БЛИЗОСТЬ ВЕРОЯТНОСТНЫХ РАСПРЕДЕЛЕНИЙ В ТЕРМИНАХ ПРЕОБРАЗОВАНИЙ ФУРЬЕ-СТИЛТЬЕСА [1 2]

СЕРГЕЙ Г. БОБКОВ

Дается обзор ряда результатов по неравенствам сглаживания для различных вероятностных метрик, в частности, для расстояния Колмогорова и приводятся некоторые аналоги в классе функций ограниченной вариации

1 Введение

Обозначим через $\mathcal{F}$ пространство всех функций распределения на вещественной прямой, то есть, неубывающих, непрерывных справа функций $F : \mathbf{R} \to [0,1]$, таких что $F(-\infty) = 0$, $F(\infty) = 1$. Существует множество популярных метрик и псевдо-метрик d на $\mathcal{F}$, и в тех или иных задачах аппроксимации возникают вопросы об оценивании расстояния $d(F,G)$ (например, в теории суммирования независимых случайных величин). Неравенства сглаживания позволяют оценить $d(F,G)$ в терминах соответствующих характеристических функций (преобразований Фурье-Стилтьеса)

$$f(t) = \int_{-\infty}^{\infty} e^{itx}\, dF(x), \qquad g(t) = \int_{-\infty}^{\infty} e^{itx}\, dG(x) \qquad (t \in \mathbf{R}), \tag{1.1}$$

[1]Ключевые слова: Вероятностные метрики, неравенства сглаживания
[2]Работа поддержана Фондом Александра фон Гумбольдта и грантом NSF DMS-1612961

хотя возможно и при дополнительных условиях типа гладкости, налагаемых на G.

Важнейшим примером является равномерное расстояние (расстояние Колмогорова)

$$\rho(F,G) = \sup_x |F(x) - G(x)|.$$

Интегрируя в (1.1) по частям, имеем равенство

$$\frac{f(t)-g(t)}{-it} = \int_{-\infty}^{\infty} e^{itx}\,(F(x)-G(x))\,dx, \qquad t \neq 0,$$

где, вообще говоря, интеграл должен пониматься как главное значение. Если функция $\frac{f(t)-g(t)}{t}$ интегрируема, то функция $F-G$ должна быть равномерно непрерывной, и можно применить обратное преобразование Фурье:

$$F(x) - G(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\,\frac{f(t)-g(t)}{-it}\,dt. \tag{1.2}$$

Отсюда получаем элементарную оценку, справедливую без каких-либо предположений гладкости:

$$\rho(F,G) \leq \frac{1}{2\pi}\int_{-\infty}^{\infty} \Big|\frac{f(t)-g(t)}{t}\Big|\,dt. \tag{1.3}$$

Однако во многих задачах интеграл справа расходится, так что эта оценка оказывается бесполезной. Тем не менее, при некоторых условиях гладкости и за счет малой ошибки интеграл в (1.3) удается заменить на интеграл по конечному интервалу $[-T,T]$, на котором может быть известно или показано, что f и g достаточно близки. Это достигается сглаживанием распределений F и G с помощью подходящего (специального выбранного) распределения H, зависящего от параметра T. То есть, рассматривают свертки

$$(F*H)(x) = \int_{-\infty}^{\infty} F(x-y)\,dH(y)$$

и устанавливают соотношения (неравенства сглаживания) вида

$$\rho(F,G) \ \leq \ c\,\rho(F*H,\, G*H) + \varepsilon,$$

где обычно c – абсолютная постоянная, а ε может зависеть от T и некоторых параметров, определяемых F и/или G.

Аналогичные соотношения изучаются и для других метрик d. В итоге получают оценки для $d(F,G)$ в терминах близости f к g; такие оценки тоже называют неравенствами сглаживания (см. напр. [1-3]). В данном обзоре собраны некоторые результаты на эту тему, обсуждаются стандартные подходы к ним и возможные уточнения или обобщения, при этом мы не претендуем на полноту изложения. Будут также представлены несколько новых неравенств для таких расстояний как квадратичное расстояние Канторовича, расстояние в метрике L^1 и расстояние Колмогорова (при дополнительных условиях гладкости, а также с полиномиальным весом). В конце обзора приводятся несколько результатов о скорости сходимости в центральной предельной теореме для слабых метрик (т.е., отвечающих за сходимость вероятностных распределений в слабой топологии).

2 Расстояние Колмогорова

Первый результат о близости распределений в терминах соответствующих преобразований Фурье-Стилтьеса был получен Эссееном в классической работе [4], где доказана следующая важная теорема.

Теорема 2.1. *Если функция распределения G дифференцируема, и $|G'(x)| \le L$ при всех $x \in \mathbf{R}$, то для любого $T > 0$ и всех $b > \frac{1}{2\pi}$*

$$\rho(F,G) \le b \int_{-T}^{T} \Big| \frac{f(t) - g(t)}{t} \Big|\, dt + c(b)\, \frac{L}{T}, \tag{2.1}$$

причем постоянная $c(b)$ зависит только от b.

Основные идеи доказательства теоремы 2.1 появились в более ранней работе Берри [5], рассматривавшего частный случай стандартной нормальной функции распределения $G = \Phi$. Поэтому неравенство (2.1) называют неравенством Берри–Эссеена, так же как и вытекающую из него оценку скорости сходимости в центральной предельной теореме (теорема 20.1). Оно остается в силе для более широкого класса функций: F может быть произвольной ограниченной неубывающей функцией, а G – функцией ограниченной вариации, причем $F(-\infty) = G(-\infty) = 0$ и $|G'(x)| \le L$ (условие монотонности на G снимается, [2–4]). При этом определение равномерного расстояния остается прежним:

$$\rho(F,G) = \sup_x |F(x) - G(x)|.$$

Такое обобщение позволяет работать с разложениями Эджворта для распределений сумм независимых случайных величин.

Вывод (2.1) (или аналогичного ему соотношения) может быть основан на неравенстве сглаживания

$$\rho(F,G) \ \le \ \frac{1}{1-2\gamma}\ \rho(F * H_T, G * H_T) + 2l\, \frac{1-\gamma}{1-2\gamma}\, \frac{L}{T}, \tag{2.2}$$

где $H_T(x) = H(Tx)$, H – вероятностное распределение с плотностью

$$\psi(x) = \frac{1}{2\pi} \left(\frac{\sin \frac{x}{2}}{\frac{x}{2}} \right)^2 = \frac{1 - \cos x}{\pi x^2}, \tag{2.3}$$

причем параметр $l > 0$ может быть любым числом, удовлетворяющим условию $1 - \gamma = H[-l,l] \equiv H(l) - H(-l) < \frac{1}{2}$ (см. доказательство леммы 16.1).

Функция распределения H с плотностью (2.3) и ее сверточные степени – самые популярные в неравенствах сглаживания. Она имеет треугольную характеристическую функцию $h(t) = (1 - |t|)_+$, так что характеристическая функция распределения H_T имеет компактный носитель $[-T,T]$. Применение (1.3) к сглаженным распределениям приводит к оценке

$$\rho(F * H_T, G * H_T) \le \frac{1}{2\pi} \int_{-T}^{T} \Big| \frac{f(t) - g(t)}{t} \Big| \Big(1 - \frac{|t|}{T}\Big)\, dt,$$

что вкупе с (2.2) влечет теорему Эссеена.

Условие на производную в (2.1) можно ослабить до условия липшицевости: $\|G\|_{\mathrm{Lip}} \leq L$. Однако в некоторых задачах (в которых функция G не обязательно непрерывна) желательно его заменить на условие гладкости в “среднем"или “интегральном"смысле. В связи с задачами из вероятностной теории чисел обобщение такого рода было предложено А. С. Файнлейбом [6]. Приведем его для более широкого класса пар (F, G), как и в теореме 2.1 (см. [3], с. 149).

Теорема 2.2. *Пусть F – неубывающая ограниченная функция, и G – функция ограниченной вариации, такие что $F(-\infty) = G(-\infty) = 0$. Для любого $T > 0$ и всех $b > \frac{1}{2\pi}$*

$$\rho(F,G) \leq b \int_{-T}^{T} \Big| \frac{f(t)-g(t)}{t} \Big| \, dt + bT \sup_x \int_{|u| \leq r(b)/T} |G(x+u) - G(x)| \, du, \tag{2.4}$$

причем постоянная $r(b)$ зависит только от b.

