \dfrac{dq_{\alpha,p}}{dq_{1,p}}\right| \leqslant \dfrac{|\alpha-1|}{\alpha} \ln \dfrac{1}{\min\limits_{x\colon p(x)>0} p(x)}$.

Границы, указанные в лемме 16, получаются из (27) элементарными преобразованиями.

Лемма 17. *Для любого канала $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ и любого распределения на входе $p \in \mathcal{P}(\mathcal{X})$ справедливы следующие утверждения*:

(a) *Либо $(\alpha-1)I_\alpha(p;W)$ – строго выпуклая по α функция из $\mathbb{R}_+$ в $[-\hbar(p),\infty)$, либо $I_\alpha(p;W) = \sum\limits_x p(x)\ln\gamma(x)$ для некоторой $\gamma\colon \mathcal{X} \to [1,\infty)$, такой что $\dfrac{dW(x)}{dq_{1,p}} = \gamma(x)$ $W(x)$-п.н. для всех $x \in \operatorname{supp} p$, и $q_{\alpha,p} = q_{1,p}$ для всех $\alpha \in \mathbb{R}_+$;*

(b) *$\dfrac{1-\alpha}{\alpha} I_\alpha(p;W)$ – невозрастающая непрерывная по α функция из $\mathbb{R}_+$ в $\mathbb{R}$;*

(c) *$I_\alpha(p;W)$ – неубывающая непрерывная по α функция из $\mathbb{R}_+$ в $[0,\hbar(p)]$;*

(d) *$\left\{\ln \dfrac{dq_{\alpha,p}}{dq_{1,p}}\right\}_{y\in\mathcal{Y}}$ – равностепенно непрерывное по α семейство функций на $\mathbb{R}_+$;*

(e) *$I_\alpha(p;W)$ – непрерывно дифференцируемая по α функция из $\mathbb{R}_+$ в $[0,\hbar(p)]$, такая что*

$$\frac{\partial}{\partial\alpha} I_\alpha(p;W)\bigg|_{\alpha=\phi} = \frac{\partial}{\partial\alpha} D_\alpha(W \,\|\, q_{\phi,p} \,|\, p)\bigg|_{\alpha=\phi} =$$

$$= \begin{cases} \dfrac{1}{(\phi-1)^2} D_1\big(W_\phi^{q_{\phi,p}} \,\|\, W \,|\, p\big), & \phi \neq 1,\\ \sum\limits_x \dfrac{p(x)}{2}\, \mathbf{E}_{W(x)}\left[\left(\ln \dfrac{dW(x)}{dq_{1,p}} - D_1(W(x) \,\|\, q_{1,p})\right)^2\right], & \phi = 1; \end{cases}$$

(f) *Если функция $(\alpha-1)I_\alpha(p;W)$ строго выпукла по α, то функция $I_1\big(p;W_\alpha^{q_{\alpha,p}}\big)$, т.е. $D_1\big(W_\alpha^{q_{\alpha,p}} \,\|\, q_{\alpha,p} \,|\, p\big)$, – монотонно возрастающая и непрерывная по α на $\mathbb{R}_+$; в противном случае $I_1\big(p;W_\alpha^{q_{\alpha,p}}\big) = \sum\limits_x p(x)\ln\gamma(x)$ (т.е. $D_1\big(W_\alpha^{q_{\alpha,p}} \,\|\, q_{\alpha,p} \,|\, p\big) = \sum\limits_x p(x)\ln\gamma(x)$) для некоторой $\gamma\colon \mathcal{X} \to [1,\infty)$, такой что $\dfrac{dW(x)}{dq_{1,p}} = \gamma(x)$ $W(x)$-п.н. для всех $x \in \operatorname{supp} p$, и $q_{\alpha,p} = q_{1,p}$ для всех $\alpha \in \mathbb{R}_+$;*

(g) *$\lim\limits_{\alpha\downarrow 0} I_1\big(p;W_\alpha^{q_{\alpha,p}}\big) = \lim\limits_{\alpha\downarrow 0} I_\alpha(p;W)$.*

(Строгая) выпуклость функции $(\alpha-1)I_\alpha(p;W)$ по α на $\mathbb{R}_+$ равносильная (строгой) вогнутости функции $sI_{\frac{1}{1+s}}(p;W)$ по s на $(-1,\infty)$ – см. доказательство утверждения (f). Вогнутость функции $sI_{\frac{1}{1+s}}(p;W)$ по s на $(-1,\infty)$ и утверждения (b), (c) леммы 17 были ранее получены Августином в [6, лемма 34.3] для порядков от нуля до единицы. Утверждения (a) и (d)–(g) леммы 17, насколько нам известно, являются новыми. Лемма 17 в основном касается информации Августина как функции порядка при заданном распределении на входе. Утверждение (d), т.е. равностепенная непрерывность $\left\{\ln \dfrac{dq_{\alpha,p}}{dq_{1,p}}\right\}_{y\in\mathcal{Y}}$ как семейства функций порядка α, выведено как необходимый инструмент для доказательства непрерывности производной информации Августина, т.е. утверждения (e). Отметим, что в лемме 16(c) уже установлена эта равностепенная непрерывность при $\alpha = 1$.

3.4. Сравнение информаций Августина и Реньи. Информация Августина – не единственная информационная величина, определяемая через расхождение Реньи, существуют и другие. По-видимому, наиболее известной среди них благодаря своему операциональному смыслу, установленному Галлагером в [14], является информация Реньи, введенная вначале Галлагером[15] [14], а затем Сибсоном [35].

Определение 8. Для любых $\alpha \in \mathbb{R}_+$, $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ и $p \in \mathcal{P}(\mathcal{X})$ *информацией Реньи порядка α для распределения на входе p* называется

$$I_\alpha^{\mathrm{g}}(p;W) \triangleq \inf_{q\in\mathcal{P}(\mathcal{Y})} D_\alpha(p \circledast W \,\|\, p\otimes q). \tag{33}$$

Как отмечено в [35], непосредственной проверкой можно убедиться, что

$$D_\alpha(p \circledast W \,\|\, p\otimes q) = \\ = D_\alpha\bigl(p \circledast W \,\|\, p\otimes q_{\alpha,p}^{\mathrm{g}}\bigr) + D_\alpha\bigl(q_{\alpha,p}^{\mathrm{g}} \,\|\, q\bigr) \quad \forall p \in \mathcal{P}(\mathcal{X}),\ q \in \mathcal{P}(\mathcal{Y}),\ \alpha \in \mathbb{R}_+,$$

где $q_{\alpha,p}^{\mathrm{g}}$ – среднее Реньи, определенное в (19). По лемме 2 отсюда получаем

$$I_\alpha^{\mathrm{g}}(p;W) = D_\alpha\bigl(p \circledast W \,\|\, p\otimes q_{\alpha,p}^{\mathrm{g}}\bigr) \quad \forall p \in \mathcal{P}(\mathcal{X}),\ \alpha \in \mathbb{R}_+, \tag{34}$$

$$D_\alpha(p \circledast W \,\|\, p\otimes q) = I_\alpha^{\mathrm{g}}(p;W) + D_\alpha\bigl(q_{\alpha,p}^{\mathrm{g}} \,\|\, q\bigr) \quad \forall p \in \mathcal{P}(\mathcal{X}),\ q \in \mathcal{P}(\mathcal{Y}),\ \alpha \in \mathbb{R}_+. \tag{35}$$

Для порядков, отличных от единицы, выражение в замкнутом виде, приведенное в (34), равно следующему, которое иногда берется за определение информации Реньи:

$$I_\alpha^{\mathrm{g}}(p;W) = \frac{\alpha}{\alpha-1}\ln\|\mu_{\alpha,p}\|, \quad \alpha \in \mathbb{R}_+ \setminus \{1\}.$$

Заметим, что в отличие от августиновского среднего порядка α среднее Реньи порядка α имеет выражение в замкнутом виде и для порядков, отличных от единицы. Кроме того, неравенства (21), (23) и (24) из леммы 13 заменяются равенством (35). Обсуждение информации Реньи, аналогичное вышеизложенному в этом параграфе для информации Августина, можно найти в [13].

Информация Реньи порядка 1 равна информации Августина порядка 1 для всех распределений на входе. Для других порядков это равенство не имеет места для произвольных распределений на входе. Однако информацию Августина и информацию Реньи можно характеризовать друг через друга, используя подходящие вариационные формы. Характеризация информации Августина в вариационном виде через информацию Реньи особенно полезна, так как информация Августина не имеет выражения в замкнутом виде, в то время как информация Реньи имеет. Из этой характеризации также вытекает другая вариационная характеризация информации Августина.

Лемма 18. *Пусть W – канал вида $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$, а $p \in \mathcal{P}(\mathcal{X})$ – распределение на входе.*

(a) *Пусть $u_{\alpha,p} \in \mathcal{P}(\mathcal{X})$ имеет вид $u_{\alpha,p}(x) = \dfrac{p(x)e^{(1-\alpha)D_\alpha(W(x)\,\|\,q_{\alpha,p})}}{\sum_{\widetilde{x}} p(\widetilde{x})e^{(1-\alpha)D_\alpha(W(\widetilde{x})\,\|\,q_{\alpha,p})}}$ для всех x; тогда*

$$I_\alpha(p;W) = I_\alpha^{\mathrm{g}}(u_{\alpha,p};W) + \frac{1}{\alpha-1}D_1(p\,\|\,u_{\alpha,p}) = \tag{36}$$

[15] Галлагер использует другую параметризацию и ограничивается случаем $\alpha \in (0,1)$.

$$= \begin{cases} \sup\limits_{u \in \mathcal{P}(\mathcal{X})} I_\alpha^{\mathrm{g}}(u; W) + \dfrac{1}{\alpha-1} D_1(p \,\|\, u), & \alpha \in (0,1), \\ \inf\limits_{u \in \mathcal{P}(\mathcal{X})} I_\alpha^{\mathrm{g}}(u; W) + \dfrac{1}{\alpha-1} D_1(p \,\|\, u), & \alpha \in (1,\infty); \end{cases} \tag{37}$$

(b) *Пусть* $a_{\alpha,p} \in \mathcal{P}(\mathcal{X})$ *имеет вид* $a_{\alpha,p}(x) = \dfrac{p(x) e^{(\alpha-1) D_\alpha\left(W(x) \,\|\, q^{\mathrm{g}}_{\alpha,p}\right)}}{\sum\limits_{\widetilde{x}} p(\widetilde{x}) e^{(\alpha-1) D_\alpha\left(W(\widetilde{x}) \,\|\, q^{\mathrm{g}}_{\alpha,p}\right)}}$ *для всех* x; *тогда*

$$I_\alpha^{\mathrm{g}}(p; W) = I_\alpha(a_{\alpha,p}; W) - \frac{1}{\alpha-1} D_1(a_{\alpha,p} \,\|\, p) = \tag{38}$$

$$= \begin{cases} \inf\limits_{a \in \mathcal{P}(\mathcal{X})} I_\alpha(a; W) - \dfrac{1}{\alpha-1} D_1(a \,\|\, p), & \alpha \in (0,1), \\ \sup\limits_{a \in \mathcal{P}(\mathcal{X})} I_\alpha(a; W) - \dfrac{1}{\alpha-1} D_1(a \,\|\, p), & \alpha \in (1,\infty); \end{cases} \tag{39}$$

(c) *Пусть* $f_{\alpha,p} \colon \mathcal{X} \to \mathbb{R}$ *имеет вид* $f_{\alpha,p}(x) = [D_\alpha(W(x) \,\|\, q_{\alpha,p}) - I_\alpha(p; W)] \mathbb{1}_{p(x)>0}$ *для всех* x; *тогда*

$$I_\alpha(p; W) = \frac{\alpha}{\alpha-1} \ln \mathbf{E}_\nu \left[\left(\sum_x p(x) e^{(1-\alpha) f_{\alpha,p}(x)} \left[\frac{dW(x)}{d\nu} \right]^\alpha \right)^{1/\alpha} \right] =$$

$$= \frac{\alpha}{\alpha-1} \ln \inf_{f \colon \mathbf{E}_p[f]=0} \mathbf{E}_\nu \left[\left(\sum_x p(x) e^{(1-\alpha) f(x)} \left[\frac{dW(x)}{d\nu} \right]^\alpha \right)^{1/\alpha} \right].$$

Лемма 18(a) была впервые доказана Полтыревым [19, теорема 3.4] в несколько другом виде для случая $\alpha \in [1/2, 1)$ в предположении, что $\mathcal{Y}$ конечно. Формула (39) леммы 18(b) была впервые получена в [10, теорема 1] для случая конечного $\mathcal{Y}$, однако в [10] не только не было получено выражение для оптимального $a_{\alpha,p}$, но и не утверждалось его существование. Лемма 18(c) была впервые доказана Августином [6, лемма 35.7] для порядков, меньших единицы.[16]

Следующие неравенства получаются как с помощью точки $u = p$ в вариационной характеризации леммы 18(a), так и с помощью точки $a = p$ в вариационной характеризации леммы 18(b):

$$I_\alpha(p; W) \geqslant I_\alpha^{\mathrm{g}}(p; W), \quad \alpha \in (0,1], \tag{40}$$

$$I_\alpha(p; W) \leqslant I_\alpha^{\mathrm{g}}(p; W), \quad \alpha \in [1,\infty). \tag{41}$$

Их также можно вывести, используя неравенство Йенсена и вогнутость функции натурального логарифма.

§ 4. Августиновская пропускная способность

В § 3 были определены и изучены информация Августина и августиновское среднее. В этом параграфе мы собираемся сделать это же для августиновской пропускной способности и августиновского центра. В п. 4.1 устанавливается существование и единственность августиновского центра для любых выпуклых множеств ограничений с конечной августиновской пропускной способностью и изучаются следствия существования августиновского центра для заданного порядка и множества ограничений. В п. 4.2 августиновские пропускная способность и центр анализируются как

[16] Утверждение [6, лемма 35.7(d)] вытекает из более сильных неравенств, получаемых с помощью соотношения (23) и леммы 18(c).

функции порядка для заданного множества ограничений. В п. 4.3 устанавливается граница на августиновскую пропускную способность выпуклой оболочки набора множеств ограничений на данном канале через августиновские пропускные способности индивидуальных множеств ограничений и определяются августиновские пропускные способности для произведений множеств ограничений на канале-произведении. Доказательства предложений из данного параграфа содержатся в [22, Приложение D].

В [6, §§ 33, 34] Августин провел анализ, аналогичный содержанию настоящего параграфа, и получил многие из ключевых результатов, такие как существование и единственность августиновского центра и его непрерывность как функции порядка (см. [6, леммы 34.6–34.8]), но для порядков не выше единицы. При этом Августин рассматривал пропускную способность и центр лишь для подмножеств $\mathcal{P}(\mathcal{X})$, определяемых через ограничения по стоимости. Этот важный специальный случай мы рассматриваем более подробно в § 5.

4.1. Существование и единственность августиновского центра.

Определение 9. Для любых $\alpha \in \mathbb{R}_+$, $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ и $\mathcal{A} \subset \mathcal{P}(\mathcal{X})$ *августиновской пропускной способностью порядка* α *канала* W *для множества ограничений* $\mathcal{A}$ называется

$$C_{\alpha,W,\mathcal{A}} \triangleq \sup_{p\in\mathcal{A}} I_\alpha(p;W).$$

Когда множество ограничений $\mathcal{A}$ совпадает со всем $\mathcal{P}(\mathcal{X})$, августиновская пропускная способность порядка α обозначается через $C_{\alpha,W}$, т.е. $C_{\alpha,W} \triangleq C_{\alpha,W,\mathcal{P}(\mathcal{X})}$.

Используя определение информации Августина $I_\alpha(p;W)$, данное в (17), получаем следующее выражение для $C_{\alpha,W,\mathcal{A}}$:

$$C_{\alpha,W,\mathcal{A}} = \sup_{p\in\mathcal{A}} \inf_{q\in\mathcal{P}(\mathcal{Y})} D_\alpha(W \,\|\, q \,|\, p). \tag{42}$$

Следующая теорема показывает, что по крайней мере для выпуклых $\mathcal{A}$ можно поменять порядок инфимума и супремума в данном выражении.

Теорема 1. *Для любых порядка* $\alpha \in \mathbb{R}_+$, *канала* W *вида* $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ *и выпуклого множества ограничений* $\mathcal{A} \subset \mathcal{P}(\mathcal{X})$ *имеет место равенство*

$$\sup_{p\in\mathcal{A}} \inf_{q\in\mathcal{P}(\mathcal{Y})} D_\alpha(W \,\|\, q \,|\, p) = \inf_{q\in\mathcal{P}(\mathcal{Y})} \sup_{p\in\mathcal{A}} D_\alpha(W \,\|\, q \,|\, p). \tag{43}$$

Если выражение в левой части (43) *конечно, т.е. если* $C_{\alpha,W,\mathcal{A}} \in \mathbb{R}_{\geqslant 0}$, *то* $\exists!\, q_{\alpha,W,\mathcal{A}} \in \mathcal{P}(\mathcal{Y})$, *называемая августиновским центром порядка* α *канала* W *для множества ограничений* $\mathcal{A}$, *такая что*

$$C_{\alpha,W,\mathcal{A}} = \sup_{p\in\mathcal{A}} D_\alpha(W \,\|\, q_{\alpha,W,\mathcal{A}} \,|\, p). \tag{44}$$

Кроме того, для любой последовательности распределений на входе $\{p^{(i)}\}_{i\in\mathbb{Z}_+} \subset \mathcal{A}$, *удовлетворяющей равенству* $\lim_{i\to\infty} I_\alpha(p^{(i)};W) = C_{\alpha,W,\mathcal{A}}$, *соответствующая последовательность* $\{q_{\alpha,p^{(i)}}\}_{i\in\mathbb{Z}_+}$ *августиновских средних порядка* α *является последовательностью Коши в метрике полной вариации на* $\mathcal{P}(\mathcal{Y})$, *а* $q_{\alpha,W,\mathcal{A}}$ *– единственной предельной точкой этой последовательности Коши.*

Наше доказательство теоремы 1 следует программе, предложенной в [12] для доказательства аналогичного результата в случае $\alpha = 1$ и $\mathcal{A} = \mathcal{P}(\mathcal{X})$. Вначале теорема 1 доказывается в предположении, что множество входов конечно, а затем этот результат обобщается на случай произвольных множеств входов. В случае, когда множество $\mathcal{X}$ конечно, можно также установить существование оптимального распределения на входе, для которого информация Августина равна августиновской пропускной способности.

Лемма 19. *Для любых порядка $\alpha \in \mathbb{R}_+$, канала W вида $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ с конечным множеством входов $\mathcal{X}$ и замкнутого выпуклого множества ограничений $\mathcal{A} \subset \mathcal{P}(\mathcal{X})$ существует $\widetilde{p} \in \mathcal{A}$, такое что $I_\alpha(\widetilde{p}; W) = C_{\alpha,W,\mathcal{A}}$, и $\exists!\, q_{\alpha,W,\mathcal{A}} \in \mathcal{P}(\mathcal{Y})$, такая что*

$$D_\alpha(W \,\|\, q_{\alpha,W,\mathcal{A}} \,|\, p) \leqslant C_{\alpha,W,\mathcal{A}} \quad \forall p \in \mathcal{A}.$$

Кроме того, $q_{\alpha,\widetilde{p}} = q_{\alpha,W,\mathcal{A}}$ для всех $\widetilde{p} \in \mathcal{A}$, таких что $I_\alpha(\widetilde{p}; W) = C_{\alpha,W,\mathcal{A}}$.

Если $\mathcal{A}$ совпадает с $\mathcal{P}(\mathcal{X})$, то выражение в правой части (44) равно радиусу Реньи $S_{\alpha,W}$, который мы сейчас определим. Тогда из теоремы 1 следует $C_{\alpha,W} = S_{\alpha,W}$.

Определение 10. Для любых $\alpha \in \mathbb{R}_+$ и $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ *радиусом Реньи порядка α канала W* называется

$$S_{\alpha,W} \triangleq \inf_{q \in \mathcal{P}(\mathcal{Y})} \sup_{x \in \mathcal{X}} D_\alpha(W(x) \,\|\, q).$$

Теорема 1 утверждает существование и единственность августиновского центра порядка α для выпуклых множеств ограничений с конечной августиновской пропускной способностью. Однако вероятностная мера $q_{\alpha,W,\mathcal{A}}$, удовлетворяющая (44), т.е. августиновский центр порядка α, может в принципе существовать и для невыпуклых множеств ограничений.