Можно упростить это неравенство в терминах модуля непрерывности

$$Q_G(h) = \sup_{|u| \leq h} |G(x+u) - G(u)|, \qquad h \geq 0. \tag{2.5}$$

Когда G – функция распределения, Q_G называется функцией концентрации, и в этом случае ее можно оценить сверху с помощью хорошо известного неравенства

$$Q_G(h) \, \leq \, \Big(\frac{96}{95}\Big)^2 h \int_{-1/h}^{1/h} |g(t)| \, dt, \qquad h > 0$$

(см. [3], с. 59). Таким образом, теорема 2.2 влечет:

Следствие 2.3. *В условиях теоремы 2.2 для любого $T > 0$*

$$c\,\rho(F,G) \, \leq \, \int_{-T}^{T} \Big| \frac{f(t)-g(t)}{t} \Big| \, dt + Q_G\Big(\frac{1}{T}\Big), \tag{2.6}$$

где $c > 0$ – абсолютная постоянная. Более того, если G не убывает, то

$$c\,\rho(F,G) \, \leq \, \int_{-T}^{T} \Big| \frac{f(t)-g(t)}{t} \Big| \, dt + \frac{1}{T} \int_{-T}^{T} |g(t)| \, dt. \tag{2.7}$$

Преимущество неравенства (2.7) перед (2.1) в том, что условие гладкости на G полностью заменяется на функционал, напрямую зависящий от характеристической функции g. В этой связи следует отметить еще одно неравенство

$$\rho(F,G) \, \leq \, \frac{1}{2\pi} \int_{-T}^{T} \Big| \frac{f(t)-g(t)}{t} \Big| \, dt + \frac{1}{T} \int_{-T}^{T} (|f(t)| + |g(t)|) \, dt, \tag{2.8}$$

справедливое для произвольных функций распределения F, G с характеристическими функциями f, g (заметим, что с точностью до абсолютного множителя оно слабее (2.7)). Неравенство (2.8) было получено в работе В. Бенткуса и Ф. Гётце [7] как следствие одного неравенства сглаживания Правица (см. (18.7)); оно успешно применялось в задачах о количестве целых точек внутри многомерных эллипсоидов, а также для изучения асимптотического поведения распределения квадратичных форм от сумм независимых случайных векторов.

3 Расстояние Леви

Расстояние П. Леви между функциями распределения F и G определяется равенством

$$L(F,G) \,=\, \inf\{h \ge 0 :\, G(x-h) - h \le F(x) \le G(x+h) + h,\ \forall\, x \in \mathbf{R}\}.$$

Как легко убедиться, $L(F,G)$ – длина стороны наибольшего квадрата, заключенного между графиками F и G (стороны квадрата параллельны координатным осям).

Расстояние Леви превращает $\mathcal{F}$ в метрическое пространство с топологией слабой сходимости. Именно, $L(F_n,F) \to 0$ при $n \to \infty$ тогда и только тогда, когда $\lim_{n\to\infty} F_n(x) = F(x)$ во всех точках непрерывности F. Поэтому данная метрика находит все больше приложений в задачах, где требуется оценить близость распределений в смысле слабой топологии.

Расстояние Колмогорова связано с расстоянием Леви очевидным соотношением

$$0 \le L(F,G) \le \rho(F,G) \le 1,$$

при этом индуцируя в $\mathcal{F}$ более сильную топологию. С другой стороны, если известно, что $|G'(x)| \le L$ (как в теореме Эссеена), то имеется обратное неравенство

$$\rho(F,G) \le (1+L)\, L(F,G).$$

В этом случае неравенства сглаживания для этих расстояний равносильны. Однако в общем случае естественно ожидать, что оценки близости между F и G в метрике Леви в терминах близости их характеристических функций должны содержать менее ограничительные условия типа гладкости, или же таковые должны полностью отсутствовать. Впервые результаты такого рода были получены Боманом [8], доказавшим следующее утверждение.

Теорема 3.1. *Пусть F и G – функции распределения с характеристическими функциями f и g соответственно. Если $|f(t)-g(t)| \le \lambda|t|$ при всех $t \in \mathbf{R}$, то для любых $x \in \mathbf{R}$ и $h > 0$*

$$G(x-h) - \frac{2\lambda}{h} \le F(x) \le G(x+h) + \frac{2\lambda}{h}. \tag{3.1}$$

В частности,

$$\frac{1}{2}\, L^2(F,G) \,\le\, \sup_t \Big|\frac{f(t)-g(t)}{t}\Big|. \tag{3.2}$$

В доказательство приведем простое рассуждение. Без ограничения общности будем предполагать, что функция $\frac{f(t)-g(t)}{t}$ интегрируема на всей числовой прямой, так что

$$\widehat{A}(t) = e^{-itx}\, \frac{f(t)-g(t)}{-it}$$

является преобразованием Фурье для функции $A(u) = F(x+u) - G(x+u)$ (значение x фиксировано). Поскольку

$$\int_{-\infty}^{\infty} e^{itu}\,(1-|u|)_+\,du \,=\, 2\,\frac{1-\cos t}{t^2},$$

$\widehat{p}(t) = 2\,\frac{1-\cos t}{t^2}$ является характеристической функцией для (вероятностной) треугольной плотности $p(u) = (1-|u|)_+$. Следовательно $\widehat{p}_h(t) = \hat{p}(ht)$ является характеристической функцией для плотности $p_h(u) = \frac{1}{h}\,p(\frac{u}{h})$.

В силу тождества Парсеваля, имеем

$$I \equiv \int_{-\infty}^{\infty} p_h(u) A(u)\,du = \frac{1}{2\pi}\int_{-\infty}^{\infty} \widehat{p}_h(t)\widehat{A}(t)\,dt.$$

Из условия $|f(t)-g(t)| \leq \lambda |t|$ вытекает $|\widehat{A}(t)| \leq \lambda$, и мы получаем

$$|I| \leq \frac{\lambda}{2\pi}\int_{-\infty}^{\infty} \widehat{p}_h(t)\,dt = \frac{\lambda}{2\pi h}\int_{-\infty}^{\infty} \widehat{p}(t)\,dt = \frac{\lambda}{2\pi h}\cdot(2\pi)\,p(0) = \frac{\lambda}{h}.$$

С другой стороны, используя монотонность F и G и тот факт, что p_h имеет носитель $[-h,h]$, получаем нижнюю оценку

$$I = \int_{-h}^{h} p_h(u)\Big(F(x+u) - G(x+u)\Big)\,du \geq F(x-h) - G(x+h).$$

Две оценки приводят к $F(x-h) \leq G(x+h) + \frac{\lambda}{h}$, что эквивалентно правому неравенству в (3.1). Аналогично получаем оценку $-I \geq G(x-h) - F(x+h)$, что приводит к левому неравенству в (3.1).

Дальнейшие обобщения и уточнения результатов Бомана были сделаны В. М. Золотаревым. В частности, ему принадлежит семейство оценок

$$L(F,G) \ \leq\ C_\gamma\bigg(\sup_{t>0}\, \frac{|f(t)-g(t)|}{t^\gamma}\bigg)^{\frac{1}{1+\gamma}} \qquad (\gamma > 0),$$

где можно положить $C_\gamma = \frac{2}{\gamma}\,(1+\gamma)^2\,\pi^{-\frac{1}{1+\gamma}}$ (см. [9], [10]). Но наиболее популярным является следующее неравенство, доказанное в [11].

Теорема 3.2. *Пусть F и G – функции распределения с характеристическими функциями f и g, соответственно. Для любого $T > 1.3$*

$$L(F,G) \ \leq\ \frac{1}{2\pi}\int_{-T}^{T}\Big|\frac{f(t)-g(t)}{t}\Big|\,dt + 2e\frac{\log T}{T}. \tag{3.3}$$

Как позднее было показано А. Ю. Зайцевым [12], присутствие логарифмического множителя в (3.3) существенно. Вывод самого неравенства был основан на неравенстве сглаживания

$$L(F,G) \ \leq\ L(F*H, G*H) + \max\{2\varepsilon, 1 - H(\varepsilon) + H(-\varepsilon)\},$$

где H может быть произвольной функцией распределения, и $\varepsilon > 0$ – тоже любое. Оно применялось к сверточным степеням треугольной плотности с изменением масштаба так, чтобы H как мера была сосредоточена на интервале $[-1,1]$.

Приведем другое доказательство (варианта) неравенства (3.3) на основе сглаживания F и G с помощью нормальной функции распределения

$$\Phi_\sigma(x) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x} e^{-y^2/2\sigma^2}\,dy \qquad (x \in \mathbf{R}),$$

в которой параметр $\sigma > 0$ будет выбран в зависимости от T.

Нам потребуется оценить сверху выражение

$$\mathcal{L}(x,h) = \max\{F(x-h) - G(x+h),\, G(x-h) - F(x+h)\} \qquad (x \in \mathbf{R},\ h \geq 0),$$

связанное с расстоянием Леви через импликацию

$$\sup_x \mathcal{L}(x,h) \leq b + 2h \implies L(F,G) \leq b + 2h \qquad (b \geq 0,\ h \geq 0). \tag{3.4}$$

Для сверток $F_\sigma = F * \Phi_\sigma$ и $G_\sigma = G * \Phi_\sigma$ рассмотрим отклонение

$$I \equiv F_\sigma(x) - G_\sigma(x) \ = \ \int_{-\infty}^{\infty} (F(x-\sigma y) - G(x-\sigma y))\, d\Phi(y),$$

где $\Phi = \Phi_1$ – стантартная нормальная функция распределения. Используя расстояние Колмогорова и неравенство (1.3) для пары (F_σ, G_σ), получаем равномерную оценку

$$|I| \ \leq \ \rho(F_\sigma, G_\sigma) \ \leq \ \frac{1}{2\pi} \int_{-\infty}^{\infty} \Big|\frac{f(t)-g(t)}{t}\Big|\, e^{-\sigma^2 t^2/2}\, dt. \tag{3.5}$$

Оценим теперь I снизу, разбивая этот интеграл на две части: $I = I_0 + I_1$, где

$$\begin{aligned} I_0 \ &= \ \int_{|y| \leq l} (F(x-\sigma y) - G(x-\sigma y))\, d\Phi(y), \\ I_1 \ &= \ \int_{|y| \geq l} (F(x-\sigma y) - G(x-\sigma y))\, d\Phi(y) \end{aligned}$$

с параметром $l > 0$. Как и I, второй интеграл допускает оценку

$$|I_1| \leq 2\,(1 - \Phi(l)) \equiv \gamma, \tag{3.6}$$

в то время как для первого интеграла, используя монотонность F и G, имеем нижнюю оценку

$$I_0 \geq (F(x-\sigma l) - G(x+\sigma l))\,(1-\gamma).$$

Аналогично

$$-I_0 \geq (G(x-\sigma l) - F(x+\sigma l))\,(1-\gamma),$$

так что $|I_0| \geq (1-\gamma)\,\mathcal{L}(x,\sigma l)$. Поскольку $|I_0| \leq |I| + |I_1|$, из (3.5)–(3.6) получаем

$$\sup_x \mathcal{L}(x,\sigma l) \leq \frac{1}{1-\gamma}\,\rho(F_\sigma, G_\sigma) + \frac{\gamma}{1-\gamma}. \tag{3.7}$$

Оценим теперь интеграл в (3.5), используя элементарное неравенство

$$\int_T^{\infty} \frac{1}{t}\, e^{-\sigma^2 t^2/2}\, dt \ < \ \frac{1}{\sigma^2 T^2}\, e^{-\sigma^2 T^2/2}.$$

Так как $|f(t) - g(t)| \leq 2$, имеем

$$\int_T^{\infty} \frac{|f(t)-g(t)|}{t}\, e^{-\sigma^2 t^2/2}\, dt \ \leq \ \frac{2}{\sigma^2 T^2}\, e^{-\sigma^2 T^2/2},$$

и, следовательно, правая часть в (3.5) не превосходит

$$\frac{1}{2\pi}\int_{-T}^{T}\Big|\frac{f(t)-g(t)}{t}\Big|\,dt+\frac{2}{\pi\sigma^2T^2}\,e^{-\sigma^2T^2/2}.$$

Возвращаясь к (3.7), получаем

$$\sup_x \mathcal{L}(x,\sigma l)\ \le\ \frac{1}{2\pi(1-\gamma)}\int_{-T}^{T}\Big|\frac{f(t)-g(t)}{t}\Big|\,dt+\frac{2}{\pi(1-\gamma)\,\sigma^2T^2}\,e^{-\sigma^2T^2/2}+\frac{\gamma}{1-\gamma}.$$

Таким образом, согласно замечанию (3.4) с $h=\sigma l$, если мы покажем, что

$$\frac{2}{\pi(1-\gamma)\,\sigma^2T^2}\,e^{-\sigma^2T^2/2}+\frac{\gamma}{1-\gamma}\ \le\ 2\sigma l, \tag{3.8}$$

то придем к оценке

$$L(F,G)\ \le\ \frac{1}{2\pi(1-\gamma)}\int_{-T}^{T}\Big|\frac{f(t)-g(t)}{t}\Big|\,dt+2\sigma l. \tag{3.9}$$

Положим

$$\sigma=\frac{1}{T}\sqrt{2\log(1+T)},\qquad l=\sqrt{2\log(1+T)}.$$

При таком выборе

$$\gamma=2\,(1-\Phi(l))<e^{-l^2/2}=\frac{1}{1+T},\qquad \frac{\gamma}{1-\gamma}\le\frac{1}{T}.$$

Если, например, $T\ge 1$, то $l\ge\sqrt{\log 4}>1.17$, $\gamma<2\,(1-\Phi(1.17))<0.26$, $\frac{1}{\pi(1-\gamma)}<0.44$, и следовательно

$$\begin{aligned}\frac{2}{\pi(1-\gamma)\,\sigma^2T^2}\,e^{-\sigma^2T^2/2}+\frac{\gamma}{1-\gamma}\ &=\ \frac{1}{\pi(1-\gamma)\,(1+T)\log(1+T)}+\frac{\gamma}{1-\gamma}\\ &<\ \frac{0.44}{(1+T)\log(1+T)}+\frac{2}{1+T}\\ &<\ \frac{4\log(1+T)}{1+T}\ <\ \frac{4\log(1+T)}{T}\ =\ 2\sigma l.\end{aligned}$$

Здесь предпоследнее неравенство нужно проверять лишь при $T=1$.

Таким образом, условие (3.8) выполнено, и мы приходим к (3.9): для всех $T\ge 1$

$$L(F,G)\ \le\ \frac{1}{\pi}\int_{-T}^{T}\Big|\frac{f(t)-g(t)}{t}\Big|\,dt+\frac{4\log(1+T)}{T}.$$

При $0<T<1$ последний член в правой части превосходит 1, и поэтому полученное неравенство справедливо при всех $T>0$.

Отметим одно непосредственное следствие этого неравенства.

Следствие 3.3. *Если характеристические функции f и g совпадают на $[0,T]$, то*

$$L(F,G)\ \le\ \frac{4\log(1+T)}{T}.$$

Используя свойство $|f(t)-g(t)|\le 2$, неравенство (3.2) приводит к асимптотически более слабой оценке $L(F,G)\ \le\ \frac{2}{\sqrt{T}}$.

4 Расстояние по вариации

При изучении других, более сильных, метрик на $\mathcal{F}$ близость сглаженных распределений обычно оценивается по полной вариации (являющейся одной из наиболее сильных метрик). Это вполне оправдано, так как сглаженные распределения имеют как правило гладкие плотности, и расстояния между ними в разных метриках часто имеют одинаковый порядок. Поэтому остановимся на одной стандартной оценке для расстояния по вариации в терминах характеристических функций.

Теорема 4.1. *Для любых функций распределения F и G с непрерывно дифференцируемыми характеристическими функциями f и g*

$$\|F-G\|_{\mathrm{TV}}^4 \le \int_{-\infty}^{\infty} |f(t)-g(t)|^2\,dt \int_{-\infty}^{\infty} |f'(t)-g'(t)|^2\,dt. \tag{4.1}$$

Иногда (4.1) записывают в формально более слабом виде (см. напр. [13]):

$$\|F-G\|_{\mathrm{TV}}^2 \le \frac{1}{2}\int_{-\infty}^{\infty} |f(t)-g(t)|^2\,dt + \frac{1}{2}\int_{-\infty}^{\infty} |f'(t)-g'(t)|^2\,dt. \tag{4.2}$$

Однако, изменяя масштаб, то есть применяя это неравенство к $F_r(x)=F(x/r)$, $G_r(x)=G(x/r)$ и затем оптимизируя по $r>0$ правую часть в (4.2), мы приходим к (4.1).

Неравенство (4.1) сохраняет силу для произвольных функций ограниченной вариации F и G. При этом конечность интегралов в (4.1) обеспечивает абсолютную непрерывность функции $A=F-G$. Так что, если G имеет плотность q, F должна иметь плотность p, и тогда (4.1) превращается в

$$\bigg(\int_{-\infty}^{\infty} |p(x)-q(x)|\,dx\bigg)^4 \le \int_{-\infty}^{\infty} |f(t)-g(t)|^2\,dt \int_{-\infty}^{\infty} |f'(t)-g'(t)|^2\,dt.$$

То есть, теорема 4.1 допускает следующую, более общую формулировку, в которой, кроме того, условие непрерывной дифференцируемости преобразований Фурье можно ослабить до абсолютной непрерывности.

Сначала напомним, что полная вариация $\|A\|_{\mathrm{TV}}$ комплекснозначной функции A, заданной на вещественной прямой, определяется как точная верхняя грань сумм

$$\sum_{k=1}^{n} |A(x_k)-A(x_{k-1})|$$

по всем конечным наборам $x_0<x_1<\dots<x_n$, и ограниченность вариации означает, что $\|A\|_{\mathrm{TV}}<\infty$. В этом случае пределы $A(-\infty)=\lim_{x\to-\infty}A(x)$ и $A(\infty)=\lim_{x\to\infty}A(x)$ конечны, причем без нарушения общности всегда можно считать, что A непрерывна справа и $A(-\infty)=0$. Если A абсолютно непрерывна (локально), то

$$\|A\|_{\mathrm{TV}} = \int_{-\infty}^{\infty} |A'(x)|\,dx,$$

где A' – производная в смысле Радона–Никодима. Часто удобно отождествлять A с борелевской (комплекснозначной) мерой на $\mathbf{R}$, определяемой равенством $A((x,y])=$

$A(y) - A(x)$, $x < y$. При этом A' будет производной Радона–Никодима меры A по отношению к линейной мере Лебега. Единственность A' понимается с точностью до значений на множестве меры нуль (то есть, в пространстве $L^1(\mathbf{R})$).

Теорема 4.2. *Любая (локально) абсолютно непрерывная функция* $a : \mathbf{R} \to \mathbf{C}$, *такая что* $\int |a(t)|^2\, dt < \infty$ *и* $\int |a'(t)|^2\, dt < \infty$, *является преобразованием Фурье интегрируемой функции* $b : \mathbf{R} \to \mathbf{C}$ *(единственной в пространстве* L^1 *), то есть,*

$$a(t) = \int_{-\infty}^{\infty} e^{itx}\, b(x)\, dx, \qquad t \in \mathbf{R}. \tag{4.3}$$

При этом

$$\left(\int_{-\infty}^{\infty} |b(x)|\, dx \right)^4 \le \int_{-\infty}^{\infty} |a(t)|^2\, dt \int_{-\infty}^{\infty} |a'(t)|^2\, dt. \tag{4.4}$$

В доказательство приведем стандартное рассуждение. Сначала предположим, что a и a' принадлежат $L^1(\mathbf{R}) \cap L^2(\mathbf{R})$ и введем функцию

$$b(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx}\, a(t)\, dt. \tag{4.5}$$

Покажем, что она интегрируема на всей числовой прямой.