Определение 11. Множество ограничений $\mathcal{A}$ для канала $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ имеет августиновский центр порядка α тогда и только тогда, когда $\exists q \in \mathcal{P}(\mathcal{Y})$, такая что

$$\sup_{p \in \mathcal{A}} D_\alpha(W \,\|\, q \,|\, p) = C_{\alpha,W,\mathcal{A}}. \tag{45}$$

Если $C_{\alpha,W,\mathcal{A}}$ бесконечна, то все вероятностные меры на выходном пространстве удовлетворяют (45) в силу (42) и неравенства максимина-минимакса. Таким образом, для множеств ограничений с бесконечной августиновской пропускной способностью порядка α все вероятностные меры на выходном пространстве являются августиновскими центрами порядка α. С другой стороны, некоторые множества ограничений не имеют августиновского центра порядка α. Рассмотрим, например, p_1 и p_2, такие что $q_{\alpha,p_1} \neq q_{\alpha,p_2}$ и $I_\alpha(p_1; W) = I_\alpha(p_2; W)$. Тогда (45) не выполнено ни для какой вероятностной меры для $\mathcal{A} = \{p_1, p_2\}$, и $\mathcal{A}$ не имеет августиновского центра порядка α. Лемма 20 утверждает, что если августиновский центр для множества ограничений с конечной августиновской пропускной способностью существует, то этот августиновский центр единствен.

Лемма 20. *Пусть $\mathcal{A} \subset \mathcal{P}(\mathcal{X})$ – множество ограничений, такое что $C_{\alpha,W,\mathcal{A}} \in \mathbb{R}_{\geqslant 0}$, а $q_{\alpha,W,\mathcal{A}}$ – вероятностная мера, удовлетворяющая* (45)*. Тогда для любой последовательности $\{p^{(i)}\}_{i \in \mathbb{Z}_+} \subset \mathcal{A}$, такой что $\lim\limits_{i \to \infty} I_\alpha(p^{(i)}; W) = C_{\alpha,W,\mathcal{A}}$, последовательность $\{q_{\alpha,p^{(i)}}\}_{i \in \mathbb{Z}_+}$ августиновских средних порядка α является последовательностью Коши с предельной точкой $q_{\alpha,W,\mathcal{A}}$, и августиновский центр $q_{\alpha,W,\mathcal{A}}$ порядка α единствен.*

Для любого $\mathcal{A}$, имеющего августиновский центр порядка α и конечную $C_{\alpha,W,\mathcal{A}}$, из лемм 13(b)–(d) и 20 следует, что

$$C_{\alpha,W,\mathcal{A}} - I_\alpha(p; W) \geqslant D_{\alpha \wedge 1}(q_{\alpha,p} \,\|\, q_{\alpha,W,\mathcal{A}}) \quad \forall p \in \mathcal{A}.$$

Леммы 13(b)–(d) и 20 можно также использовать для вывода нижней границы на $\sup\limits_{p \in \mathcal{A}} D_\alpha(W \,\|\, q \,|\, p)$ через августиновские пропускную способность и центр.

Лемма 21. *Для любого множества ограничений $\mathcal{A}$, имеющего августиновский центр порядка α и конечную $C_{\alpha,W,\mathcal{A}}$, справедливо неравенство*

$$\sup_{p\in\mathcal{A}} D_\alpha(W \,\|\, q \,|\, p) \geqslant C_{\alpha,W,\mathcal{A}} + D_{\alpha\wedge 1}(q_{\alpha,W,\mathcal{A}} \,\|\, q) \quad \forall q \in \mathcal{P}(\mathcal{Y}). \tag{46}$$

Заметим, что вид нижней границы (46) в некотором смысле аналогичен границам (21), (23) и (24). Граница (46) является границей ван Эрвена – Харремоеса[17] при $\alpha \in (0,1]$, но не совпадает с ней при $\alpha \in (1,\infty)$, поскольку в этом случае имеется член $D_1(q_{\alpha,W,\mathcal{A}} \,\|\, q)$, а не $D_\alpha(q_{\alpha,W,\mathcal{A}} \,\|\, q)$.

Для порядков выше единицы, используя представление Чисара для информации Августина, приведенное в (25), и определение августиновской пропускной способности, получаем следующие выражения:

$$C_{\alpha,W,\mathcal{A}} = \begin{cases} \sup\limits_{p\in\mathcal{A}} \inf\limits_{V\in\mathcal{P}(\mathcal{Y}|\mathcal{X})} \dfrac{\alpha}{1-\alpha} D_1(V \,\|\, W \,|\, p) + I_1(p;V), & \alpha \in (0,1), \\ \sup\limits_{p\in\mathcal{A}} \sup\limits_{V\in\mathcal{P}(\mathcal{Y}|\mathcal{X})} \dfrac{\alpha}{1-\alpha} D_1(V \,\|\, W \,|\, p) + I_1(p;V), & \alpha \in (1,\infty). \end{cases}$$

Тогда

$$C_{\alpha,W,\mathcal{A}} = \sup_{V\in\mathcal{P}(\mathcal{Y}|\mathcal{X})} \sup_{p\in\mathcal{A}} \frac{\alpha}{1-\alpha} D_1(V \,\|\, W \,|\, p) + I_1(p;V) \quad \forall \alpha \in (1,\infty).$$

Следующая лемма утверждает, что для $\alpha \in (0,1)$, если множество ограничений $\mathcal{A}$ имеет августиновский центр порядка α, например, когда $\mathcal{A}$ выпукло, можно изменить порядок супремума и инфимума, а также заменить инфимум на минимум, когда августиновская пропускная способность конечна.

Лемма 22. *Для любого $\alpha \in (0,1)$, если множество ограничений $\mathcal{A}$ для канала $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ имеет августиновский центр порядка α, то*

$$C_{\alpha,W,\mathcal{A}} = \inf_{V\in\mathcal{P}(\mathcal{Y}|\mathcal{X})} \sup_{p\in\mathcal{A}} \frac{\alpha}{1-\alpha} D_1(V \,\|\, W \,|\, p) + I_1(p;V). \tag{47}$$

Если $C_{\alpha,W,\mathcal{A}}$ конечна, то для $W_\alpha^{q_{\alpha,W,\mathcal{A}}}$ справедливо равенство

$$C_{\alpha,W,\mathcal{A}} = \sup_{p\in\mathcal{A}} \frac{\alpha}{1-\alpha} D_1\big(W_\alpha^{q_{\alpha,W,\mathcal{A}}} \,\|\, W \,|\, p\big) + I_1\big(p; W_\alpha^{q_{\alpha,W,\mathcal{A}}}\big). \tag{48}$$

Лемма 22 доказывается с использованием представления Чисара для информации Августина, данное в лемме 13(e), и леммы 20. В [36] Блейхут доказал аналогичный результат в предположении, что и $\mathcal{X}$, и $\mathcal{Y}$ – конечные множества, а $\mathcal{A} = \mathcal{P}(\mathcal{X})$. Даже при таких предположениях из результата Блейхута [36, теорема 16] следуют равенства (47) и (48) для всех порядков $(0,1)$, только когда $C_{\alpha,W}$ – дифференцируемая функция порядка α. Мотивацией Блейхуту служило выражение для экспоненты сферической упаковки, поэтому теорема 16 в [36] сформулирована в терминах оптимального распределения на входе при данной скорости $R \in (C_{0,W}, C_{1,W})$ и соответствующего оптимального порядка $\alpha^*(R)$.

[17] В [8] ван Эрвен и Харремоес предположили, что неравенство $\sup\limits_{x\in\mathcal{X}} D_\alpha(W(x) \,\|\, q) \geqslant C_{\alpha,W} + D_\alpha(q_{\alpha,W} \,\|\, q)$ справедливо для всех $q \in \mathcal{P}(\mathcal{Y})$, и доказали эту границу в случае $\alpha = \infty$ в предположении, что $\mathcal{Y}$ счетно [8, теорема 37]. Мы подтвердили гипотезу ван Эрвена – Харремоеса в [13, лемма 19] и обобщили ее на случай выпуклого множества ограничений для пропускной способности и центра Реньи в [13, лемма 25]. Краткое обсуждение пропускной способности и центра Реньи см. в п. 4.4; более полное обсуждение можно найти в [13].

4.2. Августиновские пропускная способность и центр как функции порядка.

Лемма 23. *Для любого канала W вида $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ и любого множества ограничений $\mathcal{A} \subset \mathcal{P}(\mathcal{X})$ справедливы следующие утверждения*:

(a) $C_{\alpha,W,\mathcal{A}}$ *– неубывающая и полунепрерывная снизу по α функция на $\mathbb{R}_+$*;

(b) $\dfrac{1-\alpha}{\alpha} C_{\alpha,W,\mathcal{A}}$ *– невозрастающая и непрерывная по α функция на*[18] $(0,1)$;

(c) $(\alpha-1)C_{\alpha,W,\mathcal{A}}$ *– выпуклая по α функция на $(1,\infty)$*;

(d) $C_{\alpha,W,\mathcal{A}}$ *– неубывающая и непрерывная по α функция на $(0,1]$ и $(1,\chi_{W,\mathcal{A}}]$, где* $\chi_{W,\mathcal{A}} \triangleq \sup\{\phi :\, C_{\phi,W,\mathcal{A}} \in \mathbb{R}_{\geqslant 0}\}$;

(e) *Если $\sup\limits_{p\in\mathcal{A}} I^{\mathrm{g}}_{\phi}(p;W) \in \mathbb{R}_{\geqslant 0}$ для $\phi > 1$, то $C_{\alpha,W,\mathcal{A}}$ – неубывающая и непрерывная по α функция на $(0,(1 \vee \chi_{W,\mathcal{A}})]$.*

Результаты о непрерывности, представленные в утверждениях (d) и (e), несколько неудовлетворительны. Хотелось бы либо установить непрерывность $C_{\alpha,W,\mathcal{A}}$ справа в точке $\alpha = 1$, когда $C_{\phi,W,\mathcal{A}}$ конечна для $\phi > 1$, либо указать канал W и множество ограничений $\mathcal{A}$, для которых $C_{\phi,W,\mathcal{A}}$ конечна для $\phi > 1$ и $\lim\limits_{\alpha\downarrow 1} C_{\alpha,W,\mathcal{A}} > C_{1,W,\mathcal{A}}$. Нам не удалось ни то, ни другое. Вместо этого мы доказали непрерывность $C_{\alpha,W,\mathcal{A}}$ справа в точке $\alpha = 1$, предполагая, что $\sup\limits_{p\in\mathcal{A}} I^{\mathrm{g}}_{\phi}(p;W)$ конечен для $\phi > 1$.

Так как $C_{\phi,W} = S_{\phi,W}$ по теореме 1, а $I^{\mathrm{g}}_{\phi}(p;W) \leqslant S_{\phi,W}$ для всех $p \in \mathcal{P}(\mathcal{X})$ согласно (33), то $\sup\limits_{p\in\mathcal{A}} I^{\mathrm{g}}_{\phi}(p;W)$ конечен для всех $\mathcal{A} \subset \mathcal{P}(\mathcal{X})$, когда $C_{\phi,W}$ конечна. Таким образом, $C_{\alpha,W,\mathcal{A}}$ – неубывающая и непрерывная по α функция на $(0,\chi_{W,\mathcal{A}}]$ для всех $\mathcal{A} \subset \mathcal{P}(\mathcal{X})$ при условии, что $C_{\phi,W}$ конечна для $\phi > 1$.

Лемма 21 позволяет использовать непрерывность $C_{\alpha,W,\mathcal{A}}$ по α и лемму 2 для доказательства непрерывности $q_{\alpha,W,\mathcal{A}}$ по α в топологии полной вариации на $\mathcal{P}(\mathcal{Y})$.

Лемма 24. *Для любого $\eta \in \mathbb{R}_+$, $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ и любого выпуклого $\mathcal{A} \subset \mathcal{P}(\mathcal{X})$, такого что $C_{\eta,W,\mathcal{A}} \in \mathbb{R}_+$, справедливо неравенство*

$$D_{\alpha\wedge 1}(q_{\alpha,W,\mathcal{A}} \,\|\, q_{\phi,W,\mathcal{A}}) \leqslant C_{\phi,W,\mathcal{A}} - C_{\alpha,W,\mathcal{A}} \quad \forall \alpha,\phi,\ \text{таких что } 0 < \alpha < \phi \leqslant \eta.$$

Следовательно, если $C_{\alpha,W,\mathcal{A}}$ непрерывна по α на $\mathcal{I}$ для некоторого $\mathcal{I} \subset (0,\eta]$, то $q_{\alpha,W,\mathcal{A}}\colon \mathcal{I} \to \mathcal{P}(\mathcal{Y})$ непрерывна по α на $\mathcal{I}$ в топологии полной вариации на $\mathcal{P}(\mathcal{Y})$.

4.3. Выпуклые оболочки и произведения множеств ограничений. Далее мы рассмотрим два вида часто встречающихся множеств ограничений, описываемых через более простые множества ограничений. В лемме 25 рассматривается выпуклая оболочка семейства множеств ограничений и граница на августиновскую пропускную способность выпуклой оболочки через августиновские пропускные способности индивидуальных множеств ограничений. В лемме 26 рассматривается канал-произведение для множества ограничений, являющегося произведением выпуклых оболочек множеств ограничений на каналы-компоненты, имеющие августиновские центры, и показывается, что августиновская пропускная способность представляется в аддитивной форме, а августиновский центр – в форме произведения.

Лемма 25. *Пусть α – положительное вещественное число, W – канал вида $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$, а $\mathcal{A}^{(i)}$ – множество ограничений, имеющее августиновский центр порядка α и конечную $C_{\alpha,W,\mathcal{A}^{(i)}}$ для всех $i \in \mathcal{T}$. Тогда*

$$\sup_{i\in\mathcal{T}} C_{\alpha,W,\mathcal{A}^{(i)}} \leqslant C_{\alpha,W,\mathcal{A}} \leqslant \ln \sum_{i\in\mathcal{T}} e^{C_{\alpha,W,\mathcal{A}^{(i)}}},$$

[18] Случай $\alpha = 1$ мы исключаем, поскольку не хотим предполагать $C_{1,W,\mathcal{A}}$ конечной.

где $\mathcal{A}$ – выпуклая оболочка объединения, т.е. $\mathcal{A} = \mathrm{ch}\Big(\bigcup\limits_{i\in\mathcal{T}} \mathcal{A}^{(i)}\Big)$. *Кроме того,*

- $C_{\alpha,W,\mathcal{A}^{(i)}} = C_{\alpha,W,\mathcal{A}} < \infty \Leftrightarrow \sup\limits_{p\in\mathcal{A}} D_\alpha(W \,\|\, q_{\alpha,W,\mathcal{A}^{(i)}} \,|\, p) \leqslant C_{\alpha,W,\mathcal{A}^{(i)}} \Rightarrow q_{\alpha,W,\mathcal{A}} = = q_{\alpha,W,\mathcal{A}^{(i)}}$;
- $C_{\alpha,W,\mathcal{A}} = \ln \sum\limits_{i\in\mathcal{T}} e^{C_{\alpha,W,\mathcal{A}^{(i)}}} < \infty \Leftrightarrow q_{\alpha,W,\mathcal{A}^{(i)}} \perp q_{\alpha,W,\mathcal{A}^{(j)}}\ \forall i \neq j$ *и* $|\mathcal{T}| < \infty \Rightarrow$ $\Rightarrow q_{\alpha,W,\mathcal{A}} = \sum\limits_{i\in\mathcal{T}} \dfrac{e^{C_{\alpha,W,\mathcal{A}^{(i)}}}}{e^{C_{\alpha,W,\mathcal{A}}}} q_{\alpha,W,\mathcal{A}^{(i)}}$.

Отметим, что если $\mathcal{A}^{(i)}$ выпукло, а $C_{\alpha,W,\mathcal{A}^{(i)}}$ конечна, то $\mathcal{A}^{(i)}$ имеет единственный августиновский центр порядка α по теореме 1.

Лемма 26. *Для любых $\alpha \in \mathbb{R}_+$, канала-произведения $W_{[1,n]}\colon \mathcal{X}_1^n \to \mathcal{P}(\mathcal{Y}_1^n)$ длины n и множеств ограничений $\mathcal{A}_t \subset \mathcal{P}(\mathcal{X}_t)$, имеющих августиновские центры порядка α,*

$$C_{\alpha,W_{[1,n]},\mathcal{A}} = C_{\alpha,W_{[1,n]},\mathcal{A}_1^n} = \sum_{t=1}^{n} C_{\alpha,W_t,\mathcal{A}_t},$$

где $\mathcal{A} = \big\{p \in \mathcal{P}(\mathcal{X}_1^n) : \ p_t \in \mathrm{ch}\,\mathcal{A}_t\ \forall t \in \{1,\ldots,n\}\big\}$, т.е. $p \in \mathcal{P}(\mathcal{X}_1^n)$ принадлежит $\mathcal{A}$ тогда и только тогда, когда для всех $t \in \{1,\ldots,n\}$ его маргинальное распределение p_t на $\mathcal{X}_t$ лежит в выпуклой оболочке $\mathcal{A}_t$. Кроме того, если $C_{\alpha,W_t,\mathcal{A}_t}$ конечна для всех $t \in \{1,\ldots,n\}$, то $q_{\alpha,W_{[1,n]},\mathcal{A}} = q_{\alpha,W_{[1,n]},\mathcal{A}_1^n} = \bigotimes\limits_{t=1}^{n} q_{\alpha,W_t,\mathcal{A}_t}$.

Замечание 2. Заметим, что выпуклая оболочка любого подмножества $\mathcal{A}$ является подмножеством $\mathcal{A}$, поскольку $\mathcal{A}$ выпукло по определению. В частности, $\mathcal{A}_1^n \subset \subset \mathrm{ch}\,\mathcal{A}_1^n \subset \mathcal{A}$. Тогда $C_{\alpha,W_{[1,n]},\mathrm{ch}\,\mathcal{A}_1^n} = \sum\limits_{t=1}^{n} C_{\alpha,W_t,\mathcal{A}_t}$ по лемме 26. Кроме того, если $C_{\alpha,W_t,\mathcal{A}_t}$ конечна для всех $t \in \{1,\ldots,n\}$, то $q_{\alpha,W_{[1,n]},\mathrm{ch}\,\mathcal{A}_1^n} = \bigotimes\limits_{t=1}^{n} q_{\alpha,W_t,\mathcal{A}_t}$ по лемме 25.

Замечание 3. Множество ограничений $\mathcal{A}_1^n$, описанное в лемме 26, может не быть выпуклым, но обязано $\mathcal{A}_1^n$ иметь августиновский центр порядка α.

4.4. Сравнения пропускных способностей Августина и Реньи. Используя информацию Реньи вместо информации Августина, можно определить пропускную способность Реньи следующим образом.

Определение 12. Для любых $\alpha \in \mathbb{R}_+$, $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ и $\mathcal{A} \subset \mathcal{P}(\mathcal{X})$ *пропускной способностью Реньи порядка α канала W для множества ограничений $\mathcal{A}$* называется

$$C^{\mathrm{g}}_{\alpha,W,\mathcal{A}} \triangleq \sup_{p\in\mathcal{A}} I^{\mathrm{g}}_\alpha(p;W).$$

Когда множество ограничений $\mathcal{A}$ совпадает со всем $\mathcal{P}(\mathcal{X})$, пропускная способность Реньи порядка α обозначается через $C^{\mathrm{g}}_{\alpha,W}$, т.е. $C^{\mathrm{g}}_{\alpha,W} \triangleq C^{\mathrm{g}}_{\alpha,W,\mathcal{P}(\mathcal{X})}$.

Так как $I_1(p;W) = I^{\mathrm{g}}_1(p;W)$, то $C^{\mathrm{g}}_{1,W,\mathcal{A}} = C_{1,W,\mathcal{A}}$ по определению. Для других порядков утверждать этого нельзя; согласно (40), (41) имеем

$$\begin{aligned} &C^{\mathrm{g}}_{\alpha,W,\mathcal{A}} \leqslant C_{\alpha,W,\mathcal{A}}, \quad \alpha \in (0,1],\\ &C^{\mathrm{g}}_{\alpha,W,\mathcal{A}} \geqslant C_{\alpha,W,\mathcal{A}}, \quad \alpha \in [1,\infty). \end{aligned}$$

Из определений информации Реньи и пропускной способности Реньи следует, что

$$C^{\mathrm{g}}_{\alpha,W,\mathcal{A}} = \sup_{p\in\mathcal{A}} \inf_{q\in\mathcal{P}(\mathcal{Y})} D_\alpha(p \circledast W \,\|\, p \otimes q).$$

Для пропускной способности Реньи имеет место теорема о минимаксе [13, теорема 2], аналогичная теореме 1: Для любого выпуклого множества ограничений $\mathcal{A} \subset \mathcal{P}(\mathcal{X})$

$$\sup_{p \in \mathcal{A}} \inf_{q \in \mathcal{P}(\mathcal{Y})} D_\alpha(p \circledast W \,\|\, p \otimes q) = \inf_{q \in \mathcal{P}(\mathcal{Y})} \sup_{p \in \mathcal{A}} D_\alpha(p \circledast W \,\|\, p \otimes q).$$

Если $C^{\mathrm{g}}_{\alpha,W,\mathcal{A}}$ конечна, то $\exists!\, q^{\mathrm{g}}_{\alpha,W,\mathcal{A}} \in \mathcal{P}(\mathcal{Y})$, называемая *центром Реньи порядка* α *канала* W *для множества ограничений* $\mathcal{A}$, для которой

$$C^{\mathrm{g}}_{\alpha,W,\mathcal{A}} = \sup_{p \in \mathcal{A}} D_\alpha\bigl(p \circledast W \,\|\, p \otimes q^{\mathrm{g}}_{\alpha,W,\mathcal{A}}\bigr).$$

Следовательно, пропускная способность Реньи равна радиусу Реньи при условии, что $\mathcal{A} = \mathcal{P}(\mathcal{X})$. Отсюда $C^{\mathrm{g}}_{\alpha,W} = C_{\alpha,W}$ и $q^{\mathrm{g}}_{\alpha,W} = q_{\alpha,W}$ по теореме 1. Остальные наблюдения, представленные в этом параграфе, также имеют свои аналоги для пропускной способности и центра Реньи; ср., например, лемму 21 и [13, лемма 25].

§ 5. Задача при наличии ограничений по стоимости

В § 4 мы определили августиновскую пропускную способность для произвольных множеств ограничений и доказали существование и единственность августиновского центра для любого выпуклого множества ограничений с конечной августиновской пропускной способностью. Представляющие содержательный интерес выпуклые множества ограничений часто определяются через ограничения по стоимости, и основная цель настоящего параграфа – исследовать этот важный частный случай более подробно. В п. 5.1 рассматриваются прямые следствия определения августиновской пропускной способности при наличии ограничений по стоимости и уточнения некоторых приведенных выше. В п. 5.2 определяются и исследуются понятия информации, пропускной способности, радиуса и центра Августина – Лежандра (АЛ). Результаты п. 5.2 являются обобщением некоторых результатов из [5, гл. 8] о супремуме взаимной информации для дискретных каналов с одним ограничением по стоимости, т.е. для случая $\alpha = 1$, $|\mathcal{X}| < \infty$, $|\mathcal{Y}| < \infty$, $\ell = 1$. В п. 5.3 определяются и изучаются понятия информации, среднего, пропускной способности, радиуса и центра Реньи – Галлагера (РГ). Наиболее важный результат анализа в п. 5.3 – равенство АЛ-пропускной способности и АЛ-центра, соответственно, РГ-пропускной способности и РГ-центру. В п. 5.4 показано, как можно использовать результаты пп. 5.1–5.3 для нахождения августиновских пропускной способности и центра для функции вероятностей переходов при ограничениях по стоимости. Доказательства утверждений, представленных в пп. 5.1–5.3, можно найти в [22, Приложение E].

В [6, § 34] Августин рассматривал пропускную способность при наличии ограничений по стоимости $C_{\alpha,W,\varrho}$ для случая, когда функция стоимости ρ имеет вид $\rho\colon \mathcal{X} \to [0,1]^\ell$, а $\alpha \in (0,1]$. Там же были исследованы некоторые величины, тесно связанные с РГ-информацией и РГ-пропускной способностью, которые впервые появились у Галлагера при изучении экспонент вероятности ошибки в каналах при наличии ограничений по стоимости [14, § 6; 15, §§ 7.3–7.5]. В отличие от Августина Галлагер не предполагал функцию ρ ограниченной, однако Галлагер ограничивался случаем одного ограничения по стоимости, т.е. случаем $\ell = 1$, и не рассматривал РГ-пропускную способность как величину, имеющую самостоятельный интерес. РГ-информацию и РГ-пропускную способность рассматривали некоторые другие авторы при изучении вопросов, связанных с ограничениями по стоимости [24, § IV; 25–27]. Однако, насколько нам известно, для порядков, отличных от единицы, АЛ-информационные величины, получаемые более прямым применением выпуклого сопряжения, ранее не изучались.

5.1. Августиновские пропускная способность и центр при наличии ограничений по стоимости. Обозначим множество всех функций вероятности, удовлетворяющих ограничению по стоимости ϱ, через $\mathcal{A}(\varrho)$, т.е.

$$\mathcal{A}(\varrho) \triangleq \{p \in \mathcal{P}(\mathcal{X}):\ \mathbf{E}_p[\rho] \leqslant \varrho\}.$$

При этом $\mathcal{A}(\varrho) \neq \varnothing$ тогда и только тогда, когда $\varrho \in \Gamma_\rho$, где Γ_ρ определено в (6) как множество всех допустимых ограничений по стоимости для функции стоимости ρ. Функция $\mathcal{A}(\varrho)$ не убывает по ϱ, т.е. если $\varrho_1 \leqslant \varrho_2$, то $\mathcal{A}(\varrho_1) \subset \mathcal{A}(\varrho_2)$. Августиновской пропускной способностью порядка α канала W при ограничении по стоимости ϱ называется

$$C_{\alpha,W,\varrho} \triangleq \begin{cases} \sup\limits_{p \in \mathcal{A}(\varrho)} I_\alpha(p;W), & \text{если } \varrho \in \Gamma_\rho, \\ -\infty, & \text{если } \varrho \in \mathbb{R}^\ell_{\geqslant 0} \setminus \Gamma_\rho, \end{cases} \qquad \forall \alpha \in \mathbb{R}_+. \tag{49}$$

Мы определили $C_{\alpha,W,\varrho}$ для недопустимых значений ϱ, чтобы иметь возможность пользоваться стандартными результатами без изменений. Так как множество $\mathcal{A}(\varrho)$ выпукло, для него справедлива теорема 1. Обозначим[19] августиновский центр порядка α канала W при ограничении по стоимости ϱ через $q_{\alpha,W,\varrho}$.

Для заданного порядка α августиновская пропускная способность $C_{\alpha,W,\varrho}$ является вогнутой функцией ограничения по стоимости ϱ. Следовательно, если она конечна во внутренней точке множества Γ_ρ, то она является непрерывной функцией ограничения по стоимости ϱ ниже касательных плоскостей к внутренним точкам Γ_ρ. Эти наблюдения обобщает следующая

Лемма 27. *Пусть W – канал вида $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ с функцией стоимости ρ вида $\rho\colon \mathcal{X} \to \mathbb{R}^\ell_{\geqslant 0}$.*

(a) *Для любого $\alpha \in \mathbb{R}_+$ функция $C_{\alpha,W,\varrho}$ не убывает и вогнута по ϱ на $\mathbb{R}^\ell_{\geqslant 0}$ и при этом либо бесконечна в любой точке из $\operatorname{int}\Gamma_\rho$, либо конечна и непрерывна на $\operatorname{int}\Gamma_\rho$;*

(b) *Если $C_{\alpha,W,\varrho}$ конечна на $\operatorname{int}\Gamma_\rho$ для некоторого $\alpha \in \mathbb{R}_+$, то для любого $\varrho \in \operatorname{int}\Gamma_\rho$ существует $\lambda_{\alpha,W,\varrho} \in \mathbb{R}^\ell_{\geqslant 0}$, такое что*

$$C_{\alpha,W,\varrho} + \lambda_{\alpha,W,\varrho} \cdot (\widetilde{\varrho} - \varrho) \geqslant C_{\alpha,W,\widetilde{\varrho}} \quad \forall \widetilde{\varrho} \in \mathbb{R}^\ell_{\geqslant 0} \tag{50}$$

Кроме того, множество всех таких $\lambda_{\alpha,W,\varrho}$ выпукло и компактно;

(c) *Либо $C_{\alpha,W,\varrho} = \infty$ для всех $(\alpha,\varrho) \in (0,1) \times \operatorname{int}\Gamma_\rho$, либо $C_{\alpha,W,\varrho}$ и $q_{\alpha,W,\varrho}$ непрерывны по (α,ϱ) на $(0,1) \times \operatorname{int}\Gamma_\rho$ в топологии полной вариации на $\mathcal{P}(\mathcal{Y})$.*

Если функция стоимости для канала-произведения аддитивна, то августиновская пропускная способность при наличии ограничений по стоимости для канала-произведения равна супремуму суммы августиновских пропускных способностей при наличии ограничений по стоимости для каналов-компонентов по всем допустимым распределениям стоимости. Кроме того, если существует оптимальное распределение стоимости, то августиновский центр канала-произведения является произведением мер. Эти наблюдения формализует следующая

Лемма 28. *Для любого канала-произведения $W_{[1,n]}\colon \mathcal{X}_1^n \to \mathcal{P}(\mathcal{Y}_1^n)$ длины n и любой аддитивной функции стоимости $\rho_{[1,n]}\colon \mathcal{X}_1^n \to \mathbb{R}^\ell_{\geqslant 0}$ имеет место равенство*[20]

[19] Эта небольшая вольность в обозначениях, которой можно было бы избежать, используя $C_{\alpha,W,\mathcal{A}(\varrho)}$ и $q_{\alpha,W,\mathcal{A}(\varrho)}$ вместо $C_{\alpha,W,\varrho}$ и $q_{\alpha,W,\varrho}$, упрощает обозначения и не приводит к недоразумениям.

[20] Если $C_{\alpha,W_t,\varrho_t} = -\infty$ для всех $t \in \{1,\ldots,n\}$, то сумма $\sum\limits_{t=1}^{n} C_{\alpha,W_t,\varrho_t}$ равна $-\infty$, даже если одна или несколько других C_{α,W_t,ϱ_t} равны ∞.

$$C_{\alpha,W_{[1,n]},\varrho} = \sup\left\{\sum_{t=1}^{n} C_{\alpha,W_t,\varrho_t} : \sum_{t=1}^{n} \varrho_t \leqslant \varrho,\ \ \varrho_t \in \mathbb{R}^{\ell}_{\geqslant 0}\right\} \quad \forall \varrho \in \mathbb{R}^{\ell}_{\geqslant 0}, \alpha \in \mathbb{R}_+.$$

Если $C_{\alpha,W_{[1,n]},\varrho} \in \mathbb{R}_{\geqslant 0}$ *для некоторого* $\alpha \in \mathbb{R}_+$ *и* $\exists(\varrho_1, \ldots, \varrho_n)$, *такие что* $C_{\alpha,W_{[1,n]},\varrho} = \sum_{t=1}^{n} C_{\alpha,W_t,\varrho_t}$, *то* $q_{\alpha,W_{[1,n]},\varrho} = \bigotimes_{t=1}^{n} q_{\alpha,W_t,\varrho_t}$.

Так как августиновская пропускная способность вогнута как функция ограничения по стоимости согласно лемме 27(a), то $C_{\alpha,W_{[1,n]},\varrho} = \sum_{t=1}^{n} C_{\alpha,W_t,\varrho/n}$, когда $W_{[1,n]}$ – стационарный канал и $\rho_t = \rho_1$ для всех $t \in \{1, \ldots, n\}$. Напротив, если Γ_{ρ_t} – замкнутые множества, а $C_{\alpha,W_t,\varrho}$ – полунепрерывные сверху функции ϱ на множествах Γ_{ρ_t}, то теорема об экстремуме[21] для полунепрывных функций дает существование $(\varrho_1, \ldots, \varrho_n)$, удовлетворяющих условиям $C_{\alpha,W_{[1,n]},\varrho} = \sum_{t=1}^{n} C_{\alpha,W_t,\varrho_t}$ и $\sum_{t=1}^{n} \varrho_t \leqslant \varrho$. Однако это утверждение о существовании в общем случае не верно – см. пример 3.

5.2. Информационные величины Августина – Лежандра. Августиновскую пропускную способность при наличии ограничений по стоимости $C_{\alpha,W,\varrho}$ и августиновский центр $q_{\alpha,W,\varrho}$ можно также охарактеризовать, используя выпуклое сопряжение. В этом пункте мы вводим понятия информации, пропускной способности, центра и радиуса Августина – Лежандра, позволяющие прийти к более полному пониманию этой характеризации. Подходящим методом нам видится стандартное применение техники выпуклого сопряжения для характеризации августиновской пропускной способности при наличии ограничений по стоимости. Однако этот метод не является общепринятым. Начиная с оригинальной работы Галлагера [14], стандартным способом применения техники множителей Лагранжа для характеризации августиновской пропускной способности стал более подходящий для данного случая метод, основанный на информации Реньи – см. [6, § 35; 25; 26]. Этот подход обсуждается в п. 5.3. Теорема 2, приведенная ниже, и теорема 3 из п. 5.3 доказывают эквивалентность этих подходов, устанавливая равенство пропускной способности и центра Августина – Лежандра, соответственно, пропускной способности и центру Реньи – Галлагера.

Определение 13. Для любых $\alpha \in \mathbb{R}_+$, канала W вида $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ с функцией стоимости $\rho\colon \mathcal{X} \to \mathbb{R}^{\ell}_{\geqslant 0}$, $p \in \mathcal{P}(\mathcal{X})$ и $\lambda \in \mathbb{R}^{\ell}_{\geqslant 0}$ *информацией Августина – Лежандра порядка* α *для распределения на входе* p *и множителя Лагранжа* λ называется

$$I^{\lambda}_{\alpha}(p; W) \triangleq I_{\alpha}(p; W) - \lambda \cdot \mathbf{E}_p[\rho]. \tag{51}$$

Заметим, что немедленным следствием определения АЛ-информации является равенство

$$\inf_{\lambda \geqslant 0} I^{\lambda}_{\alpha}(p; W) + \lambda \cdot \varrho = \xi_{\alpha,p}(\varrho),$$

где $\xi_{\alpha,p}(\cdot)\colon \mathbb{R}^{\ell}_{\geqslant 0} \to [-\infty, \infty)$ определяется как

$$\xi_{\alpha,p}(\varrho) \triangleq \begin{cases} I_{\alpha}(p; W), & \varrho \geqslant \mathbf{E}_p[\rho], \\ -\infty & \text{в противном случае.} \end{cases} \tag{52}$$

---

[21] Рассмотрим функцию $f(\varrho_1, \ldots, \varrho_n)$, равную $\sum_{t=1}^{n} C_{\alpha,W_t,\varrho_t}$, если $\sum_{t=1}^{n} \varrho_t \leqslant \varrho$ и $\varrho_t \in \Gamma_{\rho_t}$ для всех $t \in \{1, \ldots, n\}$, и равную $-\infty$ в противном случае. Выбирая достаточно большое, но ограниченное множество, используя вектор ϱ, получаем компактное множество под знаком супремума.

Тогда информацию Августина – Лежандра $I_\alpha^\lambda(p;W)$ можно также задать в виде

$$I_\alpha^\lambda(p;W) = \sup_{\varrho \geqslant 0} \xi_{\alpha,p}(\varrho) - \lambda \cdot \varrho. \tag{53}$$

Замечание 4. Заметим, что если функции $f\colon \mathbb{R}^\ell_{\geqslant 0} \to (-\infty,\infty]$ и $f^*\colon (-\infty,0]^\ell \to \mathbb{R}$ определить как $f(\varrho) \triangleq -\xi_{\alpha,p}(\varrho)$ и $f^*(\gamma) \triangleq I_\alpha^{-\gamma}(p;W)$, то f^* является выпукло сопряженной функцией, т.е. преобразованием Лежандра, для f. Именно по этой причине мы называем $I_\alpha^\lambda(p;W)$ информацией Августина – Лежандра.

О п р е д е л е н и е 14. Для любых $\alpha \in \mathbb{R}_+$, канала W вида $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ с функцией стоимости $\rho\colon \mathcal{X} \to \mathbb{R}^\ell_{\geqslant 0}$ и $\lambda \in \mathbb{R}^\ell_{\geqslant 0}$ *пропускной способностью Августина – Лежандра (АЛ-пропускной способностью) порядка* α *для множител Лагранжа* λ называется

$$C_{\alpha,W}^\lambda \triangleq \sup_{p \in \mathcal{P}(\mathcal{X})} I_\alpha^\lambda(p;W). \tag{54}$$

Тогда из (52), (53) вытекает

$$C_{\alpha,W}^\lambda = \sup_{\varrho \geqslant 0} C_{\alpha,W,\varrho} - \lambda \cdot \varrho \quad \forall \lambda \in \mathbb{R}^\ell_{\geqslant 0}. \tag{55}$$

Следовательно, из неравенства максимина-минимакса получаем

$$C_{\alpha,W,\varrho} \leqslant \inf_{\lambda \geqslant 0} C_{\alpha,W}^\lambda + \lambda \cdot \varrho \quad \forall \varrho \in \mathbb{R}^\ell_{\geqslant 0}. \tag{56}$$

Поэтому $C_{\alpha,W,\varrho} < \infty$ для всех $\varrho \in \mathbb{R}^\ell_{\geqslant 0}$ при условии, что $C_{\alpha,W}^\lambda < \infty$ для некоторого $\lambda \in \mathbb{R}_{\geqslant 0}$. Но равенство $C_{\alpha,W}^\lambda = \infty$ может выполняться для достаточно малых λ, когда $C_{\alpha,W,\varrho} < \infty$ для всех $\varrho \in \mathbb{R}^\ell_{\geqslant 0}$, – см. пример 1.

Замечание 5. В [6, §§ 33–35], Августин рассмотрел случай, когда функция стоимости ρ является ограниченной функцией вида $\rho\colon \mathcal{X} \to [0,1]^\ell$. В этом случае $C_{\alpha,W}^\lambda < \infty$ для всех $\lambda \in \mathbb{R}^\ell_{\geqslant 0}$ при условии, что $C_{\alpha,W,\varrho} < \infty$ для некоторого $\varrho \in \operatorname{int}\Gamma_\rho$, поскольку $C_{\alpha,W,\mathbb{1}} < \infty$ согласно лемме 27(b), а $C_{\alpha,W,\mathbb{1}} = C_{\alpha,W}$ и $C_{\alpha,W}^\lambda \leqslant C_{\alpha,W}^0 = C_{\alpha,W}$ для всех $\lambda \in \mathbb{R}^\ell_{\geqslant 0}$ по определению.

Неравенство (56) во многих содержательных случаях обращается в равенство, как показывает следующая лемма. Однако в общем случае оно равенством не является – см. пример 2.