Предположение $a \in L^1(\mathbf{R}) \cap L^2(\mathbf{R})$ гарантирует, что $b \in L^2(\mathbf{R})$, причем справедлива формула Планшереля

$$\int_{-\infty}^{\infty} |b(x)|^2\, dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} |a(t)|^2\, dt. \tag{4.6}$$

Условия $a \in L^1(\mathbf{R})$ и $a' \in L^1(\mathbf{R})$ позволяют вывести из (4.5) равенство

$$(ix)\, b(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx}\, a'(t)\, dt. \tag{4.7}$$

Действительно, в силу интегрируемости a', функция a имеет ограниченную вариацию, поэтому пределы $a(-\infty)$, $a(\infty)$ существуют и конечны, и более того, $a(-\infty) = a(\infty) = 0$ в силу интегрируемости a. Следовательно, выбирая произвольное $N > 0$, можно проинтегрировать по частям:

$$\int_{-N}^{N} e^{-itx}\, a(t)\, dt \;=\; a(t)\, \frac{e^{-itx}}{-ix} \bigg|_{t=-N}^{t=N} + \frac{1}{ix} \int_{-N}^{N} e^{-itx}\, a'(t)\, dt \qquad (x \neq 0).$$

Устремляя $N \to \infty$, приходим к (4.7).

Условие $a' \in L^1(\mathbf{R}) \cap L^2(\mathbf{R})$ позволяет применить формулу Планшереля на основе (4.7), и мы получаем

$$\int_{-\infty}^{\infty} x^2\, |b(x)|^2\, dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |a'(t)|^2\, dt.$$

Комбинируя это равенство с (4.6), имеем

$$\int_{-\infty}^{\infty} |1 + ix|^2\, |b(x)|^2\, dx = \frac{1}{2\pi} \left(\int_{-\infty}^{\infty} |a(t)|^2\, dt + \int_{-\infty}^{\infty} |a'(t)|^2\, dt \right).$$

Следовательно, в силу неравенства Коши-Буняковского,

$$\begin{aligned}\left(\int_{-\infty}^{\infty}|b(x)|\,dx\right)^2 &= \left(\int_{-\infty}^{\infty}|1+ix|\,|b(x)|\,\frac{1}{|1+ix|}\,dx\right)^2\\ &\leq \int_{-\infty}^{\infty}|1+ix|^2\,|b(x)|^2\,dx\int_{-\infty}^{\infty}\frac{1}{|1+ix|^2}\,dx\\ &\leq \frac{1}{2}\int_{-\infty}^{\infty}|a(t)|^2\,dt+\frac{1}{2}\int_{-\infty}^{\infty}|a'(t)|^2\,dt.\end{aligned}$$

Применение полученного неравенства к $a_r(x) = a(rx)$ и оптимизация по r приводит к (4.4). Таким образом, функция b действительно интегрируема и, значит, можно применить обратное преобразование Фурье на основе (4.5), что приводит к (4.3). Теорема доказана при сделанных дополнительных предположениях.

Рассмотрим теперь общий случай и сначала переформулируем теорему на операторном языке. Будем обозначать через $\|\cdot\|_p$ обычную L^p-норму функций (по мере Лебега). Основное функциональное пространство в теореме 4.2 – это пространство Соболева $W_1^2 = W_1^2(\mathbf{R})$ всех абсолютно непрерывных комплекснозначных функций $a(t)$ на вещественной прямой с конечной нормой

$$\|a\|_{W_1^2} = \left(\frac{1}{2}\,\|a\|_2^2+\frac{1}{2}\,\|a'\|_2^2\right)^{1/2}.$$

Нам потребуется:

Lemma 4.3. *Линейное подпространство* $H = \{a \in W_1^2 : a \in L^1(\mathbf{R}),\ a' \in L^1(\mathbf{R})\}$ *плотно в* W_1^2.

Действительно, каждая функция $a \in H$ может быть аппроксимирована по норме W_1^2 функциями

$$a_\sigma(t) = a(t)\,e^{-\sigma^2t^2/2} \qquad (\sigma > 0).$$

Очевидно, $a_\sigma \in L^1(\mathbf{R}) \cap L^2(\mathbf{R})$, и так как $|a_\sigma(t)| \leq |a(t)|$, по теореме Лебега о мажорирующей сходимости

$$\int_{-\infty}^{\infty}|a_\sigma(t)-a(t)|^2\,dt \to 0 \qquad (\sigma \to 0).$$

Кроме того, так как $a_\sigma'(t) = a'(t)\,e^{-\sigma^2t^2/2} - a(t)\cdot\sigma^2\,te^{-\sigma^2t^2/2}$ (где равенство понимается в смысле Радона-Никодима), имеем

$$|a_\sigma'(t)-a'(t)| \leq |a'(t)|\,(1-e^{-\sigma^2t^2/2}) + |a(t)|\cdot\sigma^2|t|\,e^{-\sigma^2t^2/2}.$$

Следовательно, используя $xe^{-x^2/2} \leq 1/\sqrt{e}$ $(x \geq 0)$, получаем

$$|a_\sigma'(t)-a'(t)|^2 \leq 2\,|a'(t)|^2\,(e^{-\sigma^2t^2/2}-1)^2 + \sigma^2\,|a(t)|^2,$$

и снова по теореме Лебега

$$\int_{-\infty}^{\infty}|a_\sigma'(t)-a'(t)|^2\,dt \to 0 \qquad (\sigma \to 0).$$

Таким образом, $\|a_\sigma - a\|_{W_1^2} \to 0$ при $\sigma \to 0$, так что H плотно в W_1^2.

Продолжая доказательство теоремы 4.2, обозначим через $\mathcal{P}b = \widehat{b}$ преобразование Фурье, действующее на функциях $b \in L^1(\mathbf{R})$. Для функций a из подпространства H на первом шаге мы рассмотрели функции $b = \mathcal{G}a$, определяемые по формуле (4.5), для которых имеем тождество (4.6) и неравенство (4.4). Следовательно,

$$\|\mathcal{G}a\|_1 \leq \|a\|_{W_1^2}, \qquad \|\mathcal{G}a\|_2 \leq \frac{1}{\sqrt{2\pi}}\,\|a\|_2 \leq \|a\|_{W_1^2}. \tag{4.8}$$

Кроме того, $\mathcal{P}\mathcal{G}a = a$. В частности, $\mathcal{G} : H \to L^1(\mathbf{R}) \cap L^2(\mathbf{R})$ представляет собой линейный непрерывный оператор по отношению к норме $\|b\| = \|b\|_1 + \|b\|_2$ в $L^1(\mathbf{R}) \cap L^2(\mathbf{R})$. Но H плотно в W_1^2, и поэтому $\mathcal{G}$ допускает продолжение по непрерывности на все пространство W_1^2, сохраняя при этом неравенства в (4.8). Так как оператор $\mathcal{P}$ непрерывен по норме L^2, тождество $\mathcal{P}\mathcal{G}a = a$ тоже сохраняется. Итак, для всех $a \in W_1^2$

$$\|b\|_1 \leq \left(\frac{1}{2}\,\|a\|_2^2 + \frac{1}{2}\,\|a'\|_2^2\right)^{1/2}, \qquad \mathcal{P}b = a,\ b = \mathcal{G}a.$$

Остается сделать это неравенство однородным по пространственной переменной, как уже было упомянуто на первом шаге. Теорема 4.2 доказана полностью.

Замечание. Использование производных в неравенстве (4.1) необязательно, если известно, что распределения F и G имеют компактный носитель. Действительно, переходя к обозначениям теоремы 4.3, если функция b в (4.3) сосредоточена на интервале $[-R, R]$, $R > 0$, то, применяя (4.6), получаем

$$\begin{aligned}\left(\int_{-\infty}^{\infty} |b(x)|\,dx\right)^2 \ &= \ \left(\int_{-R}^{R} |b(x)|\,dx\right)^2 \\ &\leq \ 2R\int_{-R}^{R} |b(x)|^2\,dx \ = \ \frac{R}{\pi}\int_{-\infty}^{\infty} |a(t)|^2\,dt.\end{aligned} \tag{4.9}$$

В случае $a = f - g$, $a \in L^2$ функция $A = F - G$ абсолютно непрерывна, и как мера, она имеет плотность $b = A'$, так что $\|A\|_{\mathrm{TV}} = \int_{-\infty}^{\infty} |b(x)|\,dx$. Следовательно, применяя (4.9) к b, неравенство (4.1) упрощается:

Теорема 4.4. *Для любых функций распределения F и G, сосредоточенных на интервале $[-R, R]$ и имеющих характеристические функции f и g,*

$$\|F - G\|_{\mathrm{TV}}^2 \ \leq \ \frac{R}{\pi}\int_{-\infty}^{\infty} |f(t) - g(t)|^2\,dt.$$

Если интеграл справа конечен, то функция $A = F - G$ абсолютно непрерывна. Как мера, она имеет плотность $b = A'$, так что $\|A\|_{\mathrm{TV}} = \int_{-\infty}^{\infty} |b(x)|\,dx$.

5 Расстояние Кульбака-Лейблера

Еще более сильным расстоянием по сравнению с полной вариацией разности распределений является расстояние Кульбака-Лейблера, также называемое относительной энтропией или величиной информационной расходимости. Для абсолютно непрерывных вероятностных распределений F и G на прямой с плотностями p и q это расстояние определяется равенством

$$D(F||G) = \int_{-\infty}^{\infty} p(x) \log \frac{p(x)}{q(x)}\, dx$$

в предположении, что $q(x) = 0 \Rightarrow p(x) = 0$ для почти всех x (то есть, как мера, распределение F абсолютно непрерывно относительно меры G). Во всех остальных случаях полагают $D(F||G) = \infty$.

В общем случае $0 \leq D(F||G) \leq \infty$ и $D(F||G) = 0 \Leftrightarrow F = G$. Однако этот функционал несимметричен по (F, G) и поэтому не является метрикой в пространстве $\mathcal{F}$. Тем не менее, во многих задачах величина $D(F||G)$ служит удобной мерой близости F к G. Она связана с расстоянием по вариации хорошо известным неравенством (типа) Пинскера

$$D(F||G) \,\geq\, \frac{1}{2}\, \|F - G\|_{\mathrm{TV}}^2 \,=\, \frac{1}{2} \left(\int_{-\infty}^{\infty} |p(x) - q(x)|\, dx \right)^2.$$

Отметим еще связь с классической энтропией

$$h(X) = -\int_{-\infty}^{\infty} p(x) \log p(x)\, dx,$$

где X – случайная величина с плотностью p. Именно, если Z – гауссовская случайная величина с распределением G, причем X и Z имеют одинаковые математические ожидания и дисперсии, то

$$D(F||G) = h(Z) - h(X).$$

Вопрос об оценивании близости F к G в смысле расстояния Кульбака-Лейблера в терминах характеристических функций еще слабо изучен. Приведем один результат для важного случая, когда распределение $G = \Phi$ – стандартное нормальное, то есть, имеет характеристическую функцию $g(t) = e^{-t^2/2}$. Положим

$$g_\alpha(t) = e^{-t^2/2} \left(1 + \alpha\, \frac{(it)^3}{3!}\right), \qquad t, \alpha \in \mathbf{R}.$$

Функция g_α является преобразованием Фурье для плотности

$$\varphi_\alpha(x) = \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2} \left(1 + \alpha\, \frac{x^3 - 3x}{3!}\right)$$

“поправленного"гауссовского распределения (заметим, что эта плотность может принимать отрицательные значения).

Как и раньше, $\|u\|_2 = (\int_{-\infty}^{\infty} |u(t)|^2\, dt)^{1/2}$ обозначает L^2-норму функции u.

Теорема 5.1. *Пусть F – вероятностное распределение на числовой прямой с характеристической функцией f, причем $\int_{-\infty}^{\infty} |x|^3\, dF(x) < \infty$. Для любого $\alpha \in \mathbf{R}$ имеем*

$$D(F||\Phi) \,\leq\, \alpha^2 + 4 \Big(\|f - g_\alpha\|_2 + \|f''' - g_\alpha'''\|_2 \Big). \tag{5.1}$$

Условие конечности 3-го абсолютного момента F гарантирует, что характеристическая функция f имеет 3 непрерывные производные.

Неравенство (5.1) доказано в [14] при изучении энтропийного варианта центральной предельной теоремы. При $\alpha = 0$ оно упрощается,

$$D(F||\Phi) \ \le \ 4\Big(\|f-g\|_2 + \|f''' - g'''\|_2\Big),$$

что можно рассматривать как полный аналог теоремы 4.1. Однако выбор других значений α приводит к более точным оценкам для $D(F||\Phi)$.

6 Расстояние Леви-Прохорова

Вернемся теперь к слабым вероятностным метрикам (то есть, метрикам, отвечающим за слабую сходимость вероятностных мер).

Естественной модификацией расстояния Леви для борелевских вероятностных мер F и G на произвольном метрическом пространстве (M,d) является расстояние Леви-Прохорова $\pi(F,G)$, введенное в [15] для метризации слабой сходимости. Оно определяется как точная нижняя грань значений $h \ge 0$, таких что

$$F(A) \le G(A_h) + h \tag{6.1}$$

для всех борелевских множеств $A \subset M$. Здесь A_h обозначает открытую h-окрестность A в метрике d, то есть,

$$A_h = \{x \in M : \exists\, y \in M,\ d(x,y) < h\}.$$

Хотя формально в этом определении имеется несимметрия, дуальное неравенство $G(A) \le F(A_h) + h$ легко вытекает из (6.1).

В случае $M = \mathbf{R}$, если ограничиваться в (6.1) только полуосями A, то мы бы возвратились к расстоянию Леви. Поэтому расстояние Леви-Прохорова сильнее: $L(F,G) \le \pi(F,G)$ (в случае вещественной прямой будем отождествлять вероятностные меры F и ассоциированные с ними функции распределения $x \to F(-\infty, x]$). Тем не менее, обе эти метрики порождают одну и ту же топологию слабой сходимости в $\mathcal{F}$. В случае евклидова пространства $M = \mathbf{R}^k$ другим обобщением метрики Леви может служить расстояние, определяемое с помощью (6.1) для класса всех полупространств; оно иногда называется расстоянием Цирельсона.

В. В. Юринским [16] был предложен один вариант оценивания близости вероятностных распределений на $\mathbf{R}^k$ в метрике Леви-Прохорова в терминах близости соответствующих характеристических функций при дополнительных моментных условиях

$$\int_{\mathbf{R}^k} \|x\|^{[\frac{k}{2}]+1}\, dF(x) < \infty, \qquad \int_{\mathbf{R}^k} \|x\|^{[\frac{k}{2}]+1}\, dG(x) < \infty \tag{6.2}$$

и в предположении, что G имеет плотность q, удовлетворяющую

$$\int_{\mathbf{R}^k} |q(x+h) - q(x)|\, dx \le \lambda\, \|h\| \tag{6.3}$$

(условие типа гладкости по аналогии с теоремой 2.1). При этих условиях им установлено неравенство сглаживания

$$\pi(F,G) \,\le\, c_1\,\pi(F*H,G*H)+c_2(1+\lambda)\int_{\mathbf{R}^k}\|x\|\,dH(x),$$

справедливое для любой вероятностной меры H на $\mathbf{R}^k$ с постоянными c_1 и c_2, зависящими только от размерности k. Расстояние Леви-Прохорова между сглаженными распределениями можно оценить по вариации (теорема 4.1 при $k=1$), что в итоге приводит к следующему результату ([16]). Мы приведем его в одномерной ситуации с учетом одного замечания В. А. Абрамова о возможном ослаблении условия (6.2) в терминах псевдомоментов.

Теорема 6.1. *Пусть F и G – функции распределения с характеристическими функциями f и g, соответственно, такие что*

$$\int_{-\infty}^{\infty}|x|\,|F-G|(dx)<\infty. \tag{6.4}$$

Если выполнено условие (6.3)*, то для любого $T>0$ с некоторой абсолютной постоянной $c>0$*

$$c\,\pi(F,G) \,\le\, \left(\int_{-T}^{T}|f(t)-g(t)|^2+|(f(t)-g(t))'|^2\,dt\right)^{1/2}+\frac{1+\lambda}{T}. \tag{6.5}$$

Здесь $|F-G|$ – вариация функции $F-G$, рассматриваемая как конечная положительная мера на прямой. Отметим, что непрерывная дифференцируемость функции $f-g$ гарантируется моментным условием (6.4).

Хотя первое слагаемое в правой части (6.5) в общем случае намного больше аналогичного интегрального слагаемого в неравенстве Эссеена, главной особенностью (6.5) является убывание второго слагаемого (что достигается за счет (6.3)). Поэтому в приложениях, связанных со скоростью сходимости в центральной предельной теореме (когда $G=\Phi$), неравенство Юринского приводит к правильным асимптотикам и таким образом существенно усиливает утверждение о нормальной аппроксимации для сумм независимых величин (по сравнению с утверждением о нормальной аппроксимации в смысле расстояния Колмогорова).

В общем случае, то есть, без дополнительных предположений типа гладкости оценки близости $\pi(F,G)$ в терминах характеристических функций изучались В. А. Абрамовым [17] и А. Ю. Зайцевым [18], [12]; см. также [19]. Приведем один из результатов, полученных в [18], [12].

Теорема 6.2. *Пусть F и G – функции распределения, удовлеворяющие условию* (6.2) *при $k=1$, с характеристическими функциями f и g. Для любого $T>e$ имеем*

$$c\,\pi(F,G) \,\le\, I(F,G)+\frac{\log T}{T}, \tag{6.6}$$

где $c>0$ – абсолютная постоянная, и

$$I(F,G)=\left(\int_{-T}^{T}|f(t)-g(t)|^2\,dt\,\int_{-T}^{T}\Big(\frac{\log T}{T^2}\,|f(t)-g(t)|^2+|f'(t)-g'(t)|^2\Big)\,dt\right)^{1/4}.$$

Наличие члена $\frac{\log T}{T}$ в правой части (6.6) сближает это неравенство с неравенством Золотарева для расстояния Леви, в то время как функционал $I(F,G)$ напоминает оценку в (4.1) для полной вариации (правда, с той существенной разницей, что интегралы теперь берутся по конечному интервалу $[-T,T]$). Усиление неравенства Золотарева проявляется, например, в следующем уточнении следствия 3.3 в терминах расстояния Леви-Прохорова.