Л е м м а 29. *Пусть* $\alpha \in \mathbb{R}_+$, *и пусть* W *– канал вида* $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ *с функцией стоимости* $\rho\colon \mathcal{X} \to \mathbb{R}^\ell_{\geqslant 0}$. *Тогда*

(a) $C_{\alpha,W}^\lambda$ *– выпуклая невозрастающая полунепрерывная снизу по* λ *функция на* $\mathbb{R}^\ell_{\geqslant 0}$, *непрерывная по* λ *на множестве* $\{\lambda : \exists \varepsilon > 0,$ *такое что* $C_{\alpha,W}^{\lambda-\varepsilon\mathbb{1}} < \infty\}$;

(b) *Если* $\mathcal{X}$ *– конечное множество, то* $C_{\alpha,W,\varrho} = \inf_{\lambda \geqslant 0} C_{\alpha,W}^\lambda + \lambda \cdot \varrho$;

(c) *Если* $\varrho \in \operatorname{int}\Gamma_\rho$, *то* $C_{\alpha,W,\varrho} = \inf_{\lambda \geqslant 0} C_{\alpha,W}^\lambda + \lambda \cdot \varrho$. *Если при этом* $C_{\alpha,W,\varrho} < \infty$, *то существует непустое выпуклое компактное множество значений* $\lambda_{\alpha,W,\varrho}$, *для которых выполнены условия* (50) *и* $C_{\alpha,W,\varrho} = C_{\alpha,W}^{\lambda_{\alpha,W,\varrho}} + \lambda_{\alpha,W,\varrho} \cdot \varrho$;

(d) *Если* $C_{\alpha,W,\varrho}$ *конечна и* $C_{\alpha,W,\varrho} = C_{\alpha,W}^\lambda + \lambda \cdot \varrho$ *для некоторых* $\varrho \in \Gamma_\rho$ *и* $\lambda \in \mathbb{R}^\ell_{\geqslant 0}$, *то* $\lim_{i\to\infty} I_\alpha^\lambda(p^{(i)};W) = C_{\alpha,W}^\lambda$ *для всех* $\{p^{(i)}\}_{i\in\mathbb{Z}_+} \in \mathcal{A}(\varrho)$, *таких что* $\lim_{i\to\infty} I_\alpha(p^{(i)};W) = C_{\alpha,W,\varrho}$.

Используя определения величин $I_\alpha(p;W)$, $I_\alpha^\lambda(p;W)$ и $C_{\alpha,W}^\lambda$, данные в (17), (51) и (54), получаем следующее выражение для $C_{\alpha,W}^\lambda$:

$$C_{\alpha,W}^\lambda = \sup_{p\in\mathcal{P}(\mathcal{X})} \inf_{q\in\mathcal{P}(\mathcal{Y})} D_\alpha(W \,\|\, q \,|\, p) - \lambda\cdot \mathbf{E}_p[\rho]. \tag{57}$$

АЛ-пропускная способность удовлетворяет теореме о минимаксе, аналогичной теореме о минимаксе для августиновской пропускной способности, что позволяет утверждать существование и единственность АЛ-центра в случае, когда АЛ-пропускная способность конечна.

Теорема 2. *Для любых $\alpha\in\mathbb{R}_+$, канала $W\colon \mathcal{X}\to\mathcal{P}(\mathcal{Y})$ с функцией стоимости $\rho\colon \mathcal{X}\to\mathbb{R}_{\geqslant 0}^\ell$ и множителя Лагранжа $\lambda\in\mathbb{R}_{\geqslant 0}^\ell$ справедливы равенства*

$$\sup_{p\in\mathcal{P}(\mathcal{X})} \inf_{q\in\mathcal{P}(\mathcal{Y})} D_\alpha(W \,\|\, q \,|\, p) - \lambda\cdot \mathbf{E}_p[\rho] = \inf_{q\in\mathcal{P}(\mathcal{Y})} \sup_{p\in\mathcal{P}(\mathcal{X})} D_\alpha(W \,\|\, q \,|\, p) - \lambda\cdot \mathbf{E}_p[\rho] = \tag{58}$$

$$= \inf_{q\in\mathcal{P}(\mathcal{Y})} \sup_{x\in\mathcal{X}} D_\alpha(W(x) \,\|\, q) - \lambda\cdot\rho(x). \tag{59}$$

Если выражение в левой части (58) конечно, т.е. если $C_{\alpha,W}^\lambda<\infty$, то $\exists!\, q_{\alpha,W}^\lambda \in \in \mathcal{P}(\mathcal{Y})$, называемая центром Августина – Лежандра порядка α канала W для множителя Лагранжа λ, такая что

$$C_{\alpha,W}^\lambda = \sup_{p\in\mathcal{P}(\mathcal{X})} D_\alpha\big(W \,\|\, q_{\alpha,W}^\lambda \,|\, p\big) - \lambda\cdot \mathbf{E}_p[\rho] = \tag{60}$$

$$= \sup_{x\in\mathcal{X}} D_\alpha\big(W(x) \,\|\, q_{\alpha,W}^\lambda\big) - \lambda\cdot\rho(x). \tag{61}$$

Кроме того, для любой последовательности распределений на входе $\{p^{(i)}\}_{i\in\mathbb{Z}_+}\subset \subset \mathcal{P}(\mathcal{X})$, такой что $\lim\limits_{i\to\infty} I_\alpha^\lambda(p^{(i)};W) = C_{\alpha,W}^\lambda$, соответствующая последовательность $\{q_{\alpha,p^{(i)}}\}_{i\in\mathbb{Z}_+}$ августиновских средних порядка α является последовательностью Коши в метрике полной вариации на $\mathcal{P}(\mathcal{Y})$, а $q_{\alpha,W}^\lambda$ – единственная предельная точка этой последовательности Коши.

Заметим, что теорема 2 при $\lambda=0$ есть ни что иное, как теорема 1 для $\mathcal{A}=\mathcal{P}(\mathcal{X})$. Доказательство теоремы 2 также очень похоже на доказательство теоремы 1, но вместо леммы 19 оно использует следующую лемму 30. Отметим, что лемма 30 при $\lambda=0$ также есть ни что иное, как лемма 19 для $\mathcal{A}=\mathcal{P}(\mathcal{X})$.

Лемма 30. *Для любых $\alpha\in\mathbb{R}_+$, канала $W\colon \mathcal{X}\to\mathcal{P}(\mathcal{Y})$ с функцией стоимости $\rho\colon \mathcal{X}\to\mathbb{R}_{\geqslant 0}^\ell$ для конечного множества входов $\mathcal{X}$ и множителя Лагранжа $\lambda\in\mathbb{R}_{\geqslant 0}^\ell$ существует $\widetilde{p}\in\mathcal{P}(\mathcal{X})$, такое что $I_\alpha^\lambda(\widetilde{p};W) = C_{\alpha,W}^\lambda$, и $\exists!\, q_{\alpha,W}^\lambda\in\mathcal{P}(\mathcal{Y})$, такое что*

$$D_\alpha\big(W \,\|\, q_{\alpha,W}^\lambda \,|\, p\big) - \lambda\cdot \mathbf{E}_p[\rho] \leqslant C_{\alpha,W}^\lambda \quad \forall p\in\mathcal{P}(\mathcal{X}).$$

Кроме того, $q_{\alpha,\widetilde{p}} = q_{\alpha,W}^\lambda$ для всех $\widetilde{p}\in\mathcal{P}(\mathcal{X})$, таких что $I_\alpha^\lambda(\widetilde{p};W) = C_{\alpha,W}^\lambda$.

Заметим, что выражение в левой части (58) есть ни что иное, как АЛ-пропускная способность. Таким образом, теорема 2 устанавливает равенство АЛ-пропускной способности и АЛ-радиуса, определяемого следующим образом.

Определение 15. Для любых $\alpha\in\mathbb{R}_+$, канала $W\colon \mathcal{X}\to\mathcal{P}(\mathcal{Y})$ с функцией стоимости $\rho\colon \mathcal{X}\to\mathbb{R}_{\geqslant 0}^\ell$ и $\lambda\in\mathbb{R}_{\geqslant 0}^\ell$ *радиусом Августина – Лежандра порядка α канала W для множителя Лагранжа λ* называется

$$S_{\alpha,W}^\lambda \triangleq \inf_{q\in\mathcal{P}(\mathcal{Y})} \sup_{x\in\mathcal{X}} D_\alpha(W(x) \,\|\, q) - \lambda\cdot\rho(x). \tag{62}$$

Если $C^{\lambda}_{\alpha,W}$ конечна, то из леммы 13(b)–(d), теоремы 2 и определения величины $I^{\lambda}_{\alpha}(p;W)$, данного в (51), следует, что

$$C^{\lambda}_{\alpha,W} - I^{\lambda}_{\alpha}(p;W) \geqslant D_{\alpha\wedge 1}\big(q_{\alpha,p} \,\|\, q^{\lambda}_{\alpha,W}\big) \quad \forall p \in \mathcal{P}(\mathcal{X}).$$

С помощью леммы 13 и теоремы 2 можно также получить границу, аналогичную границе из леммы 21. Однако этого мы делать не будем, поскольку несколько более сильные результаты можно получить, используя характеризацию АЛ-пропускной способности и АЛ-центра через РГ-пропускную способность и РГ-центр, описанную в п. 5.3, – см. лемму 35 и последующее обсуждение.

Как следствие леммы 29(c), мы знаем, что если $C_{\alpha,W,\varrho}$ конечна для некоторого $\varrho \in \operatorname{int}\Gamma_{\rho}$, то существует по крайней мере один множитель Лагранжа $\lambda_{\alpha,W,\varrho}$, для которого выполнено равенство $C_{\alpha,W,\varrho} = C^{\lambda_{\alpha,W,\varrho}}_{\alpha,W} + \lambda_{\alpha,W,\varrho} \cdot \varrho$. Следующая лемма 31 утверждает, что для любого такого множителя Лагранжа соответствующий АЛ-центр порядка α должен быть равен августиновскому центру порядка α для ограничения по стоимости ϱ. Таким образом, если существуют несколько $\lambda_{\alpha,W,\varrho}$, таких что $C_{\alpha,W,\varrho} = C^{\lambda_{\alpha,W,\varrho}}_{\alpha,W} + \lambda_{\alpha,W,\varrho} \cdot \varrho$, то все они имеют один и тот же АЛ-центр порядка α.

Лемма 31. *Для любых $\alpha \in \mathbb{R}_{+}$, канала $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ с функцией стоимости $\rho\colon \mathcal{X} \to \mathbb{R}^{\ell}_{\geqslant 0}$ и ограничения по стоимости $\varrho \in \Gamma_{\rho}$, таких что $C_{\alpha,W,\varrho} < \infty$, если $C_{\alpha,W,\varrho} = C^{\lambda}_{\alpha,W} + \lambda \cdot \varrho$ для некоторого $\lambda \in \mathbb{R}^{\ell}_{\geqslant 0}$, то $q_{\alpha,W,\varrho} = q^{\lambda}_{\alpha,W}$.*

Для произведений множеств ограничений в каналах-произведениях августиновская пропускная способность представляется в аддитивном виде, а августиновский центр – в мультипликативном (когда он существует) по лемме 26. Однако ограничения по стоимости для аддитивных функций стоимости не имеют вид произведений. Чтобы вычислить августиновскую пропускную способность при наличии ограничений по стоимости для каналов-произведений с аддитивными функциями стоимости, требуется провести оптимизацию по допустимым распределениям стоимости по каналам-компонентам согласно лемме 28. При этом августиновский центр при наличии ограничений по стоимости для канала-произведения можно выразить с помощью леммы 28 в виде произведения августиновских центров при наличии ограничений по стоимости для каналов-компонентов, только когда существует допустимое распределение стоимости, на котором достигается оптимальное значение. С другой стороны, для АЛ-пропускной способности и АЛ-центра имеется значительно более четкая картина: для каналов-произведений с аддитивными функциями стоимости АЛ-пропускная способность аддитивна, а АЛ-центр мультипликативен, если он существует.

Лемма 32. *Для любого канала-произведения $W_{[1,n]}\colon \mathcal{X}_1^n \to \mathcal{P}(\mathcal{Y}_1^n)$ длины n и любой аддитивной функции стоимости $\rho_{[1,n]}\colon \mathcal{X}_1^n \to \mathbb{R}^{\ell}_{\geqslant 0}$ выполнено равенство*

$$C^{\lambda}_{\alpha,W_{[1,n]}} = \sum_{t=1}^{n} C^{\lambda}_{\alpha,W_t} \quad \forall \lambda \in \mathbb{R}^{\ell}_{\geqslant 0},\ \alpha \in \mathbb{R}_{+}.$$

Кроме того, если $C^{\lambda}_{\alpha,W_{[1,n]}} < \infty$, то $q^{\lambda}_{\alpha,W_{[1,n]}} = \bigotimes_{t=1}^{n} q^{\lambda}_{\alpha,W_t}$.

Из аддитивности функции стоимости $\rho_{[1,n]}$ следует, что для любого $p \in \mathcal{P}(\mathcal{X}_1^n)$

$$\mathbf{E}_p[\rho_{[1,n]}] = \sum_{t=1}^{n} \mathbf{E}_{p_t}[\rho_t],$$

где $p_t \in \mathcal{P}(\mathcal{X}_t)$ – маргинальное распределение для p на $\mathcal{X}_t$. Таким образом, из леммы 14 и определения АЛ-информации следует, что

$$I_\alpha^\lambda\bigl(p;W_{[1,n]}\bigr) \leqslant I_\alpha^\lambda\bigl(p_1\otimes\ldots\otimes p_n;W_{[1,n]}\bigr) = \sum_{t=1}^{n} I_\alpha^\lambda(p_t;W_t). \tag{63}$$

Лемма 32 доказывается с помощью соотношения (63) и теоремы 2.

5.3. Информационные величины Реньи – Галлагера. В п. 5.2 была дана характеризация августиновских пропускной способности и центра при наличии ограничений по стоимости через АЛ-пропускную способность и АЛ-центр. АЛ-пропускная способность определяется как супремум АЛ-информации. В [14, формулы (103), (116)] Галлагер (неявно) предложил другую меру информации с множителем Лагранжа. Через супремум этой информации Августин в [6, леммы 35.4(b), 35.8(b)] охарактеризовал августиновскую пропускную способность при наличии ограничений по стоимости в предположении, что функция стоимости ограничена. Мы называем этот супремум РГ-пропускной способностью. Основная цель этого пункта – установить равенство АЛ-пропускной способности и АЛ-центра, соответственно, РГ-пропускной способности и РГ-центру. Мы также выведем границу ван Эрвена – Харремоеса для АЛ-пропускной способности и АЛ-центра и применим ее для доказательства непрерывности АЛ-центра как функции множителя Лагранжа λ.

Определение 16. Для любых $\alpha \in \mathbb{R}_+$, канала $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ с функцией стоимости $\rho\colon \mathcal{X} \to \mathbb{R}^\ell_{\geqslant 0}$, $p \in \mathcal{P}(\mathcal{X})$ и $\lambda \in \mathbb{R}^\ell_{\geqslant 0}$ *информацией Реньи – Галлагера (РГ-информацией) порядка α для распределения на входе p и множителя Лагранжа λ* называется

$$I_\alpha^{\mathrm{g}\lambda}(p;W) \triangleq \begin{cases} \inf\limits_{q\in\mathcal{P}(\mathcal{Y})} D_\alpha\bigl(p \circledast W e^{\frac{1-\alpha}{\alpha}\lambda\cdot\rho} \,\|\, p\otimes q\bigr), & \alpha \in \mathbb{R}_+\setminus\{1\},\\ \inf\limits_{q\in\mathcal{P}(\mathcal{Y})} D_1(p\circledast W \,\|\, p\otimes q) - \lambda\cdot \mathbf{E}_p[\rho], & \alpha = 1. \end{cases} \tag{64}$$

Если λ – нулевой вектор, то РГ-информация совпадает с информацией Реньи. Как и для информации Реньи, для РГ-информации имеется выражение в замкнутом виде через вероятностную меру, доставляющую минимум в ее определении.

Определение 17. Для любых $\alpha \in \mathbb{R}_+$, канала $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ с функцией стоимости $\rho\colon \mathcal{X} \to \mathbb{R}^\ell_{\geqslant 0}$, $p \in \mathcal{P}(\mathcal{X})$ и $\lambda \in \mathbb{R}^\ell_{\geqslant 0}$ *средней мерой порядка α для распределения на входе p и множителя Лагранжа λ* называется

$$\frac{d\mu_{\alpha,p}^\lambda}{d\nu} \triangleq \left[\sum_x p(x) e^{(1-\alpha)\lambda\cdot\rho(x)} \left(\frac{dW(x)}{d\nu}\right)^{\alpha}\right]^{\frac{1}{\alpha}}.$$

Средним Реньи – Галлагера (РГ-средним) порядка α для распределения на входе p и множителя Лагранжа λ называется

$$q_{\alpha,p}^{\mathrm{g}\lambda} \triangleq \frac{\mu_{\alpha,p}^\lambda}{\|\mu_{\alpha,p}^\lambda\|}.$$

Как $\mu_{\alpha,p}^\lambda$, так и $q_{\alpha,p}^{\mathrm{g}\lambda}$ зависят от множителя Лагранжа λ при $\alpha \in \mathbb{R}_+\setminus\{1\}$. Кроме того, непосредственной подстановкой можно убедиться, что

$$\begin{aligned} &D_\alpha\bigl(p\circledast W e^{\frac{1-\alpha}{\alpha}\lambda\cdot\rho} \,\|\, p\otimes q\bigr) = \\ &= D_\alpha\bigl(p\circledast W e^{\frac{1-\alpha}{\alpha}\lambda\cdot\rho} \,\|\, p\otimes q_{\alpha,p}^{\mathrm{g}\lambda}\bigr) + D_\alpha\bigl(q_{\alpha,p}^{\mathrm{g}\lambda}\,\|\, q\bigr), \quad \alpha\in\mathbb{R}_+\setminus\{1\}. \end{aligned} \tag{65}$$

Тогда по лемме 2 получаем

$$I_\alpha^{\mathrm{g}\lambda}(p;W) = D_\alpha\bigl(p \circledast W e^{\frac{1-\alpha}{\alpha}\lambda\cdot\rho} \,\|\, p \otimes q_{\alpha,p}^{\mathrm{g}\lambda}\bigr) = \tag{66}$$

$$= \frac{\alpha}{\alpha-1}\ln\|\mu_{\alpha,p}^\lambda\|, \quad \alpha \in \mathbb{R}_+ \setminus \{1\}. \tag{67}$$

Ни $\mu_{1,p}^\lambda$, ни $q_{1,p}^{\mathrm{g}\lambda}$ не зависят от множителя Лагранжа λ. Кроме того, непосредственной подстановкой можно убедиться, что

$$D_1(p \circledast W \,\|\, p \otimes q) - \lambda \cdot \mathbf{E}_p[\rho] = D_1\bigl(p \circledast W \,\|\, p \otimes q_{1,p}^{\mathrm{g}\lambda}\bigr) - \lambda \cdot \mathbf{E}_p[\rho] + D_1\bigl(q_{1,p}^{\mathrm{g}\lambda} \,\|\, q\bigr). \tag{68}$$

Тогда по лемме 2 получаем

$$I_1^{\mathrm{g}\lambda}(p;W) = D_1\bigl(p \circledast W \,\|\, p \otimes q_{1,p}^{\mathrm{g}\lambda}\bigr) - \lambda \cdot \mathbf{E}_p[\rho]. \tag{69}$$

Используя определения АЛ- и РГ-информации, данные в (51) и (64), с учетом неравенства Йенсена и вогнутости функции натурального логарифма получаем

$$I_\alpha^\lambda(p;W) \geqslant I_\alpha^{\mathrm{g}\lambda}(p;W), \quad \alpha \in (0,1],$$
$$I_\alpha^\lambda(p;W) \leqslant I_\alpha^{\mathrm{g}\lambda}(p;W), \quad \alpha \in [1,\infty).$$

Эти соотношения можно усилить, выражая АЛ- и РГ-информацию друг через друга следующим образом.