Следствие 6.3. *Если характеристические функции f и g совпадают на интервале $[0,T]$, $T > e$, то*

$$\pi(F,G) \, \leq \, c\,\frac{\log T}{T},$$

где c – абсолютная постоянная.

7 Расстояние в метрике L^p

Другим важным расстоянием в пространстве $\mathcal{F}$ является расстояние в метрике $L^p(\mathbf{R})$

$$\|F-G\|_p \, = \, \bigg(\int_{-\infty}^{\infty} |F(x)-G(x)|^p \, dx\bigg)^{1/p} \qquad (1 \leq p < \infty).$$

В отличие от расстояний Леви и Леви-Прохорова, оно однородно порядка $1/p$ по пространственной переменной: если случайные величины X и Y имеют функции распределения F и G, то для функций распределения $F_r(x) = F(x/r)$, $G_r(x) = G(x/r)$ случайных величин rX, rY $(r > 0)$, имеем

$$\|F_r - G_r\|_p \, = \, r^{1/p}\, \|F-G\|_p.$$

Функция $p \rightarrow \|F-G\|_p^p$ не возрастает, так что $\|F-G\|_p \leq \|F-G\|_1^{1/p}$. Поэтому, для конечности L^p-расстояния достаточно потребовать конечность абсолютных моментов $\mathbf{E}\,|X|$ и $\mathbf{E}\,|Y|$. Более того, если $\mathbf{E}\,|X|^\varepsilon < \infty$ и $\mathbf{E}\,|Y|^\varepsilon < \infty$ при некотором $\varepsilon > 0$, то как нетрудно видеть,

$$\lim_{p\rightarrow\infty} \|F-G\|_p = \|F-G\|_\infty = \rho(F,G).$$

Отметим еще одно элементарное соотношение

$$L(F,G) \leq \|F-G\|_p^{p/(p+1)}. \tag{7.1}$$

Действительно, если $L(F,G) > h \geq 0$, то из определения расстояния Леви немедленно следует, что найдется точка $x_0 \in \mathbf{R}$, такая что $G(x_0-h) - h > F(x_0)$ или $F(x_0) > G(x_0+h) + h$. Пусть для определенности выполнено второе неравенство. Тогда ввиду монотонности F и G,

$$\|F-G\|_p^p \, \geq \, \int_{x_0}^{x_0+h} |F(x)-G(x)|^p \, dx \, \geq \, (F(x_0) - G(x_0+h))^p \, h \, \geq \, h^{p+1}.$$

Это доказывает неравенство (7.1).

Отсюда можно заключить, что топология, порождаемая L^p-расстоянием – более сильная, чем топология слабой сходимости в $\mathcal{F}$. Тем не менее, на любом подпространстве вероятностных распределений с ограниченным абсолютным моментом порядка $\alpha > 1$, то есть, при условии $\int_{-\infty}^{\infty} |x|^\alpha \, dF(x) \leq M$ с фиксированным параметром M сходимость в метрике L^p будет равносильна слабой сходимости.

При $p \geq 2$ можно оценить расстояние $\|F-G\|_p$ в терминах соответствующих характеристических функций, воспользовавшись классическим неравенством Хаусдорфа-Юнга

$$\|\widehat{a}\|_p \leq \|a\|_q, \qquad q = \frac{p}{p-1}, \tag{7.2}$$

имеющим силу для любой интегрируемой комлекснозначной функции a на вещественной прямой. Здесь

$$\hat{a}(x) = \int_{-\infty}^{\infty} e^{2\pi i\, tx} a(t)\, dt \qquad (x \in \mathbf{R})$$

– преобразование Фурье функции a с измененным масштабом.

Согласно (1.2), $F - G$ представляет собой преобразование Фурье функции $a(t) = \frac{1}{2\pi} \frac{f(t)-g(t)}{-it}$ в обычном смысле, когда $a(t)$ интегрируема. Следовательно, можно применить (7.2), и тогда мы приходим к следующей оценке (в которой условие интегрируемости легко снимается).

Теорема 7.1. *Для любых функций распределений F и G с характеристическими функциями f и g, соответственно, при $p \geq 2$ имеем*

$$\|F - G\|_p \, \leq \, \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} \Big|\frac{f(t) - g(t)}{t}\Big|^q \, dt\right)^{1/q}, \qquad q = \frac{p}{p-1}. \tag{7.3}$$

Это неравенство сохраняет силу для произвольных функций ограниченной вариации F и G, удовлетворяющих $F(-\infty) = G(-\infty)$, $F(\infty) = G(\infty)$ (в этом случае можно интегрировать по частям в (1.1) с тем, чтобы получить (1.2)). При этом в пределе при $p \to \infty$ (7.3) возвращает нас к оценке (1.3) для расстояния Колмогорова.

При $2 < p < \infty$ можно улучшить константу $\frac{1}{2\pi}$ в (7.3), применяя уточненное неравенство Хаусдорфа-Юнга (см. [20]). С другой стороны, при $p = 2$ это неравенство превращается в равенство в силу теоремы Планшереля:

$$\|F - G\|_2^2 \, = \, \frac{1}{2\pi} \int_{-\infty}^{\infty} \Big|\frac{f(t) - g(t)}{t}\Big|^2 \, dt. \tag{7.4}$$

Применяя это равенство в (7.1), получаем еще одну известную оценку для расстояния Леви:

$$L(F,G) \, \leq \, \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} \Big|\frac{f(t) - g(t)}{t}\Big|^2 \, dt\right)^{1/3}.$$

Однако, эта оценка вряд ли может конкурировать с неравенством Золотарева.

С учетом того, что $|f(t) - g(t)| \leq 2$, чтобы судить о близости F и G в метрике L^p, достаточно, чтобы характеристические функции f и g были близки на большом интервале. Действительно,

$$\int_{|t| \geq T} \Big|\frac{f(t) - g(t)}{t}\Big|^q \, dt \, \leq \, 2^{q+1} \int_T^{\infty} \frac{1}{t^q} \, dt \, = \, \frac{2^{q+1}}{q-1} \, \frac{1}{T^{q-1}}.$$

Используя
$$\left(\frac{2^{q+1}}{(q-1)\,T^{q-1}}\right)^{1/q} \ = \ 2^{1+1/q}\,(p-1)^{\frac{p-1}{p}}\ \frac{1}{T^{1/p}} \ \le \ \frac{4(p-1)}{T^{1/p}}$$
и $(\frac{1}{2\pi})^{1/q} < \frac{1}{2}$, получаем из (7.3) следующее:

Следствие 7.2. *Пусть F и G – функции распределения с характеристическими функциями f и g. При $p \ge 2$ для любого $T > 0$ имеем*
$$\|F-G\|_p \ \le \ \frac{1}{2}\left(\int_{-T}^{T}\Big|\frac{f(t)-g(t)}{t}\Big|^q\,dt\right)^{1/q} + \frac{4(p-1)}{T^{1/p}}. \tag{7.5}$$
В частности, если f и g совпадают на интервале $[0,T]$, то $\|F-G\|_p \ \le \ \frac{4(p-1)}{T^{1/p}}$.

8 Расстояние в метрике L^1

При $1 \le p < 2$ неравенство (7.3) перестает быть верным, и нужны другие подходы к оцениванию близости распределений в метрике L^p в терминах характеристических функций. Рассмотрим наиболее интересный случай $p = 1$, когда мы имеем дело со средним расстоянием в пространстве $\mathcal{F}$. В виду взаимосвязи расстояния в метрике L^1 с расстоянием Канторовича и другими транспортными метриками, будем использовать еще одно стандартное обозначение
$$W_1(F,G) = \|F-G\|_1 = \int_{-\infty}^{\infty}|F(x)-G(x)|\,dx \qquad (F,G\in\mathcal{F}).$$
Напомним, что согласно неравенству (7.1) при $p = 1$, имеет место соотношение $L^2(F,G) \le W_1(F,G)$.

Пусть f и g – характеристические функции для функций распределения F и G. Если величина $W_1(F,G)$ конечна, то $f-g$ непрерывно дифференцируема, причем функция $a(t) = \frac{f(t)-g(t)}{-it}$ является преобразованием Фурье функции $b = F - G$. Поэтому можно применить теорему 4.2, которая приводит к оценке
$$W_1^2(F,G) \ \le \ \frac{1}{2}\int_{-\infty}^{\infty}\Big|\frac{f(t)-g(t)}{t}\Big|^2\,dt + \frac{1}{2}\int_{-\infty}^{\infty}\Big|\frac{d}{dt}\ \frac{f(t)-g(t)}{t}\Big|^2\,dt.$$
Так же, как это было сделано перед следствием 7.2, здесь можно сузить интегрирование до интервала $[-T,T]$, и тогда в лучшем случае мы получим оценку типа
$$W_1(F,G) \le \left(\int_{-T}^{T}\Big|\frac{f(t)-g(t)}{t}\Big|^2\,dt\right)^{1/2} + \left(\int_{-T}^{T}\Big|\frac{d}{dt}\ \frac{f(t)-g(t)}{t}\Big|^2\,dt\right)^{1/2} + \frac{c}{\sqrt{T}} \tag{8.1}$$
по аналогии с неравенством (7.5) при $p = 2$ (причем с дополнительным интегралом, содержащим производную характеристических функций).