Лемма 33. *Пусть W – канал вида $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ с функцией стоимости $\rho\colon \mathcal{X} \to \mathbb{R}_{\geqslant 0}^\ell$, $p \in \mathcal{P}(\mathcal{X})$ – распределение на входе, а $\lambda \in \mathbb{R}_{\geqslant 0}^\ell$ – множитель Лагранжа.*

(a) *Пусть $u_{\alpha,p}^\lambda \in \mathcal{P}(\mathcal{X})$ имеет вид $u_{\alpha,p}^\lambda(x) = \dfrac{p(x)e^{(1-\alpha)D_\alpha(W(x)\,\|\,q_{\alpha,p})+(\alpha-1)\lambda\cdot\rho(x)}}{\sum_{\widetilde{x}} p(\widetilde{x})e^{(1-\alpha)D_\alpha(W(\widetilde{x})\,\|\,q_{\alpha,p})+(\alpha-1)\lambda\cdot\rho(x)}}$ для всех x, тогда*

$$I_\alpha^\lambda(p;W) = I_\alpha^{\mathrm{g}\lambda}(u_{\alpha,p};W) + \frac{1}{\alpha-1}D_1(p \,\|\, u_{\alpha,p}) =$$
$$= \begin{cases} \sup\limits_{u\in\mathcal{P}(\mathcal{X})} I_\alpha^{\mathrm{g}\lambda}(u;W) + \dfrac{1}{\alpha-1}D_1(p\,\|\,u), & \alpha \in (0,1), \\ \inf\limits_{u\in\mathcal{P}(\mathcal{X})} I_\alpha^{\mathrm{g}\lambda}(u;W) + \dfrac{1}{\alpha-1}D_1(p\,\|\,u), & \alpha \in (1,\infty); \end{cases}$$

(b) *Пусть $a_{\alpha,p}^\lambda \in \mathcal{P}(\mathcal{X})$ имеет вид $a_{\alpha,p}^\lambda(x) = \dfrac{p(x)e^{(\alpha-1)D_\alpha\left(W(x)\,\|\,q_{\alpha,p}^{\mathrm{g}\lambda}\right)+(1-\alpha)\lambda\cdot\rho(x)}}{\sum_{\widetilde{x}} p(\widetilde{x})e^{(\alpha-1)D_\alpha\left(W(\widetilde{x})\,\|\,q_{\alpha,p}^{\mathrm{g}\lambda}\right)+(1-\alpha)\lambda\cdot\rho(x)}}$ для всех x, тогда*

$$I_\alpha^{\mathrm{g}\lambda}(p;W) = I_\alpha^\lambda\bigl(a_{\alpha,p}^\lambda;W\bigr) - \frac{1}{\alpha-1}D_1\bigl(a_{\alpha,p}^\lambda \,\|\, p\bigr) =$$
$$= \begin{cases} \inf\limits_{a\in\mathcal{P}(\mathcal{X})} I_\alpha^\lambda(a;W) - \dfrac{1}{\alpha-1}D_1(a\,\|\,p), & \alpha \in (0,1), \\ \sup\limits_{a\in\mathcal{P}(\mathcal{X})} I_\alpha^\lambda(a;W) - \dfrac{1}{\alpha-1}D_1(a\,\|\,p), & \alpha \in (1,\infty); \end{cases}$$

(c) *Пусть* $f^{\lambda}_{\alpha,p}\colon \mathcal{X} \to \mathbb{R}$ *имеет вид* $f^{\lambda}_{\alpha,p}(x) = [D_\alpha(W(x) \,\|\, q_{\alpha,p}) - \lambda \cdot \rho(x) - I^{\lambda}_{\alpha}(p;W)] \times$ $\times\, \mathbb{1}_{p(x)>0}$ *для всех* x, *тогда*

$$I^{\lambda}_{\alpha}(p;W) = \frac{\alpha}{\alpha-1} \ln \mathbf{E}_\nu \left[\left(\sum_x p(x) e^{(1-\alpha)(f^{\lambda}_{\alpha,p}(x)+\lambda\cdot\rho(x))} \left[\frac{dW(x)}{d\nu}\right]^\alpha \right)^{1/\alpha} \right] =$$

$$= \frac{\alpha}{\alpha-1} \ln \inf_{f\colon \mathbf{E}_p[f]=0} \mathbf{E}_\nu \left[\left(\sum_x p(x) e^{(1-\alpha)(f(x)+\lambda\cdot\rho(x))} \left[\frac{dW(x)}{d\nu}\right]^\alpha \right)^{1/\alpha} \right].$$

Лемма 33 при $\lambda = 0$ превращается в лемму 18, что обсуждалось ранее в [6,10,19].

Определение 18. Для любых $\alpha \in \mathbb{R}_+$, канала $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ с функцией стоимости $\rho\colon \mathcal{X} \to \mathbb{R}^{\ell}_{\geqslant 0}$ и $\lambda \in \mathbb{R}^{\ell}_{\geqslant 0}$ *пропускной способностью Реньи – Галлагера* (*РГ-пропускной способностью*) *порядка* α *для множителя Лагранжа* λ называется

$$C^{\mathrm{g}\lambda}_{\alpha,W} \triangleq \sup_{p\in\mathcal{P}(\mathcal{X})} I^{\mathrm{g}\lambda}_{\alpha}(p;W).$$

Используя определение $I^{\mathrm{g}\lambda}_{\alpha}(p;W)$, данное в (64), получаем следующее выражение для $C^{\mathrm{g}\lambda}_{\alpha,W}$:

$$C^{\mathrm{g}\lambda}_{\alpha,W} = \begin{cases} \sup\limits_{p\in\mathcal{P}(\mathcal{X})} \inf\limits_{q\in\mathcal{P}(\mathcal{Y})} D_\alpha\bigl(p \circledast W e^{\frac{1-\alpha}{\alpha}\lambda\cdot\rho} \,\|\, p\otimes q\bigr), & \alpha \in \mathbb{R}_+ \setminus \{1\}, \\ \sup\limits_{p\in\mathcal{P}(\mathcal{X})} \inf\limits_{q\in\mathcal{P}(\mathcal{Y})} D_\alpha(p \circledast W \,\|\, p\otimes q) - \lambda\cdot \mathbf{E}_p[\rho], & \alpha = 1. \end{cases}$$

Для РГ-пропускной способности имеет место теорема о минимаксе, аналогичная теореме о минимаксе для АЛ-пропускной способности, т.е. теореме 2. Поскольку утверждение и доказательство этих теорем для АЛ-пропускной способности порядка 1 и РГ-пропускной способности порядка 1 полностью идентичны, мы сформулируем теорему о минимаксе для РГ-пропускной способности лишь для конечных положительных порядков, отличных от единицы.

Теорема 3. *Для любых* $\alpha \in \mathbb{R}_+ \setminus \{1\}$, *канала* $W\colon \mathcal{X} \to \mathcal{P}(\mathcal{Y})$ *с функцией стоимости* $\rho\colon \mathcal{X} \to \mathbb{R}^{\ell}_{\geqslant 0}$ *и множителя Лагранжа* $\lambda \in \mathbb{R}^{\ell}_{\geqslant 0}$ *справедливы равенства*

$$\sup_{p\in\mathcal{P}(\mathcal{X})} \inf_{q\in\mathcal{P}(\mathcal{Y})} D_\alpha\bigl(p \circledast W e^{\frac{1-\alpha}{\alpha}\lambda\cdot\rho} \,\|\, p\otimes q\bigr) = \tag{70}$$

$$= \inf_{q\in\mathcal{P}(\mathcal{Y})} \sup_{p\in\mathcal{P}(\mathcal{X})} D_\alpha\bigl(p \circledast W e^{\frac{1-\alpha}{\alpha}\lambda\cdot\rho} \,\|\, p\otimes q\bigr) =$$

$$= \inf_{q\in\mathcal{P}(\mathcal{Y})} \sup_{x\in\mathcal{X}} D_\alpha(W(x) \,\|\, q) - \lambda\cdot\rho(x). \tag{71}$$

Если выражение (70) *конечно, т.е. если* $C^{\mathrm{g}\lambda}_{\alpha,W} < \infty$, *то* $\exists!\, q^{\mathrm{g}\lambda}_{\alpha,W} \in \mathcal{P}(\mathcal{Y})$, *называемая центром Реньи – Галлагера порядка* α *канала* W *для множителя Лагранжа* λ, *для которой*

$$C^{\mathrm{g}\lambda}_{\alpha,W} = \sup_{p\in\mathcal{P}(\mathcal{X})} D_\alpha\bigl(p \circledast W e^{\frac{1-\alpha}{\alpha}\lambda\cdot\rho} \,\|\, p\otimes q^{\mathrm{g}\lambda}_{\alpha,W}\bigr) = \tag{72}$$

$$= \sup_{x\in\mathcal{X}} D_\alpha\bigl(W(x) \,\|\, q^{\mathrm{g}\lambda}_{\alpha,W}\bigr) - \lambda\cdot\rho(x). \tag{73}$$

Кроме того, для любой последовательности распределений на входе $\{p^{(i)}\}_{i\in\mathbb{Z}_+} \subset$ $\subset \mathcal{P}(\mathcal{X})$, *такой что* $\lim\limits_{i\to\infty} I^{\mathrm{g}\lambda}_{\alpha}(p^{(i)};W) = C^{\mathrm{g}\lambda}_{\alpha,W}$, *соответствующая последователь-*

ность $\{q^{\mathrm{g}\lambda}_{\alpha,p^{(i)}}\}_{i\in\mathbb{Z}_+}$ *средних Реньи–Галлагера порядка* α *является последовательностью Коши в метрике полной вариации на* $\mathcal{P}(\mathcal{Y})$, *а* $q^{\mathrm{g}\lambda}_{\alpha,W}$ *– единственная предельная точка этой последовательности Коши.*

Доказательство теоремы 3 в значительной степени аналогично доказательствам теорем 1 и 2. Вместо лемм 19 или 30 в нем используется следующая

Лемма 34. *Для любых* $\alpha\in\mathbb{R}_+\setminus\{1\}$, *канала* $W\colon\mathcal{X}\to\mathcal{P}(\mathcal{Y})$ *с функцией стоимости* $\rho\colon\mathcal{X}\to\mathbb{R}^{\ell}_{\geqslant 0}$ *для конечного множества входов* $\mathcal{X}$ *и множителя Лагранжа* $\lambda\in\mathbb{R}^{\ell}_{\geqslant 0}$ *существует* $\widetilde{p}\in\mathcal{P}(\mathcal{X})$, *такое что* $I^{\lambda}_{\alpha}(\widetilde{p};W)=C^{\lambda}_{\alpha,W}$, *и* $\exists!\,q^{\lambda}_{\alpha,W}\in\mathcal{P}(\mathcal{Y})$, *такая что*

$$D_{\alpha}\big(p\circledast We^{\frac{1-\alpha}{\alpha}\lambda\cdot\rho}\,\|\,p\otimes q^{\mathrm{g}\lambda}_{\alpha,W}\big)\leqslant C^{\mathrm{g}\lambda}_{\alpha,W}\quad\forall p\in\mathcal{P}(\mathcal{X}).$$

Кроме того, $q^{\mathrm{g}\lambda}_{\alpha,\widetilde{p}}=q^{\mathrm{g}\lambda}_{\alpha,W}$ *для всех* $\widetilde{p}\in\mathcal{P}(\mathcal{X})$, *таких что* $I^{\mathrm{g}\lambda}_{\alpha}(\widetilde{p};W)=C^{\mathrm{g}\lambda}_{\alpha,W}$.

Выражение (70) задает РГ-пропускную способность, в то время как (71) задает АЛ-радиус, определенный в (62). Таким образом, из теорем 2, 3 следует, что

$$C^{\lambda}_{\alpha,W}=S^{\lambda}_{\alpha,W}=C^{\mathrm{g}\lambda}_{\alpha,W}\quad\forall\alpha\in\mathbb{R}_+,\ \lambda\in\mathbb{R}^{\ell}_{\geqslant 0}.$$

Кроме того, когда $C^{\lambda}_{\alpha,W}$ конечна, единственный АЛ-центр, описанный в (61), также равен единственному РГ-центру, описанному в (73), по теоремам 2 и 3:

$$q^{\lambda}_{\alpha,W}=q^{\mathrm{g}\lambda}_{\alpha,W}\quad\forall\alpha\in\mathbb{R}_+,\ \lambda\in\mathbb{R}^{\ell}_{\geqslant 0},\ \text{такие что } C^{\lambda}_{\alpha,W}<\infty.$$

Чтобы не использовать разные названия для одной и той же величины, всюду далее мы будем формулировать утверждения в терминах АЛ-пропускной способности и АЛ-центра.

Если $C^{\lambda}_{\alpha,W}$ конечна, то из формул (65), (66) и теоремы 3 для $\alpha\in\mathbb{R}_+\setminus\{1\}$, а также формул (68), (69) и теоремы 2 для $\alpha=1$ следует, что

$$C^{\lambda}_{\alpha,W}-I^{\mathrm{g}\lambda}_{\alpha}(p;W)\geqslant D_{\alpha}\big(q^{\mathrm{g}\lambda}_{\alpha,p}\,\|\,q^{\lambda}_{\alpha,W}\big)\quad\forall p\in\mathcal{P}(\mathcal{X}).$$

С помощью этих же рассуждений можно также доказать границу ван Эрвена–Харремоеса для АЛ-пропускной способности.

Лемма 35. *Для любых* $\alpha\in\mathbb{R}_+$, *канала* $W\colon\mathcal{X}\to\mathcal{P}(\mathcal{Y})$ *с функцией стоимости* $\rho\colon\mathcal{X}\to\mathbb{R}^{\ell}_{\geqslant 0}$ *и множителя Лагранжа* $\lambda\in\mathbb{R}^{\ell}_{\geqslant 0}$, *таких что* $C^{\lambda}_{\alpha,W}<\infty$,

$$\sup_{x\in\mathcal{X}}D_{\alpha}(W(x)\,\|\,q)-\lambda\cdot\rho(x)\geqslant C^{\lambda}_{\alpha,W}+D_{\alpha}\big(q^{\lambda}_{\alpha,W}\,\|\,q\big)\quad\forall q\in\mathcal{P}(\mathcal{Y}).$$

Аналогичный, но более слабый результат можно получить с помощью леммы 13 и теоремы 2. Последний член соответствующей границы в этом случае будет равен $D_{\alpha\wedge 1}\big(q^{\lambda}_{\alpha,W}\,\|\,q\big)$ вместо $D_{\alpha}\big(q^{\lambda}_{\alpha,W}\,\|\,q\big)$.

Из леммы 35 и непрерывности АЛ-пропускной способности $C^{\lambda}_{\alpha,W}$ как функции λ, установленной в лемме 29(a), по лемме 2 вытекает непрерывность АЛ-центра $q^{\lambda}_{\alpha,W}$ по λ в топологии полной вариации на $\mathcal{P}(\mathcal{Y})$.

Лемма 36. *Для любых* $\alpha\in\mathbb{R}_+$, *канала* $W\colon\mathcal{X}\to\mathcal{P}(\mathcal{Y})$ *с функцией стоимости* $\rho\colon\mathcal{X}\to\mathbb{R}^{\ell}_{\geqslant 0}$ *и множителя Лагранжа* $\lambda_0\in\mathbb{R}^{\ell}_{\geqslant 0}$, *таких что* $C^{\lambda_0}_{\alpha,W}<\infty$,

$$D_{\alpha}\big(q^{\lambda_2}_{\alpha,W}\,\|\,q^{\lambda_1}_{\alpha,W}\big)\leqslant C^{\lambda_1}_{\alpha,W}-C^{\lambda_2}_{\alpha,W}\quad\forall\lambda_1,\lambda_2\in\mathbb{R}^{\ell}_{\geqslant 0},\ \textit{таких что }\lambda_0\leqslant\lambda_1\leqslant\lambda_2.$$

Кроме того, функция $q^{\lambda}_{\alpha,W}$ *непрерывна по* λ *на множестве* $\{\lambda:\exists\varepsilon>0,$ *такое что* $C^{\lambda-\varepsilon\mathbb{1}}_{\alpha,W}<\infty\}$ *в топологии полной вариации на* $\mathcal{P}(\mathcal{Y})$.

5.4. Информационные величины для функций вероятностей переходов. Мы определили условное расхождение Реньи, информацию Августина, АЛ- и РГ-информацию лишь для распределений на входе, лежащих в $\mathcal{P}(\mathcal{X})$, т.е. для функций вероятности, равных нулю всюду, кроме конечного числа элементов множества $\mathcal{X}$. Однако во многих практически значимых и аналитически интересных моделях множество входов $\mathcal{X}$ является несчетно бесконечным множеством, снабженным σ-алгеброй $\mathcal{X}$. Среди важнейших примеров таких каналов можно назвать гауссовские каналы, возможно, с многими входными и выходными антеннами и замиранием, а также пуассоновские каналы. Для таких каналов зачастую бывает желательно продолжить определения информации Августина и АЛ-информации с $\mathcal{P}(\mathcal{X})$ на $\mathcal{P}(\mathcal{X})$. Например, в аддитивных гауссовских каналах, описанных в примерах 4 и 5, равенство $I_\alpha(p;W) = C_{\alpha,W,\varrho}$ не выполняется для любой функции вероятности p, удовлетворяющей ограничению по стоимости, но выполняется для гауссовского распределения с нулевым средним и дисперсией ϱ.

Далее мы вначале покажем, что если $\mathcal{Y}$ – счетно порожденная σ-алгебра, то определения условного расхождения Реньи, информации Августина и АЛ-информации можно обобщить с $\mathcal{P}(\mathcal{X})$ на $\mathcal{P}(\mathcal{X})$ при условии, что W и Q – не просто функции из $\mathcal{X}$ в $\mathcal{P}(\mathcal{Y})$, а функции вероятностей переходов из $(\mathcal{X},\mathcal{X})$ в $(\mathcal{Y},\mathcal{Y})$. Затем мы покажем, что если при этом $\mathcal{X}$ счетно отделимо, то супремум АЛ-информации $I_\alpha^\lambda(p;W)$ по $\mathcal{P}(\mathcal{X})$ равен АЛ-радиусу $S_{\alpha,W}^\lambda$ (теорема 4). Из этого будет следовать, что августиновская пропускная способность при наличии ограничений по стоимости $C_{\alpha,W,\varrho}$, определенная в (49), также равна супремуму информации Августина $I_\alpha(p;W)$ по элементам $\mathcal{P}(\mathcal{X})$, таким что $\mathbf{E}_p[\rho] \leqslant \varrho$, по крайней мере, для ограничений по стоимости, принадлежащих внутренней области множества всех допустимых ограничений (теорема 5).

Вначале напомним определение функции вероятностей переходов. Мы придерживаемся определения Богачева [21, определение 10.7.1] с небольшой поправкой: используем $W(\mathcal{E}\,|\,x)$ вместо $W(x\,|\,\mathcal{E})$.

Определение 19. Пусть $(\mathcal{X},\mathcal{X})$ и $(\mathcal{Y},\mathcal{Y})$ – измеримые пространства. Тогда функция $W\colon \mathcal{Y}\times\mathcal{X}\to[0,1]$ называется функцией вероятностей переходов (стохастическим ядром, марковским ядром) из $(\mathcal{X},\mathcal{X})$ в $(\mathcal{Y},\mathcal{Y})$, если она удовлетворяет следующим условиям:

(i) Для всех $x\in\mathcal{X}$ функция $W(\cdot\,|\,x)\colon \mathcal{Y}\to[0,1]$ является вероятностной мерой на $(\mathcal{Y},\mathcal{Y})$;

(ii) Для всех $\mathcal{E}\in\mathcal{Y}$ функция $W(\mathcal{E}\,|\,\cdot)\colon \mathcal{X}\to[0,1]$ является $(\mathcal{X},\mathcal{B}([0,1]))$-измеримой.

Обозначим множество всех функций вероятностей переходов из $(\mathcal{X},\mathcal{X})$ в $(\mathcal{Y},\mathcal{Y})$ через $\mathcal{P}(\mathcal{Y}|\mathcal{X})$, подразумевая при этом, что множества $\mathcal{X}$ и $\mathcal{Y}$ будут ясны из контекста. Если W удовлетворяет условию (i), то $W\colon \mathcal{X}\to\mathcal{P}(\mathcal{Y})$ является каналом, т.е. W – элемент $\mathcal{P}(\mathcal{Y}\,|\,\mathcal{X})$, даже если W не удовлетворяет условию (ii). Следовательно, $\mathcal{P}(\mathcal{Y}\,|\,\mathcal{X})\subset\mathcal{P}(\mathcal{Y}\,|\,\mathcal{X})$. В связи с этим наблюдением для удобства будем обозначать вероятностную меру $W(\cdot\,|\,x)$ через $W(x)$, когда это не вызовет разночтений.

Чтобы продолжить определение условного расхождения Реньи с $\mathcal{P}(\mathcal{X})$ на $\mathcal{P}(\mathcal{X})$, убедимся в $\mathcal{X}$-измеримости $D_\alpha(W(x)\,\|\,Q(x))$ на $\mathcal{X}$ и заменим сумму в (15) интегралом. Если $(\mathcal{X},\tau)$ – топологическое пространство, а $\mathcal{X}$ – ассоциированная с ним борелевская σ-алгебра, то для проверки измеримости можно сперва убедиться в непрерывности, каковая имеет место, когда $\dfrac{dW(x)}{d\nu}$ и $\dfrac{dQ(x)}{d\nu}$ непрерывны по x для ν-почти всех y для некоторой вероятностной меры ν, для которой $(W(x)+Q(x))\prec\nu$ для всех $x\in\mathcal{X}$. Иногда это предположение о W и Q бывает нелегко проверить. С другой стороны, если W и Q – функции вероятностей переходов из $(\mathcal{X},\mathcal{X})$ в $(\mathcal{Y},\mathcal{Y})$ для счетно порожденной $\mathcal{Y}$, то требуемая измеримость вытекает из элементарных свойств измеримых функций и леммы 9, как показывает следующая

Лемма 37. *Для любых $\alpha \in \mathbb{R}_+$, счетно порожденной σ-алгебры $\mathcal{Y}$ подмножеств множества $\mathfrak{Y}$ и $W, Q \in \mathcal{P}(\mathcal{Y}\,|\,\mathcal{X})$ функция $D_\alpha(W(\cdot)\,\|\,Q(\cdot))\colon \mathfrak{X} \to [0,\infty]$ является $\mathcal{X}$-измеримой.*

Доказательство. Существует последовательность $\{\mathcal{E}_i\}_{i\in\mathbb{Z}_+} \subset \mathcal{Y}$, такая что $\mathcal{Y} = \sigma(\{\mathcal{E}_i : i \in \mathbb{Z}_+\})$, поскольку $\mathcal{Y}$ – счетно порожденная σ-алгебра. Положим

$$\mathcal{Y}_i \triangleq \sigma(\{\mathcal{E}_1, \ldots, \mathcal{E}_i\}), \quad i \in \mathbb{Z}_+.$$

Тогда $\mathcal{Y}_1 \subset \mathcal{Y}_2 \subset \ldots \subset \mathcal{Y}$, $\mathcal{Y} = \sigma\Bigl(\bigcup\limits_{i=1}^{\infty} \mathcal{Y}_i\Bigr)$, и из леммы 9 следует, что

$$D_\alpha(W(x)\,\|\,Q(x)) = \lim_{i\to\infty} D_\alpha^{\mathcal{Y}_i}(W(x)\,\|\,Q(x)) \quad \forall x \in \mathfrak{X}. \tag{74}$$

С другой стороны, множество $\mathcal{Y}_i$ конечно для любого $i \in \mathbb{Z}_+$. Таким образом, для всех $i \in \mathbb{Z}_+$ существует $\mathcal{Y}_i$-измеримое конечной разбиение $\mathscr{E}_i$ множества $\mathfrak{Y}$. Тогда с учетом определения расхождения Реньи, данного в (8), получаем

$$D_\alpha^{\mathcal{Y}_i}(W(x)\,\|\,Q(x)) = \begin{cases} \dfrac{1}{\alpha-1} \ln \sum\limits_{\mathcal{E}\in\mathscr{E}_i} (W(\mathcal{E}\,|\,x))^\alpha \,(Q(\mathcal{E}\,|\,x))^{1-\alpha}, & \alpha \in \mathbb{R}_+ \setminus \{1\}, \\ \sum\limits_{\mathcal{E}\in\mathscr{E}_i} W(\mathcal{E}\,|\,x) \ln \dfrac{W(\mathcal{E}\,|\,x)}{Q(\mathcal{E}\,|\,x)}, & \alpha = 1. \end{cases}$$

Значит, $D_\alpha^{\mathcal{Y}_i}(W(x)\,\|\,Q(x))$ является $\mathcal{X}$-измеримой функцией для любого $i \in \mathbb{Z}_+$ согласно [21, теорема 2.1.5(i)–(iv) и замечание 2.1.6], поскольку $W(\mathcal{E}\,|\,x)$ и $Q(\mathcal{E}\,|\,x)$ – $\mathcal{X}$-измеримые функции для всех $\mathcal{E} \in \mathscr{E}_i$ по условию леммы. Тогда из равенства (74) следует, что $D_\alpha(W(x)\,\|\,Q(x))$ – $\mathcal{X}$-измеримая функция согласно [21, теорема 2.1.5(v) и замечание 2.1.6].

Определение 20. Для любых $\alpha \in \mathbb{R}_+$, счетно порожденной σ-алгебры $\mathcal{Y}$ подмножеств множества $\mathfrak{Y}$, $W \in \mathcal{P}(\mathcal{Y}\,|\,\mathcal{X})$ и $p \in \mathcal{P}(\mathcal{X})$ *условным расхождением Реньи порядка α для распределения на входе p* называется

$$D_\alpha(W\,\|\,Q\,|\,p) \triangleq \int D_\alpha(W(x)\,\|\,Q(x))\, p(dx). \tag{75}$$

Если $\exists q \in \mathcal{P}(\mathcal{Y})$, такая что p-п.н. $Q(x) = q$, то будем обозначать $D_\alpha(W\,\|\,Q\,|\,p)$ через $D_\alpha(W\,\|\,q\,|\,p)$.

Теперь можно определить информацию Августина и АЛ-информацию для всех $p \in \mathcal{P}(\mathcal{X})$ при условии, что W принадлежит $\mathcal{P}(\mathcal{Y}\,|\,\mathcal{X})$ для счетно порожденной $\mathcal{Y}$, а ρ – $\mathcal{X}$-измеримая функция.

Определение 21. Для любых $\alpha \in \mathbb{R}_+$, счетно порожденной σ-алгебры $\mathcal{Y}$ подмножеств множества $\mathfrak{Y}$, $W \in \mathcal{P}(\mathcal{Y}\,|\,\mathcal{X})$ и $p \in \mathcal{P}(\mathcal{X})$ *информацией Августина порядка α для распределения на входе p* называется

$$I_\alpha(p; W) \triangleq \inf_{q\in\mathcal{P}(\mathcal{Y})} D_\alpha(W\,\|\,q\,|\,p). \tag{76}$$

Кроме того, для любой $\mathcal{X}$-измеримой функции стоимости $\rho\colon \mathfrak{X} \to \mathbb{R}^{\ell}_{\geqslant 0}$ и любого $\lambda \in \mathbb{R}^{\ell}_{\geqslant 0}$ *информацией Августина – Лежандра порядка α для распределения на входе p и множителя Лагранжа λ* называется

$$I_\alpha^\lambda(p; W) \triangleq I_\alpha(p; W) - \lambda \cdot \mathbf{E}_p[\rho], \tag{77}$$

где подразумевается, что если $\lambda \cdot \mathbf{E}_p[\rho] = \infty$, то $I_\alpha^\lambda(p; W) = -\infty$.

Хотя мы и включили случай $\lambda \cdot \mathbf{E}_p[\rho] = \infty$ в формальное определение АЛ-информации, нас будут интересовать только те p, для которых $\lambda \cdot \mathbf{E}_p[\rho]$ конечно. Введем множество $\mathcal{A}^\lambda$ всех таких p:

$$\mathcal{A}^\lambda \triangleq \{p \in \mathcal{P}(\mathcal{X}) :\ \lambda \cdot \mathbf{E}_p[\rho] < \infty\}. \tag{78}$$

Для произвольной σ-алгебры $\mathcal{X}$ одноэлементные множества не обязательно измеримы, поэтому $\mathcal{P}(\mathfrak{X})$ не обязательно является подмножеством $\mathcal{A}^\lambda$. Если $\mathcal{X}$ счетно отделима, то одноэлементные множества лежат в $\mathcal{X}$ согласно [21, теорема 6.5.7], $\mathcal{P}(\mathfrak{X}) \subset \mathcal{A}^\lambda$ и $\sup\limits_{p \in \mathcal{A}^\lambda} I_\alpha^\lambda(p;W) \geqslant C_{\alpha,W}^\lambda$. Противоположное неравенство следует из теоремы 2, поэтому получаем $\sup\limits_{p \in \mathcal{A}^\lambda} I_\alpha^\lambda(p;W) = C_{\alpha,W}^\lambda$. В следующей теореме 4 формально приведены эти наблюдения, а также аналогичные результаты об АЛ-центре с использованием теоремы о минимаксе.

Т е о р е м а 4. *Пусть $\mathcal{X}$ – счетно отделимая σ-алгебра, $\mathcal{Y}$ – счетно порожденная σ-алгебра, W – функция вероятностей переходов из $(\mathfrak{X}, \mathcal{X})$ в $(\mathfrak{Y}, \mathcal{Y})$, $\rho\colon \mathfrak{X} \to \mathbb{R}_{\geqslant 0}^\ell$ – $\mathcal{X}$-измеримая функция стоимости и $\alpha \in \mathbb{R}_+$. Тогда для всех $\lambda \in \mathbb{R}_{\geqslant 0}^\ell$ справедливы равенства*

$$\sup_{p \in \mathcal{A}^\lambda} \inf_{q \in \mathcal{P}(\mathcal{Y})} D_\alpha(W \,\|\, q \,|\, p) - \lambda \cdot \mathbf{E}_p[\rho] = \inf_{q \in \mathcal{P}(\mathcal{Y})} \sup_{p \in \mathcal{A}^\lambda} D_\alpha(W \,\|\, q \,|\, p) - \lambda \cdot \mathbf{E}_p[\rho] = \tag{79}$$

$$= \inf_{q \in \mathcal{P}(\mathcal{Y})} \sup_{x \in \mathfrak{X}} D_\alpha(W(x) \,\|\, q) - \lambda \cdot \rho(x) = \tag{80}$$

$$= C_{\alpha,W}^\lambda, \tag{81}$$

где множество $\mathcal{A}^\lambda$ определено в (78). Если $C_{\alpha,W}^\lambda$ конечна, то $\exists!\, q_{\alpha,W}^\lambda \in \mathcal{P}(\mathcal{Y})$, называемая центром Августина – Лежандра порядка α функции вероятности переходов W для множителя Лагранжа λ, такая что

$$C_{\alpha,W}^\lambda = \sup_{p \in \mathcal{A}^\lambda} D_\alpha\big(W \,\|\, q_{\alpha,W}^\lambda \,|\, p\big) - \lambda \cdot \mathbf{E}_p[\rho] = \tag{82}$$

$$= \sup_{x \in \mathfrak{X}} D_\alpha\big(W(x) \,\|\, q_{\alpha,W}^\lambda\big) - \lambda \cdot \rho(x). \tag{83}$$

Д о к а з а т е л ь с т в о. Так как $\mathcal{P}(\mathfrak{X}) \subset \mathcal{A}^\lambda$, то из неравенства максимина-минимакса следует, что

$$\sup_{p \in \mathcal{P}(\mathfrak{X})} \inf_{q \in \mathcal{P}(\mathcal{Y})} D_\alpha(W \,\|\, q \,|\, p) - \lambda \cdot \mathbf{E}_p[\rho] \leqslant \sup_{p \in \mathcal{A}^\lambda} \inf_{q \in \mathcal{P}(\mathcal{Y})} D_\alpha(W \,\|\, q \,|\, p) - \lambda \cdot \mathbf{E}_p[\rho] \leqslant$$
$$\leqslant \inf_{q \in \mathcal{P}(\mathcal{Y})} \sup_{p \in \mathcal{A}^\lambda} D_\alpha(W \,\|\, q \,|\, p) - \lambda \cdot \mathbf{E}_p[\rho] = \inf_{q \in \mathcal{P}(\mathcal{Y})} \sup_{x \in \mathfrak{X}} D_\alpha(W(x) \,\|\, q) - \lambda \cdot \rho(x).$$

Таким образом, равенства (79), (80) следуют из равенств (58), (59) теоремы 2, а (81) – из равенств (59) и (57).

Если $C_{\alpha,W}^\lambda$ конечна, то по теореме 2 существует единственная $q_{\alpha,W}^\lambda \in \mathcal{P}(\mathcal{Y})$, такая что

$$\sup_{x \in \mathfrak{X}} D_\alpha\big(W(x) \,\|\, q_{\alpha,W}^\lambda\big) - \lambda \cdot \rho(x) = C_{\alpha,W}^\lambda.$$

Теперь равенства (82), (83) вытекают из того факта, что для любой $q \in \mathcal{P}(\mathcal{Y})$

$$\sup_{p \in \mathcal{A}^\lambda} D_\alpha(W \,\|\, q \,|\, p) - \lambda \cdot \mathbf{E}_p[\rho] = \sup_{x \in \mathfrak{X}} D_\alpha(W(x) \,\|\, q) - \lambda \cdot \rho(x).$$

Пусть $\mathcal{A}(\varrho)$ – подмножество $\mathcal{P}(\mathcal{X})$, состоящее из вероятностных мер, удовлетворяющих ограничению по стоимости ϱ:

$$\mathcal{A}(\varrho) \triangleq \{p \in \mathcal{P}(\mathcal{X}) :\ \mathbf{E}_p[\rho] \leqslant \varrho\}.$$

Тогда $\mathscr{A}(\varrho) \subset \mathcal{A}(\varrho)$, и $\sup\limits_{p \in \mathcal{A}(\varrho)} I_\alpha(p; W) \geqslant C_{\alpha,W,\varrho}$, если $\mathscr{X}$ счетно отделима. Для ограничений по стоимости, принадлежащих $\operatorname{int}\Gamma_\rho$, противоположное неравенство имеет место в силу леммы 29(с) и теоремы 4, а значит, $\sup\limits_{p \in \mathscr{A}(\varrho)} I_\alpha(p; W) = C_{\alpha,W,\varrho}$. В следующей теореме 5 формально приведены эти наблюдения, а также аналогичные результаты об августиновском центре с использованием теоремы о минимаксе.

Т е о р е м а 5. *Пусть* $\mathcal{X}$ *– счетно отделимая* σ*-алгебра,* $\mathcal{Y}$ *– счетно порожденная* σ*-алгебра,* W *– функция вероятностей переходов из* $(\mathscr{X}, \mathcal{X})$ *в* $(\mathscr{Y}, \mathcal{Y})$, $\rho\colon \mathscr{X} \to \mathbb{R}^{\ell}_{\geqslant 0}$ *–* $\mathcal{X}$*-измеримая функция стоимости и* $\alpha \in \mathbb{R}_+$. *Для любого* $\varrho \in \operatorname{int}\Gamma_\rho$ *справедливы равенства*

$$\sup_{p \in \mathscr{A}(\varrho)} \inf_{q \in \mathcal{P}(\mathcal{Y})} D_\alpha(W \,\|\, q \,|\, p) = \inf_{q \in \mathcal{P}(\mathcal{Y})} \sup_{p \in \mathcal{A}(\varrho)} D_\alpha(W \,\|\, q \,|\, p) = \tag{84}$$

$$= C_{\alpha,W,\varrho}, \tag{85}$$

где $C_{\alpha,W,\varrho}$ *определена в* (49). *Если* $C_{\alpha,W,\varrho} \in \mathbb{R}_{\geqslant 0}$, *то* $\exists!\, q_{\alpha,W,\varrho} \in \mathcal{P}(\mathcal{Y})$, *называемая августиновским центром порядка* α *функции вероятностей переходов* W *для ограничения по стоимости* ϱ, *такая что*

$$C_{\alpha,W,\varrho} = \sup_{p \in \mathcal{A}(\varrho)} D_\alpha(W \,\|\, q_{\alpha,W,\varrho} \,|\, p) = \tag{86}$$

$$= \sup_{p \in \mathscr{A}(\varrho)} D_\alpha(W \,\|\, q_{\alpha,W,\varrho} \,|\, p). \tag{87}$$

Кроме того, $q_{\alpha,W,\varrho} = q^{\lambda}_{\alpha,W}$ *для всех* $\lambda \in \mathbb{R}^{\ell}_{\geqslant 0}$, *таких что* $C_{\alpha,W,\varrho} = C^{\lambda}_{\alpha,W} + \lambda \cdot \varrho$.

Д о к а з а т е л ь с т в о. Так как $\mathscr{A}(\varrho) \subset \mathcal{A}(\varrho)$, то из неравенства максимина-минимакса следует, что

$$\sup_{p \in \mathscr{A}(\varrho)} \inf_{q \in \mathcal{P}(\mathcal{Y})} D_\alpha(W \,\|\, q \,|\, p) \leqslant \sup_{p \in \mathcal{A}(\varrho)} \inf_{q \in \mathcal{P}(\mathcal{Y})} D_\alpha(W \,\|\, q \,|\, p) \leqslant$$

$$\leqslant \inf_{q \in \mathcal{P}(\mathcal{Y})} \sup_{p \in \mathcal{A}(\varrho)} D_\alpha(W \,\|\, q \,|\, p).$$

Таким образом, при $C_{\alpha,W,\varrho} = \infty$ равенства (84), (85) справедливы согласно (42). С другой стороны, по теореме 4 для любого λ с конечной $C^{\lambda}_{\alpha,W}$ существует единственная $q^{\lambda}_{\alpha,W}$, удовлетворяющая (83). Следовательно,

$$\inf_{q \in \mathcal{P}(\mathcal{Y})} \sup_{p \in \mathcal{A}(\varrho)} D_\alpha(W \,\|\, q \,|\, p) \leqslant \sup_{p \in \mathcal{A}(\varrho)} D_\alpha\big(W \,\|\, q^{\lambda}_{\alpha,W} \,|\, p\big) \leqslant$$

$$\leqslant \sup_{p \in \mathcal{A}(\varrho)} D_\alpha\big(W \,\|\, q^{\lambda}_{\alpha,W} \,|\, p\big) - \lambda \cdot \mathbf{E}_p[\rho] + \lambda \cdot \varrho \leqslant C^{\lambda}_{\alpha,W} + \lambda \cdot \varrho.$$

Кроме того, если $C_{\alpha,W,\varrho} \in \mathbb{R}$, то по лемме 29(с) существует по крайней мере один $\lambda \in \mathbb{R}^{\ell}_{\geqslant 0}$, такой что $C_{\alpha,W,\varrho} = C^{\lambda}_{\alpha,W} + \lambda \cdot \varrho$. Поэтому равенства (84), (85) выполнены при $C_{\alpha,W,\varrho} \in \mathbb{R}$, а (86) выполнено для $q_{\alpha,W,\varrho} = q^{\lambda}_{\alpha,W}$ при условии, что $C_{\alpha,W,\varrho} = C^{\lambda}_{\alpha,W} + \lambda \cdot \varrho$. С другой стороны, $q_{\alpha,W,\varrho}$ – вероятностная мера, удовлетворяющая (87) по теореме 1, и по лемме 31 $q_{\alpha,W,\varrho} = q^{\lambda}_{\alpha,W}$ для всех λ, таких что $C_{\alpha,W,\varrho} = C^{\lambda}_{\alpha,W} + \lambda \cdot \varrho$.

Счетная отделимость $\mathcal{X}$ и счетная порожденность $\mathcal{Y}$ – довольно слабые условия, которым удовлетворяют большинство функций вероятностей переходов, рассматриваемых на практике. Поэтому теоремы 4, 5 являются дополнительной мотивацией для того, чтобы первым делом изучать этот относительно несложный случай функций вероятности.

Согласно лемме 29(c), (d) и теореме 5 существование распределения на входе p, удовлетворяющего условиям $\mathbf{E}_p[\rho] \leqslant \varrho$ и $I_\alpha(p;W) = C_{\alpha,W,\varrho}$, несущественно для существования и единственности меры $q_{\alpha,W,\varrho}$ или ее характеризации через $q^\lambda_{\alpha,W}$ для λ, удовлетворяющих условию $C_{\alpha,W,\varrho} = C^\lambda_{\alpha,W} + \lambda \cdot \varrho$. Хотя в некоторых специальных случаях можно доказать существование такого p, в общем случае такого распределения на входе не существует. Поэтому мы считаем правильным отделять вопрос о существовании оптимального распределения от обсуждения $C_{\alpha,W,\varrho}$ и $q_{\alpha,W,\varrho}$ и их характеризации через $C^\lambda_{\alpha,W}$ и $q^\lambda_{\alpha,W}$, что, однако, не является общепринятой практикой [37, теорема 1].