Для того, чтобы заменить последний член $\frac{c}{\sqrt{T}}$ на $\frac{c}{T}$ (что согласовалось бы со степенью T в (7.5) при $p = 1$), Эссеен предложил использовать сглаживание с помощью специальной конечной знакопеременной меры H_T на прямой с параметром $T > 0$, у которой преобразование Фурье-Стилтьеса задается равенствами

$$\begin{aligned} h_T(t) &= 1, && |t| \le \frac{T}{2}, \\ &= 0, && |t| \ge T, \\ &= \frac{2(T-|t|)}{T}, && \frac{T}{2} \le |t| \le T. \end{aligned}$$

Заметим, что $0 \le h_T(t) \le 1$ и $|h_T'(t)| \le \frac{2}{T}$ ($|t| \ne T, T/2$). В терминах характеристических функций $v_T(t) = (1 - \frac{|t|}{T})_+$ можно записать $h_T = 2v_T - v_{T/2}$. Отсюда сразу вытекает, что $\|H_T\|_{\mathrm{TV}} \le 3$.

С некоторыми модификациями будем далее следовать в основном изложению [13]. Обозначим через V класс всех комплекснозначных функций $A = A(x)$ на вещественной прямой, имеющих ограниченную вариацию, и через $\widetilde{V}$ класс их преобразований Фурье–Стилтьеса $a = a(t)$. Положим $\|a\|_{\mathrm{tv}} = \|A\|_{\mathrm{TV}}$.

Класс V замкнут относительно операции свертки, поэтому класс $\widetilde{V}$ замкнут относительно умножения и представляет собой алгебру, в которой введенная норма обладает свойствами

$$\|a+b\|_{\mathrm{tv}} \le \|a\|_{\mathrm{tv}} + \|b\|_{\mathrm{tv}}, \quad \|ab\|_{\mathrm{tv}} \le \|a\|_{\mathrm{tv}}\, \|b\|_{\mathrm{tv}}$$

для всех $a, b \in \widetilde{V}$. При изменении масштаба норма по вариации не меняется, поэтому для $a_r(t) = a(rt)$ также имеем свойство $\|a_r\|_{\mathrm{tv}} = \|a\|_{\mathrm{tv}}$.

Любая характеристическая функция имеет норму $\|a\|_{\mathrm{tv}} = 1$, и, в частности, $\|1\|_{\mathrm{tv}} = 1$. В общем случае, когда функция a абсолютно непрерывна, условия $\int |a(t)|^2\,dt < \infty$ и $\int |a'(t)|^2\,dt < \infty$ гарантируют, что $a \in \widetilde{V}$, и по теореме 4.2

$$\|a\|_{\mathrm{tv}} \le \Big(\int_{-\infty}^{\infty} |a(t)|^2\,dt \int_{-\infty}^{\infty} |a'(t)|^2\,dt \Big)^{1/4}. \tag{8.2}$$

Пример. Рассмотрим функцию

$$\begin{aligned} u_T(t) &= \frac{4t}{-iT^2}, && |t| \le \frac{T}{2}, \\ &= \frac{1}{-it}, && |t| \ge \frac{T}{2}. \end{aligned}$$

Чтобы оценить норму этой функции, можно использовать тождество $u_T(t) = \frac{2}{T}\, u_2(\frac{2t}{T})$, что сводит задачу к случаю $T = 2$. Применяя (8.2), сразу получаем $\|u_2\|_{\mathrm{tv}} \le \sqrt{\frac{8}{3}}$, и следовательно,

$$\|u_T\|_{\mathrm{tv}} \le \frac{c}{T}, \qquad c = 2\sqrt{\frac{8}{3}}. \tag{8.3}$$

Следует отметить, что задача о минимизации $\|u\|_{\mathrm{tv}}$ среди всех функций $u \in \widetilde{V}$, таких что $u(t) = \frac{1}{-it}$ при $|t| \ge 1$, изучалась Beurling'ом (см. [4]). Как оказалось, минимальное значение равно $\frac{\pi}{2} = 1.57...$ (пример u_2 дает немного худшее значение $\sqrt{\frac{8}{3}} = 1.63...$).

Пусть теперь A – заданная функция ограниченной вариации с преобразованием Фурье-Стилтьеса

$$a(t) = \int_{-\infty}^{\infty} e^{itx}\,dA(x).$$

Если A интегрируема, то $A(-\infty) = A(\infty) = 0$, и функция a – абсолютно непрерывна (локально). Более того, можно проинтегрировать по частям:

$$b(t) \equiv \frac{a(t)}{-it} = \int_{-\infty}^{\infty} e^{itx}\, A(x)\, dx = \int_{-\infty}^{\infty} e^{itx}\, dB(x), \qquad B(x) = \int_{-\infty}^{x} A(y)\, dy.$$

Следовательно b принадлежит $\widetilde{V}$ и имеет норму

$$\|b\|_{\rm tv} = \|B\|_{\rm TV} = \int_{-\infty}^{\infty} |A(x)|\, dx.$$

С другой стороны, представляя $b = bh_T + b\cdot(1-h_T)$, по неравенству треугольника получаем

$$\|b\|_{\rm tv} \ \leq \ \|bh_T\|_{\rm tv} + \|b\cdot(1-h_T)\|_{\rm tv}. \tag{8.4}$$

Так как $1 - h_T(t) = 0$ при $|t| \leq \frac{T}{2}$, равенство

$$b(t)\cdot(1-h_T(t)) = a(t)\cdot u(t)\cdot(1-h_T(t))$$

справедливо для любой функции $u \in \widetilde{V}$, такой что $u(t) = 1/(-it)$ для $|t| \geq \frac{T}{2}$. Более того, в этом случае

$$\|b\cdot(1-h_T)\|_{\rm tv} \leq \|a\|_{\rm tv}\, \|u\|_{\rm tv}\, \|1-h_T\|_{\rm tv}.$$

Здесь $\|a\|_{\rm tv} = \|A\|_{\rm TV}$ и $\|1-h_T\|_{\rm tv} \leq 1 + \|h_T\|_{\rm tv} \leq 4$, так что $\|b\cdot(1-h_T)\|_{\rm tv} \leq 4\,\|A\|_{\rm TV}\,\|u\|_{\rm tv}$. Выбирая $u = u_T$, как в примере, из (8.4) получаем неравенство типа сглаживания

$$\int_{-\infty}^{\infty} |A(x)|\, dx \ \leq \ \|bh_T\|_{\rm tv} + \frac{4c\,\|A\|_{\rm TV}}{T}, \qquad c = 2\sqrt{\frac{8}{3}}. \tag{8.5}$$

Перейдем к первому члену в правой части этого неравенства. Применяя (8.2) и учитывая, что $h_T = 0$ вне интервала $(-T, T)$, имеем

$$\|bh_T\|_{\rm tv}^2 \ \leq \ \frac{1}{2}\int_{-T}^{T} |b(t)\, h_T(t)|^2\, dt + \frac{1}{2}\int_{-T}^{T} |(b(t)\, h_T(t))'|^2\, dt. \tag{8.6}$$

Здесь первый интеграл не превосходит $\varepsilon = \int_{-T}^{T} |b(t)|^2\, dt$. Кроме того, из тождества $(bh_T)' = b'h_T + bh_T'$ (в смысле Радона-Никодима) при $|t| < T$ вытекает неравенство

$$|(b(t)h_T(t))'|^2 \ \leq \ 2\,|b'(t)|^2 + \frac{8\,|b(t)|^2}{T^2},$$

имеющее место почти всюду. Следовательно, полагая $\delta = \int_{-T}^{T} |b'(t)|^2\, dt$, второй интеграл в (8.6) можно оценить величиной $2\delta + \frac{8}{T^2}\,\varepsilon$, и в итоге получаем

$$\|bh_T\|_{\rm tv}^2 \ \leq \ \left(\frac{1}{2} + \frac{4}{T^2}\right)\varepsilon + \delta.$$

Вместе с (8.5) приходим к следующей теореме Эссеена [21].

Теорема 8.1. *Пусть A – комплекснозначная интегрируемая функция ограниченной вариации с преобразованием Фурье-Стилтьеса a. Для всех $T > 0$*

$$\int_{-\infty}^{\infty} |A(x)|\, dx \ \leq \ c_T\bigg(\int_{-T}^{T} \left|\frac{a(t)}{t}\right|^2 dt\bigg)^{1/2} + \bigg(\int_{-T}^{T} \left|\frac{d}{dt}\frac{a(t)}{t}\right|^2 dt\bigg)^{1/2} + c\,\frac{\|A\|_{\rm TV}}{T}, \tag{8.7}$$

где $c_T = (\frac{1}{2} + \frac{4}{T^2})^{1/2}$ *и* c *– постоянная. Можно положить* $c = 16\sqrt{\frac{2}{3}} < 13.07$.