Мы предполагали $\mathcal{Y}$ счетно порожденной, чтобы гарантировать корректность определения условного расхождения Реньи, использованного в (76). Однако чтобы определить информацию Реньи, не обязательно предполагать, что $\mathcal{Y}$ счетно порождена, достаточно структуры функции вероятностей переходов. Напомним, что если $W \in \mathcal{P}(\mathcal{Y}\,|\,\mathcal{X})$, то согласно [21, теорема 10.7.2] для любого $p \in \mathcal{P}(\mathcal{X})$ существует единственная вероятностная мера $p \circledast W$ на $(\mathfrak{X} \times \mathfrak{Y}, \mathcal{X} \otimes \mathcal{Y})$, такая что

$$p \circledast W(\mathcal{E}_x \times \mathcal{E}_y) = \int\limits_{\mathcal{E}_x} W(\mathcal{E}_y\,|\,x)\,p(dx) \quad \forall \mathcal{E}_x \in \mathcal{X},\ \mathcal{E}_y \in \mathcal{Y}.$$

Таким образом, $I^{\mathrm{g}}_\alpha(p;W)$ корректно определена для любых $W \in \mathcal{P}(\mathcal{Y}\,|\,\mathcal{X})$ и $p \in \mathcal{P}(\mathcal{X})$.

К сожалению, для РГ-информации ситуация далеко не столь проста. Чтобы определить РГ-информацию, используя аналогичный подход, вначале надо показать, что $We^{\frac{1-\alpha}{\alpha}\lambda\cdot\rho}$ является переходным ядром (а не функцией вероятностей переходов, т.е. марковским ядром), а затем установить существование и единственность меры $p \circledast We^{\frac{1-\alpha}{\alpha}\lambda\cdot\rho}$ для всех $p \in \mathcal{P}(\mathcal{Y})$. Для порядков, больших единицы, полученная мера является субвероятностной, и в качестве определения РГ-информации можно использовать (64). С другой стороны, для порядков между нулем и единицей мера $p \circledast We^{\frac{1-\alpha}{\alpha}\lambda\cdot\rho}$ является σ-конечной для всех $p \in \mathcal{P}(\mathcal{X})$, но не обязательно конечной для всех $p \in \mathcal{P}(\mathcal{X})$. Таким образом, для порядков между нулем и единицей можно использовать (64) как определение РГ-информации только после того как продолжить определение расхождения Реньи на σ-конечные меры.

§ 6. Примеры

Здесь мы сперва продемонстрируем некоторые тонкости, указанные в предыдущих параграфах, а затем рассмотрим гауссовские каналы и получим выражения в замкнутом виде для их августиновских пропускной способности и центра.

6.1. Семейства, инвариантные относительно сдвига.

Пример 1 (канал с афинной пропускной способностью). Рассмотрим канал $W\colon \mathbb{R}_{\geqslant 0} \to \mathcal{P}(\mathcal{B}([0,1)))$ и связанную с ним функцию стоимости $\rho\colon \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\geqslant 0}$, имеющие вид

$$\frac{dW(x)}{d\nu} = f_{\lfloor x \rfloor}(y - x - \lfloor y - x \rfloor),$$
$$\rho(x) = \lfloor x \rfloor,$$

где ν – мера Лебега на $[0,1)$, а функции f_i имеют вид

$$f_i(y) = e^{i+1}\mathbb{1}_{y \in [0, e^{-i-1})} \quad \forall i \in \mathbb{Z}_{\geqslant 0}.$$

Пусть u_i – равномерное распределение на $[i, i+1)$, тогда непосредственной подстановкой можно убедиться, что $\mathrm{T}_{\alpha,u_i}(u_0) = u_0$. Тогда, используя неравенство Йенсена и свойство неподвижности точки, получаем[22]

$$D_\alpha(W \,\|\, q \,|\, u_i) \geqslant D_\alpha(W \,\|\, u_0 \,|\, u_i) + D_{\alpha\wedge 1}(u_0 \,\|\, q).$$

Таким образом, u_0 – единственное августиновское среднее порядка α для распределения на входе u_i, т.е. $q_{\alpha,u_i} = u_0$ и $I_\alpha(u_i; W) = D_\alpha(W \,\|\, u_0 \,|\, u_i)$, а значит, $I_\alpha(u_i; W) = i + 1$, для всех $i \in \mathbb{Z}_+$ и $\alpha \in \mathbb{R}_+$. Далее, используя тот факт, что $\mathbf{E}_{u_i}[\rho] = i$, заключаем, что $C_{\alpha,W,\varrho} \geqslant (\varrho + 1)$ не только для $\varrho \in \mathbb{Z}_{\geqslant 0}$, но и для $\varrho \in \mathbb{R}_{\geqslant 0}$, поскольку функция $C_{\alpha,W,\varrho}$ вогнута по ϱ согласно лемме 27(a). С другой стороны, непосредственной подстановкой можно убедиться, что

$$D_\alpha(W \,\|\, u_0 \,|\, p) = \mathbf{E}_p[\rho] + 1. \tag{88}$$

Итак, $I_\alpha(p; W) \leqslant (\varrho + 1)$ для любого p, удовлетворяющего ограничению по стоимости ϱ. Следовательно,

$$C_{\alpha,W,\varrho} = \varrho + 1,$$
$$q_{\alpha,W,\varrho} = u_0.$$

Поэтому в силу (55) получаем

$$C^\lambda_{\alpha,W} = \begin{cases} \infty, & \lambda \in [0,1), \\ 1, & \lambda \in [1,\infty). \end{cases}$$

Теперь, используя (88) и теорему 4, заключаем, что $q^\lambda_{\alpha,W} = u_0$ для всех $\lambda \in [1,\infty)$.

Пример 2 (канал с неполунепрерывной сверху пропускной способностью). Зададим канал $W\colon \mathbb{R} \to \mathcal{P}(\mathcal{B}([0,1)))$ и функцию стоимости $\rho\colon \mathbb{R} \to \mathbb{R}_{\geqslant 0}$, связанную с ним, как

$$\frac{dW(x)}{d\nu} = f_{\lfloor x \rfloor}(y - x - \lfloor y - x \rfloor),$$

$$\rho(x) = \begin{cases} \lfloor x \rfloor, & x \geqslant 0, \\ 2^{\lfloor x \rfloor}, & x < 0, \end{cases}$$

где ν – мера Лебега на $[0,1)$, а функции $f_i\colon [0,1) \to \mathbb{R}_{\geqslant 0}$ имеют вид

$$f_i(y) = \begin{cases} 2^{i+1}\mathbb{1}_{y \in [0, 2^{-i-1})}, & i > 0, \\ 3/2\mathbb{1}_{y \in [0,2/3)}, & i = 0, \\ 2\mathbb{1}_{y \in [0,1/2)}, & i < 0. \end{cases}$$

Рассуждениями, аналогичными приведенным выше, получаем, что

$$C_{\alpha,W,\varrho} = \begin{cases} (\varrho + 1)\ln 2, & \varrho > 0, \\ \ln 3/2, & \varrho = 0, \end{cases}$$

$$C^\lambda_{\alpha,W} = \begin{cases} \infty, & \lambda \in [0, \ln 2), \\ \ln 2, & \lambda \in [\ln 2, \infty). \end{cases}$$

Следовательно, $C_{\alpha,W,\varrho} \neq \inf\limits_{\lambda \geqslant 0} C^\lambda_{\alpha,W} + \lambda \cdot \varrho$ для $\varrho = 0$.

Пример 3 (канал-произведение без оптимального распределения стоимости). Пусть W_1 и W_2 – каналы, описанные в примерах 1 и 2, а ρ_1 и ρ_2 – связанные с ними

[22] См. вывод формул (23), (24) леммы 13(c), (d), приведенный в [22, Приложение B].

функции стоимости. Пусть $W_{[1,2]}$ – произведение этих двух каналов с аддитивной функцией стоимости $\varrho_{[1,2]}$, т.е.

$$W_{[1,2]}(x_1, x_2) = W_1(x_1) \otimes W_2(x_2),$$
$$\rho_{[1,2]}(x_1, x_2) = \rho_1(x_1) + \rho_2(x_2).$$

Тогда из леммы 28 получаем

$$C_{\alpha,W_{[1,2]},\varrho} = \begin{cases} \varrho + 1 + \ln 2, & \varrho > 0, \\ 1 + \ln \dfrac{3}{2}, & \varrho = 0. \end{cases}$$

Заметим, что для положительных значений ϱ не существует пары (ϱ_1, ϱ_2), одновременно удовлетворяющей условию $C_{\alpha,W_{[1,2]},\varrho} = C_{\alpha,W_1,\varrho_1} + C_{\alpha,W_2,\varrho_2}$ и ограничению по стоимости $\varrho_1 + \varrho_2 \leqslant \varrho$.

6.2. Гауссовские каналы. В дальнейшем будем обозначать гауссовскую вероятностную меру на $\mathcal{B}(\mathbb{R})$ с нулевым средним и дисперсией σ^2 через φ_{σ^2}. Допуская некоторую вольность обозначений, соответствующую функцию плотности вероятности будем обозначать тем же символом:

$$\varphi_{\sigma^2}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \quad \forall x \in \mathbb{R}.$$

Гауссовские каналы и соответствующие функции вероятностей переходов будем использовать взаимозаменяемым образом; по теоремам 4, 5 они имеют одинаковые августиновскую пропускную способность и центр при наличии ограничений по стоимости.

Пример 4 (скалярный гауссовский канал). Пусть W – скалярный гауссовский канал с дисперсией шума σ^2, и пусть связанная с ним функция стоимости ρ квадратична, т.е.

$$W(\mathcal{E} \,|\, x) = \int\limits_{\mathcal{E}} \varphi_{\sigma^2}(y - x)\, dy \quad \forall \mathcal{E} \in \mathcal{B}(\mathbb{R}),$$
$$\rho(x) = x^2 \quad \forall x \in \mathbb{R}.$$

Августиновские пропускная способность и центр этого канала имеют следующий вид:

$$C_{\alpha,W,\varrho} = \\ = \begin{cases} \dfrac{\alpha\varrho}{2(\alpha\theta_{\alpha,\sigma,\varrho} + (1-\alpha)\sigma^2)} + \dfrac{1}{\alpha - 1} \ln \dfrac{(\theta_{\alpha,\sigma,\varrho})^{\alpha/2}\sigma^{(1-\alpha)}}{\sqrt{\alpha\theta_{\alpha,\sigma,\varrho} + (1-\alpha)\sigma^2}}, & \alpha \in \mathbb{R}_+ \setminus \{1\}, \\ \dfrac{1}{2} \ln\left(1 + \dfrac{\varrho}{\sigma^2}\right), & \alpha = 1, \end{cases} \tag{89}$$

$$q_{\alpha,W,\varrho} = \varphi_{\theta_{\alpha,\sigma,\varrho}}, \tag{90}$$

$$\theta_{\alpha,\sigma,\varrho} \triangleq \sigma^2 + \frac{\varrho}{2} - \frac{\sigma^2}{2\alpha} + \sqrt{(\frac{\varrho}{2} - \frac{\sigma^2}{2\alpha})^2 + \varrho\sigma^2}. \tag{91}$$

Кроме того, $C_{\alpha,W,\varrho}$ непрерывно дифференцируема по ϱ, и ее производная – непрерывная убывающая по ϱ биективная функция из $\mathbb{R}_+$ в $[0, \alpha/2\sigma^2)$, имеющая вид

$$\frac{d}{d\varrho} C_{\alpha,W,\varrho} = \frac{\alpha}{2(\alpha\theta_{\alpha,\sigma,\varrho} + (1-\alpha)\sigma^2)} = \tag{92}$$

$$= \frac{\alpha}{\alpha\varrho + \sigma^2 + \sqrt{(\alpha\varrho - \sigma^2)^2 + 4\varrho\alpha^2\sigma^2}}. \tag{93}$$

Для доказательства этих фактов вначале покажем, что августиновское среднее для гауссовского распределения с нулевым средним и дисперсией ϱ является гауссовским распределением с нулевым средним и дисперсией $\theta_{\alpha,\sigma,\varrho}$, т.е. $q_{\alpha,\varphi_\varrho} = \varphi_{\theta_{\alpha,\sigma,\varrho}}$. Отсюда будет следовать, что $I_\alpha(\varphi_\varrho; W) = D_\alpha\big(W \,\|\, \varphi_{\theta_{\alpha,\sigma,\varrho}} \,|\, \varphi_\varrho\big)$. Величина $D_\alpha\big(W \,\|\, \varphi_{\theta_{\alpha,\sigma,\varrho}} \,|\, \varphi_\varrho\big)$ равна выражению в правой части равенства (89). Для доказательства соотношений (89), (90) покажем, что эта величина является наибольшим значением информации Августина по всем распределениям на входе, удовлетворяющим ограничению по стоимости ϱ. Следовательно, мы получим $C_{\alpha,W,\varrho} = I_\alpha(\varphi_\varrho; W)$ и $q_{\alpha,W,\varrho} = q_{\alpha,\varphi_\varrho}$. Затем проверим (92), используя некоторое тождество, а именно (96), полученное при выводе равенства $q_{\alpha,\varphi_\varrho} = \varphi_{\theta_{\alpha,\sigma,\varrho}}$.

Непосредственной подстановкой можно убедиться, что

$$D_\alpha(W(x) \,\|\, \varphi_\theta) = \\ = \begin{cases} \dfrac{\alpha x^2}{2(\alpha\theta + (1-\alpha)\sigma^2)} + \dfrac{1}{\alpha - 1} \ln \dfrac{\theta^{\alpha/2}\sigma^{(1-\alpha)}}{\sqrt{\alpha\theta + (1-\alpha)\sigma^2}}, & \alpha \in \mathbb{R}_+ \setminus \{1\}, \\ \dfrac{\sigma^2 + x^2 - \theta}{2\theta} + \dfrac{1}{2} \ln \dfrac{\theta}{\sigma^2}, & \alpha = 1. \end{cases} \tag{94}$$

Тогда скошенный канал $W_\alpha^{\varphi_\theta}$ порядка α, определенный в (16), также является гауссовским:

$$W_\alpha^{\varphi_\theta}(\mathcal{E} \,|\, x) = \int\limits_{\mathcal{E}} \varphi_{\frac{\sigma^2\theta}{\alpha\theta + (1-\alpha)\sigma^2}} \left(y - \frac{\alpha\theta}{\alpha\theta + (1-\alpha)\sigma^2} x \right) dy.$$

Значит, $\mathrm{T}_{\alpha,p}(q)$ является гауссовской вероятностной мерой с нулевым средним, если p и q таковы. В частности,

$$\mathrm{T}_{\alpha,\varphi_\varrho}(\varphi_\theta) = \varphi_{\left(\frac{\alpha\theta}{\alpha\theta + (1-\alpha)\sigma^2}\right)^2 \varrho + \frac{\sigma^2\theta}{\alpha\theta + (1-\alpha)\sigma^2}}. \tag{95}$$

Следовательно, если φ_θ – неподвижная точка оператора $\mathrm{T}_{\alpha,\varphi_\varrho}(\cdot)$, то θ удовлетворяет следующему равенству:

$$\theta\left[\theta^2 - \theta\Big(\varrho + \Big(2 - \frac{1}{\alpha}\Big)\sigma^2\Big) + \Big(1 - \frac{1}{\alpha}\Big)\sigma^4\right] = 0. \tag{96}$$

Величина $\theta_{\alpha,\sigma,\varrho}$, определенная в (91), является единственным корнем уравнения (96), превосходящим σ^2, при $\alpha \in \mathbb{R}_+$, а также единственным положительным корнем при $\alpha \in (0,1)$. Кроме того, с помощью (95) можно убедиться, что $\mathrm{T}_{\alpha,\varphi_\varrho}(\varphi_{\theta^2_{\alpha,\sigma,\varrho}}) = \varphi_{\theta_{\alpha,\sigma,\varrho}}$, т.е. $\varphi_{\theta_{\alpha,\sigma,\varrho}}$ является неподвижной точкой $\mathrm{T}_{\alpha,\varphi_\varrho}(\cdot)$. Тогда, используя неравенство Йенсена и свойство неподвижности точки, получаем[23]

$$D_\alpha(W \,\|\, q \,|\, \varphi_\varrho) \geqslant D_\alpha\big(W \,\|\, \varphi_{\theta_{\alpha,\sigma,\varrho}} \,|\, \varphi_\varrho\big) + D_{1\wedge\alpha}(\varphi_{\theta_{\alpha,\sigma,\varrho}} \,\|\, q) \quad \forall q \in \mathcal{P}(\mathcal{B}(\mathbb{R})).$$

Таким образом, $\varphi_{\theta_{\alpha,\sigma,\varrho}}$ является августиновским средним порядка α для распределения на входе φ_ϱ, т.е. $q_{\alpha,\varphi_\varrho} = \varphi_{\theta_{\alpha,\sigma,\varrho}}$ и $I_\alpha(\varphi_\varrho; W) = D_\alpha\big(W \,\|\, \varphi_{\theta_{\alpha,\sigma,\varrho}} \,|\, \varphi_\varrho\big)$. С другой

[23] Вывод этого неравенства аналогичен выводу соотношений (23), (24) леммы 13(c), (d), приведенному в [22, Приложение B].

стороны, из (94) следует, что

$$D_\alpha\big(W \,\|\, \varphi_{\theta_{\alpha,\sigma,\varrho}} \,|\, p\big) = \frac{\alpha(\mathbf{E}_p[\rho]-\varrho)}{2(\alpha\theta_{\alpha,\sigma,\varrho}+(1-\alpha)\sigma^2)} + I_\alpha(\varphi_\varrho; W) \quad \forall p \in \mathcal{P}(\mathcal{B}(\mathbb{R})). \tag{97}$$

Тогда $I_\alpha(p;W) \leqslant I_\alpha(\varphi_\varrho;W)$ для всех p, таких что $\mathbf{E}_p[\rho] \leqslant \varrho$. Следовательно, $C_{\alpha,W,\varrho} = I_\alpha(\varphi_\varrho;W)$ и $q_{\alpha,W,\varrho} = q_{\alpha,\varphi_\varrho}$.

Для случая $\alpha = 1$ равенство (92) очевидно. Чтобы доказать (92) в случае $\alpha \in \mathbb{R}_+ \setminus \{1\}$, заметим, что

$$\begin{aligned}
\frac{d}{d\varrho}C_{\alpha,W,\varrho} &= \frac{\alpha}{2(\alpha\theta_{\alpha,\sigma,\varrho}+(1-\alpha)\sigma^2)} + \\
&+ \left[\frac{-\alpha^2\varrho}{2(\alpha\theta_{\alpha,\sigma,\varrho}+(1-\alpha)\sigma^2)^2} + \frac{\alpha(\theta_{\alpha,\sigma,\varrho}-\sigma^2)}{2(\alpha\theta_{\alpha,\sigma,\varrho}+(1-\alpha)\sigma^2)\theta_{\alpha,\sigma,\varrho}}\right]\frac{d}{d\varrho}\theta_{\alpha,\sigma,\varrho} = \\
&= \frac{\alpha}{2(\alpha\theta_{\alpha,\sigma,\varrho}+(1-\alpha)\sigma^2)} + \frac{\alpha^2}{2(\alpha\theta_{\alpha,\sigma,\varrho}+(1-\alpha)\sigma^2)^2\theta_{\alpha,\sigma,\varrho}} \times \\
&\times \left[\theta^2_{\alpha,\sigma,\varrho} - \theta_{\alpha,\sigma,\varrho}\Big(\varrho + \Big(2-\frac{1}{\alpha}\Big)\sigma^2\Big) + \Big(1-\frac{1}{\alpha}\Big)\sigma^4\right]\frac{d}{d\varrho}\theta_{\alpha,\sigma,\varrho}.
\end{aligned}$$

Тогда (92) выполнено для $\alpha \in \mathbb{R}_+ \setminus \{1\}$, поскольку $\theta_{\alpha,\sigma,\varrho}$ – корень уравнения (96).

АЛ-пропускная способность и АЛ-центр этого канала имеют следующий вид:

$$C^\lambda_{\alpha,W} = \begin{cases} \left(\dfrac{\alpha}{\alpha-1}\ln\sqrt{\dfrac{1}{\alpha}+\dfrac{\alpha-1}{\alpha}\dfrac{2\sigma^2\lambda}{\alpha}} - \ln\sqrt{\dfrac{2\sigma^2\lambda}{\alpha}}\right)\mathbb{1}_{\lambda\in(0,\frac{\alpha}{2\sigma^2})}, & \alpha \in \mathbb{R}_+\setminus\{1\}, \\ \left(\sigma^2\lambda - \ln\sqrt{2e\sigma^2\lambda}\right)\mathbb{1}_{\lambda\in(0,\frac{1}{2\sigma^2})}, & \alpha = 1, \end{cases} \tag{98}$$

$$q^\lambda_{\alpha,W} = \varphi_{\theta^\lambda_{\alpha,\sigma}} \tag{99}$$

$$\theta^\lambda_{\alpha,\sigma} \triangleq \sigma^2 + \left|\frac{1}{2\lambda} - \frac{\sigma^2}{\alpha}\right|^+ \tag{100}$$

Тогда $C^\lambda_{\alpha,W}$ непрерывно дифференцируема по λ, и ее производная – непрерывная возрастающая по λ биективная функция из $\mathbb{R}_+$ в $(-\infty, 0]$, имеющая вид

$$\frac{d}{d\lambda}C^\lambda_{\alpha,W} = -\frac{\alpha - 2\sigma^2\lambda}{2\lambda(\alpha+(\alpha-1)2\sigma^2\lambda)}\mathbb{1}_{\lambda\leqslant\frac{\alpha}{2\sigma^2}}. \tag{101}$$

Выражения (98), (99) для АЛ-пропускной способности и АЛ-центра выводятся с помощью выражений для августиновских пропускной способности и центра, соотношений (55), (92)–(94) и леммы 31.

- Если $\lambda \in (0, \alpha/2\sigma^2)$, то в силу (93) существует единственное $\varrho^\lambda_{\alpha,W}$, такое что $\left.\frac{d}{d\varrho}C_{\alpha,W,\varrho}\right|_{\varrho=\varrho^\lambda_{\alpha,W}} = \lambda$. Кроме того, согласно (55) $\varrho^\lambda_{\alpha,W}$ удовлетворяет равенству $C^\lambda_{\alpha,W} = C_{\alpha,W,\varrho^\lambda_{\alpha,W}} - \lambda\varrho^\lambda_{\alpha,W}$, так как $\frac{d}{d\varrho}C_{\alpha,W,\varrho}$ убывает по ϱ. Тогда равенство (98) вытекает из (89), (92). С другой стороны, $q^\lambda_{\alpha,W} = q_{\alpha,W,\varrho^\lambda_{\alpha,W}}$ по лемме 31, поскольку $C_{\alpha,W,\varrho^\lambda_{\alpha,W}} = C^\lambda_{\alpha,W} + \lambda\varrho^\lambda_{\alpha,W}$. Тогда (99) вытекает из (90)–(92), (100). Кроме того, можно проверить, что $\varrho^\lambda_{\alpha,W} = -\frac{d}{d\lambda}C^\lambda_{\alpha,W}$, явным образом решая уравнение $\left.\frac{d}{d\varrho}C_{\alpha,W,\varrho}\right|_{\varrho=\varrho^{\lambda}_{\alpha,W}} = \lambda$ относительно $\varrho^{\lambda}_{\alpha,W}$. Однако нам для проверки равенств (98), (99) это явное решение не потребуется.

- Если $\lambda \in [\alpha/2\sigma^2, \infty)$, то $D_{\alpha}(W \,\|\, \varphi_{\sigma^2} \,|\, p) - \lambda\,\mathbf{E}_p[\varrho] \leqslant 0$ в силу (94). С другой стороны, $C^{\lambda}_{\alpha,W} \geqslant 0$, так как для вероятностной меры, сосредоточенной в точке $x = 0$, АЛ-информация равна нулю. Следовательно, $C^{\lambda}_{\alpha,W} = 0$ и $q^{\lambda}_{\alpha,W} = \varphi_{\sigma^2}$. Таким образом, верны равенства (98), (99).

П р и м е р 5 (параллельные гауссовские каналы). Пусть $W_{[1,n]}$ – произведение скалярных гауссовских каналов W_i с дисперсией шума σ_i для $i \in \{1, \ldots, n\}$, а $\rho_{[1,n]}$ – аддитивные функции стоимости, т.е.

$$W_{[1,n]}(\mathcal{E} \,|\, x_1^n) = \int\limits_{\mathcal{E}} \left[\prod_{i=1}^{n} \varphi_{\sigma_i^2}(y_i - x_i)\right] dy_1^n \quad \forall \mathcal{E} \in \mathcal{B}(\mathbb{R}^n),$$

$$\rho_{[1,n]}(x_1^n) = \sum_{i=1}^{n} x_i^2 \quad \forall x_1^n \in \mathbb{R}^n.$$

В силу леммы 28 августиновская пропускная способность при наличии ограничений по стоимости канала $W_{[1,n]}$ удовлетворяет равенству

$$C_{\alpha,W_{[1,n]},\varrho} = \sup_{\varrho_1,\ldots,\varrho_n\colon \sum\limits_i \varrho_i \leqslant \varrho} C_{\alpha,W_i,\varrho_i}.$$

Так как все C_{α,W_i,ϱ_i} непрерывны, строго вогнуты и возрастают по ϱ_i, то супремум достигается в единственной точке $(\varrho_{\alpha,1}, \ldots, \varrho_{\alpha,n})$. Тогда $q_{\alpha,W_{[1,n]},\varrho} = q_{\alpha,W_1,\varrho_{\alpha,1}} \otimes \otimes \ldots \otimes q_{\alpha,W_n,\varrho_{\alpha,n}}$ по лемме 28. Кроме того, так как C_{α,W_i,ϱ_i} непрерывно дифференцируемы по ϱ_i, эту единственную точку $(\varrho_{\alpha,1}, \ldots, \varrho_{\alpha,n})$ можно найти через условия на производные: $\left.\frac{d}{d\varrho_i}C_{\alpha,W_i,\varrho_i}\right|_{\varrho_i=\varrho_{\alpha,i}} = \lambda_{\alpha}$ для всех i с положительными $\varrho_{\alpha,i}$ и $\left.\frac{d}{d\varrho_i}C_{\alpha,W_i,\varrho_i}\right|_{\varrho_i=\varrho_{\alpha,i}} \leqslant \lambda_{\alpha}$ для всех i с нулевыми $\varrho_{\alpha,i}$ для некоторого $\lambda_{\alpha} \in \mathbb{R}_+$. Таким образом, используя равенство (93), получаем, что оптимальное распределение стоимости, т.е. $(\varrho_{\alpha,1}, \ldots, \varrho_{\alpha,n})$, удовлетворяет равенству

$$\varrho_{\alpha,i} = \frac{|\alpha - 2\sigma_i^2\lambda_{\alpha}|^+}{2\lambda_{\alpha}(\alpha + 2(\alpha-1)\sigma_i^2\lambda_{\alpha})} \tag{102}$$

для некоторого λ_{α}, однозначно задаваемому условием $\sum\limits_{i=1}^{n} \varrho_{\alpha,i} = \varrho$, поскольку выражение в правой части (102) не возрастает по λ_{α} для каждого i. Следовательно,

$$C_{\alpha,W_{[1,n]},\varrho} = \sum_{i=1}^{n} C_{\alpha,W_i,\varrho_{\alpha,i}} \tag{103}$$

$$q_{\alpha,W_{[1,n]},\varrho} = \bigotimes_{i=1}^{n} \varphi_{\theta_{\alpha,\sigma_i,\varrho_{\alpha,i}}}, \tag{104}$$

где $\theta_{\alpha,\sigma,\varrho}$ определено в (91). Используя те же ограничения на оптимальность распределения стоимости через производные, т.е. $\left.\frac{d}{d\varrho_i}C_{\alpha,W_i,\varrho_i}\right|_{\varrho_i=\varrho_{\alpha,i}} = \lambda_{\alpha}$ для всех i с положительными $\varrho_{\alpha,i}$ и $\left.\frac{d}{d\varrho_i}C_{\alpha,W_i,\varrho_i}\right|_{\varrho_i=\varrho_{\alpha,i}} \leqslant \lambda_{\alpha}$ для всех i с нулевыми $\varrho_{\alpha,i}$, а также равенство (92) вместо (93), получаем следующую альтернативную характеризацию

$\theta_{\alpha,\sigma_i,\varrho_{\alpha,i}}$ через σ_i и λ_α, не зависящую явно от величин $\varrho_{\alpha,i}$:

$$\theta_{\alpha,\sigma_i,\varrho_{\alpha,i}} = \sigma_i^2 + \left|\frac{1}{2\lambda_\alpha} - \frac{\sigma_i^2}{\alpha}\right|^+. \tag{105}$$

Используя лемму 32, АЛ-пропускную способность и АЛ-центр канала $W_{[1,n]}$ можно выразить через соответствующие величины для каналов-компонентов следующим образом:

$$C^\lambda_{\alpha,W_{[1,n]}} = \sum_{i=1}^n C^\lambda_{\alpha,W_i},$$
$$q^\lambda_{\alpha,W_{[1,n]}} = \bigotimes_{i=1}^n q^\lambda_{\alpha,W_i}.$$

Августиновские пропускную способность и центр при наличии ограничений по стоимости, а также АЛ-пропускную способность и АЛ-центр векторных гауссовских каналов с многими входными и выходными антеннами можно исследовать, используя этот же подход, с помощью сингулярных разложений матриц.

§ 7. Обсуждение

Так же, как и информация Реньи, информация Августина является обобщением взаимной информации, определяемым чрез расхождение Реньи. Однако в отличие от информации Реньи порядка α информация Августина порядка α не имеет выражения в замкнутом виде, кроме случая порядка 1. Поэтому некоторые свойства информации Августина, такие как ее непрерывная дифференцируемость как функции порядка α, существование и единственность августиновского среднего $q_{\alpha,p}$ порядка α, а также границы (7), доказывать труднее. Но когда эти фундаментальные свойства информации Августина уже установлены, анализ августиновской пропускной способности становится довольно очевидным и весьма похожим на аналогичный анализ для пропускной способности Реньи, представленный в [13].

Для вычисления августиновской пропускной способности при наличии ограничений по стоимости через величину $I^{\mathrm{g}\lambda}_\alpha(p;W)$, которую мы называем РГ-информацией, ранее была применена техника выпуклого сопряжения. Хотя такой подход успешно характеризует августиновскую пропускную способность при наличии ограничений по стоимости через РГ-пропускную способность, он не является стандартным и довольно сложен. Более стандартный подход, основанный на понятии АЛ-информации $I^\lambda_\alpha(p;W)$, представлен в п. 5.2. Насколько нам известно, АЛ-информация ранее не использовалась и не изучалась, тем не менее получаемая пропускная способность оказалась идентична пропускной способности, связанной с РГ-информацией. В свете этого наблюдения оптимальность подхода, основанного на РГ-информации, кажется более наглядной.

Наше изучение информации Августина и августиновской пропускной способности было первоначально мотивировано их операциональным значением в задаче кодирования каналов [6]. Мы исследовали это операциональное значение более подробно и вывели границы сферической упаковки с полиномиальными коэффициентами для двух семейств каналов без памяти – при ограничениях на композицию и при наличии ограничений по стоимости [7]. В целом, вывод границы сферической упаковки для каналов без памяти в [7] аналогичен выводу границы сферической упаковки для каналов-произведений в [38] за исключением использования августиновских пропускной способности и центра вместо пропускной способности и центра Реньи.

Автор выражает благодарность Фатьме и Мехмету Накибоглу за гостеприимство, без которого эта работа не была бы возможной. Автор также благодарит М. Далаи за указание на неявное утверждение Фано о неподвижной точке в [23] и Г. Васкеса-Вилара за указание на работу Полтырева [19] о границе случайного кодирования. Кроме того, автор благодарит рецензента за внимательнейшую рецензию, позволившую исправить ряд не совсем точных и/или нестрогих утверждений в первоначальной версии статьи.

СПИСОК ЛИТЕРАТУРЫ

1. *Nakiboğlu B.* The Augustin Center and the Sphere Packing Bound for Memoryless Channels // Proc. 2017 IEEE Int. Sympos. on Information Theory (ISIT'2017). Aachen, Germany. June 25–30, 2017. P. 1401–1405.
2. *Csiszár, I.* Generalized Cutoff Rates and Rényi's Information Measures // IEEE Trans. Inform. Theory. 1995. V. 41. № 1. P. 26–34.
3. *Dalai M.* Some Remarks on Classical and Classical-Quantum Sphere Packing Bounds: Rényi vs. Kullback–Leibler // Entropy. 2017. V. 19. № 7. P. 355 (11 pp.).
4. *Mosonyi M., Ogawa T.* Divergence Radii and the Strong Converse Exponent of Classical-Quantum Channel Coding with Constant Compositions // `arXiv:1811.10599v4 [quant-ph]`, 2018.
5. Csiszár, I. and Körner, J., *Information Theory: Coding Theorems for Discrete Memoryless Systems*, Cambridge, UK: Cambridge Univ. Press, 2011.
6. *Augustin U.* Noisy Channels // Habilitation Thesis. Universität Erlangen-Nürnberg, 1978. Avaliable at `http://bit.ly/2ID8h7m`.
7. *Nakiboğlu B.* The Sphere Packing Bound for Memoryless Channels // `arXiv:1804.06372v2 [cs.IT]`, 2018.
8. *van Erven T., Harremoës P.* Rényi Divergence and Kullback–Leibler Divergence // IEEE Trans. Inform. Theory. 2014. V. 60. № 7. P. 3797–3820.
9. *Shayevitz O.* A Note on a Characterization of Rényi Measures and Its Relation to Composite Hypothesis Testing // `arXiv:1012.4401v2 [cs.IT]`, 2010.
10. *Shayevitz O.* On Rényi Measures and Hypothesis Testing // Proc. 2010 IEEE Int. Sympos. on Information Theory (ISIT'2010). Austin, Texas, USA. June 13–18, 2010. P. 894–898.
11. *Verdú S.* α-Mutual Information // Proc. 2015 Information Theory and Applications Workshop (ITA'2015). San Diego, CA, USA. Feb. 1–6, 2015. P. 1–6. Available at `http://ita.ucsd.edu/workshop/15/files/paper/paper_374.pdf`.
12. *Kemperman J.H.B.* On the Shannon Capacity of an Arbitrary Channel // Indag. Math. 1974. V. 77. № 2. P. 101–115.
13. *Nakiboğlu B.* The Rényi Capacity and Center // IEEE Trans. Inform. Theory. 2019. V. 65. № 2. P. 841–860.
14. *Gallager R.G.* A Simple Derivation of the Coding Theorem and Some Applications // IEEE Trans. Inform. Theory. 1965. V. 11. № 1. P. 3–18.
15. *Gallager R.G.* Information Theory and Reliable Communication. New York: John Wiley & Sons, 1968.
16. *Ebert P.M.* Error Bounds for Parallel Communication Channels. Tech. Rep. № 448. Research Lab. of Electronics, MIT. Cambridge, MA, USA. 1966. Available at `https://dspace.mit.edu/handle/1721.1/4295`.
17. *Richters J.S.* Communication over Fading Dispersive Channels. Tech. Rep. № 464. Research Lab. of Electronics, MIT. Cambridge, MA, USA. 1967. Available at `https://dspace.mit.edu/handle/1721.1/4279`.
18. *Арутюнян Е.А.* Оценки экспоненты вероятности ошибки для полунепрерывного канала без памяти // Пробл. передачи информ. 1968. Т. 4. № 4. С. 37–48.
19. *Полтырев Г.Ш.* Границы случайного кодирования для дискретных каналов без памяти // Пробл. передачи информ. 1982. Т. 18. № 1. С. 12–26.
20. *Dudley R.M.* Real Analysis and Probability. Cambridge: Cambridge Univ. Press, 2002.

21. *Bogachev V.I.* Measure Theory. Berlin: Springer, 2007.
22. *Nakiboğlu B.* The Augustin Capacity and Center // `arXiv:1803.07937v3 [cs.IT]`, 2018.
23. *Fano R.M.* Transmission of Information: A Statistical Theory of Communications. New York: M.I.T. Press, 1961.
24. *Arimoto S.* Computation of Random Coding Exponent Functions // IEEE Trans. Inform. Theory. 1976. V. 22. № 6. P. 665–671.
25. *Oohama Y.* The Optimal Exponent Function for the Additive White Gaussian Noise Channel at Rates above the Capacity // Proc. 2017 IEEE Int. Sympos. on Information Theory (ISIT'2017). Aachen, Germany. June 25–30, 2017. P. 1053–1057.
26. *Oohama Y.* Exponent Function for Stationary Memoryless Channels with Input Cost at Rates above the Capacity // `arXiv:1701.06545v3 [cs.IT]`, 2017.
27. *Vazquez-Vilar G., Martinez A., Guillén i Fàbregas A.* A Derivation of the Cost-Constrained Sphere-Packing Exponent // Proc. 2015 IEEE Int. Sympos. on Information Theory (ISIT'2015). Hong Kong, China. June 14–19, 2015. P. 929–933.
28. *Rényi A.* On Measures of Entropy and Information // Proc. 4th Berkeley Sympos. on Mathematical Statistics and Probability. Berkely, CA, USA. June 20 – July 30, 1960. Berkely: Univ. of California Press, 1961.
29. *Csiszár I.* Information-type Measures of Difference of Probability Distributions and Indirect Observations // Studia Sci. Math. Hungar. 1967. V. 2. № 3–4. P. 299–318.
30. *Gilardoni G.L.* On Pinsker's and Vajda's Type Inequalities for Csiszár's f-Divergences // IEEE Trans. Inform. Theory. 2010. V. 56. № 11. P. 5377–5386.
31. *Shiryaev A.N.* Probability. New York: Springer, 1995.
32. *Polyanskiy Y., Verdú S.* Arimoto Channel Coding Converse and Rényi Divergence // Proc. 48th Annual Allerton Conf. on Communication, Control, and Computation. Sept. 29 – Oct. 1, 2010. Allerton, IL, USA. P. 1327–1333.
33. *Колмогоров А.Н., Фомин С.В.* Элементы теори функций и функционального анализа. М.: Наука, 1968.
34. *Csiszár I.* A Class of Measures of Informativity of Observation Channels // Period. Math. Hungar. 1972. V. 2. № 1–4. P. 191–213.
35. *Sibson R.* Information Radius // Z. Wahrsch. Verw. Gebiete. 1969. V. 14. № 2. P. 149–160.
36. *Blahut R.E.* Hypothesis Testing and Information Theory // IEEE Trans. Inform. Theory. 1974. V. 20. № 4. P. 405–417.
37. *Kostina V., Verdú S.* Channels with Cost Constraints: Strong Converse and Dispersion // IEEE Trans. Inform. Theory. 2015. V. 61. № 5. P. 2415–2429.
38. *Nakiboğlu B.* The Sphere Packing Bound via Augustin's Method // IEEE Trans. Inform. Theory. 2019. V. 65. № 2. P. 816–840.
39. *Krantz S.G., Parks H.R.* A Primer of Real Analytic Functions. Boston: Birkhäuser, 2002.
40. *Munkres J.R.* Topology. Upper Saddle River, NJ: Prentice Hall, 2000.
41. *Rudin W.* Principles of Mathematical Analysis. New York: McGraw-Hill, 1976.
42. *Bertsekas D.P., Nedić A., Ozdaglar A.E.* Convex Analysis and Optimization. Belmont, MA: Athena Sci., 2003.
43. *Komiya H.* Elementary Proof for Sion's Minimax Theorem // Kodai Math. J. 1988. V. 11. № 1. P. 5–7.
44. *Sion M.* On General Minimax Theorems // Pacific J. Math. 1958. V. 8. № 1. P. 171–176.

Накибоглу Барыш
Средневосточный технический университет,
Анкара, Турция
`bnakib@metu.edu.tr`

Поступила в редакцию
22.03.2018
После доработки
20.08.2019
Принята к публикации
12.11.2019