В случае $A = F - G$ с функциями распределения F и G имеем $\|A\|_{\mathrm{TV}} \leq 2$, и мы приходим к уточнению неравенства (8.1).

Следствие 8.2. *Пусть* F *и* G *– функции распределения с характеристическими функциями* f *и* g. *Если* $W_1(F,G) < \infty$, *то для всех* $T \geq 3$

$$W_1(F,G) \leq \left(\int_{-T}^{T} \Big|\frac{f(t)-g(t)}{t}\Big|^2 dt\right)^{1/2} + \left(\int_{-T}^{T} \Big|\frac{d}{dt}\,\frac{f(t)-g(t)}{t}\Big|^2 dt\right)^{1/2} + \frac{c}{T}, \qquad (8.8)$$

где c *– постоянная. В частности, если* f *и* g *совпадают на* $[0,T]$, *то* $W_1(F,G) \leq \frac{c}{T}$.

Основываясь на (8.7), в последнем неравенстве можно положить $c = 32\sqrt{\frac{2}{3}}$. Эта константа может быть улучшена на основе вышеупомянутого результата Beurling'a; как было показано Эссееном [4], $W_1(F,G) \leq \frac{\pi}{T}$ в случае, когда характеристические функции f и g совпадают на интервале $[0,T]$.

Наконец, чтобы подчеркнуть связь теоремы 8.1 с теоремой 4.2, имеет смысл переформулировать неравенство (8.7), замечая, что $\frac{a(t)}{-it}$ является преобразованием Фурье функции A. После смены обозначений, получаем:

Следствие 8.3. *Для каждой интегрируемой функции ограниченной вариации* $b : \mathbf{R} \to \mathbf{C}$, *ее преобразование Фурье* $\widehat{b}$ *есть (локально) абсолютно непрерывная функция, причем для всех* $T > 0$

$$\int_{-\infty}^{\infty} |b(x)|\,dx \leq \left(\frac{1}{2} + \frac{4}{T^2}\right)^{1/2} \left(\int_{-T}^{T} |\widehat{b}(t)|^2\,dt\right)^{1/2} + \left(\int_{-T}^{T} \Big|\frac{d}{dt}\,\widehat{b}(t)\Big|^2 dt\right)^{1/2} + \frac{14}{T}\,\|b\|_{\mathrm{TV}}.$$

Устремляя $T \to \infty$, мы возвращаемся к эквивалентной форме неравенства (4.4) с точностью до абсолютного множителя.

9 Идеальные метрики Золотарева

В середине 1970-х гг. в связи с задачами о скорости сходимости в центральной предельной теореме В. М. Золотаревым были введены так называемые идеальные метрики в пространстве вероятностных распределений на $\mathbf{R}^k$. Среди них наиболее важными являются следующие (см. напр. [22], [23]); для простоты будем рассматривать только одномерный случай.

Зафиксируем целое число $s \geq 0$. Для вероятностных распределений F и G на вещественной прямой с конечными абсолютными моментами порядка s положим

$$\zeta_s(F,G) = \sup \Big|\int_{-\infty}^{\infty} u\,dF - \int_{-\infty}^{\infty} u\,dG\Big|, \qquad (9.1)$$

где супремум берется по всем функциям $u : \mathbf{R} \to \mathbf{R}$, имеющих $s-1$ производную, причем производная порядка $s-1$ удовлетворяет условию Липшица

$$|u^{(s-1)}(x) - u^{(s-1)}(y)| \leq |x-y| \qquad (x, y \in \mathbf{R})$$

(достаточно рассматривать s раз дифференцируемые u, такие что $|u^{(s)}| \le 1$).

В случае $s = 0$ приходим к расстоянию по вариации: $\zeta_0(F,G) = \|F - G\|_{\mathrm{TV}}$.

В случае $s = 1$, как легко видеть, мы возвращаемся к метрике в L^1:

$$\zeta_1(F,G) = \sup_{\|u\|_{\mathrm{Lip}} \le 1} \left| \int_{-\infty}^{\infty} u\,dF - \int_{-\infty}^{\infty} u\,dG \right| = \int_{-\infty}^{\infty} |F(x) - G(x)|\,dx.$$

В случае $s = 2$ имеет место аналогичная формула

$$\zeta_2(F,G) = \int_{-\infty}^{\infty} \left| \int_{-\infty}^{x} (F(y) - G(y))\,dy \right| dx.$$

В общей ситуации, полагая $b_1 = F - G$, имеет место рекуррентное соотношение

$$\zeta_s(F,G) = \int_{-\infty}^{\infty} |b_s(x)|\,dx, \qquad b_s(x) = \int_{-\infty}^{x} b_{s-1}(y)\,dy, \tag{9.2}$$

откуда получаем представление

$$\zeta_s(F,G) = \frac{1}{(s-2)!} \int_{-\infty}^{\infty} \left| \int_{-\infty}^{x} (F(y) - G(y))\,(x-y)^{s-2}\,dy \right| dx, \qquad s \ge 2.$$

Отметим, что для конечности $\zeta_s(F,G)$ достаточно, чтобы F и G имели совпадающие моменты порядка до $s-1$ включительно, то есть,

$$\int_{-\infty}^{\infty} x^p\,dF(x) = \int_{-\infty}^{\infty} x^p\,dG(x), \qquad p = 1, \dots, s-1, \tag{9.3}$$

и имели конечные абсолютные моменты $\int_{-\infty}^{\infty} |x|^s\,dF(x)$, $\int_{-\infty}^{\infty} |x|^s\,dG(x)$ (что предполагалось с самого начала).

Расстояние ζ_s однородно порядка s по пространственной переменной: если случайные величины X и Y имеют функции распределения F и G, то для функций распределения F_r, G_r случайных величин rX, rY $(r > 0)$, имеем

$$\zeta_s(F_r, G_r) \,=\, r^s\,\zeta_s(F,G).$$

Эти метрики взаимосвязаны между собой; в частности,

$$\zeta_1 \,\le\, C_s\,\zeta_{1+s}^{\frac{1}{1+s}}, \qquad \zeta_2^{1/2} \,\le\, C_s'\,\zeta_{2+s}^{\frac{1}{2+s}},$$

где постоянные зависят только от s (см. [9], теорема 3). Например, $\zeta_1^2 \le 8\zeta_2$ ([23]).

С помощью ζ_2 можно оценить расстояние Леви-Прохорова:

$$\pi^3(F,G) \le c\,\zeta_2(F,G)$$

для любых $F, G \in \mathcal{F}$ с некоторой постоянной c (см. [24]).

Метрики Золотарева определяются аналогично и для нецелых $s = m + \alpha$ (m – целое, $0 < \alpha < 1$). В этом случае супремум в (9.1) берется по всем m раз дифференцируемым функциям u, у которых производные порядка m удовлетворяют условию Липшица с показателем α:

$$|u^{(m)}(x) - u^{(m)}(y)| \le |x - y|^\alpha, \qquad x, y \in \mathbf{R}.$$

По формуле (9.1) определение $\zeta_s(F,G)$ распространяется и на любые функции ограниченной вариации F и G. Для конечности $\zeta_1(F,G)$ следует предполагать, что интеграл $\int_{-\infty}^{\infty} |F(x) - G(x)|\, dx$ конечен, что обеспечивается условием

$$\int_{-\infty}^{\infty} |x|\ |d(F(x) - G(x)| < \infty,$$

и тогда с необходимостью $F(-\infty) = G(-\infty)$ и $F(\infty) = G(\infty)$.

В этой более общей ситуации перейдем к задаче оценивания расстояния Золотарева в терминах преобразований Фурье-Стилтьеса

$$f(t) = \int_{-\infty}^{\infty} e^{itx}\, dF(x), \qquad g(t) = \int_{-\infty}^{\infty} e^{itx}\, dG(x) \qquad (t \in \mathbf{R}).$$

Для оценки $\zeta_1(F,G)$ можно применить теорему 8.1 или следствие 8.3: неравенство

$$\zeta_1(F,G) \le \bigg(\int_{-T}^{T} \Big|\frac{f(t) - g(t)}{t}\Big|^2 dt\bigg)^{1/2} + \bigg(\int_{-T}^{T} \Big|\frac{d}{dt}\, \frac{f(t) - g(t)}{t}\Big|^2 dt\bigg)^{1/2} + \frac{14}{T}\, \|F - G\|_{\rm TV}$$

имеет место при всех $T \ge 3$.

Если функция $b_2(x) = \int_{-\infty}^{x} (F(y) - G(y))\, dy$ интегрируема, то в равенстве

$$\frac{f(t) - g(t)}{-it} = \int_{-\infty}^{\infty} e^{itx}\, (F(x) - G(x))\, dx = \int_{-\infty}^{\infty} e^{itx}\, db_2(x)$$

можно проинтегрировать по частям, что дает

$$\frac{f(t) - g(t)}{(-it)^2} = \int_{-\infty}^{\infty} e^{itx}\, b_2(x)\, dx.$$

Применяя теперь следствие 8.3 к $b = b_2$ и учитывая что $\|b_2\|_{\rm TV} = \zeta_1(F,G)$, получаем

$$\zeta_2(F,G) \le \bigg(\int_{-T}^{T} \Big|\frac{f(t) - g(t)}{t^2}\Big|^2 dt\bigg)^{1/2} + \bigg(\int_{-T}^{T} \Big|\frac{d}{dt}\, \frac{f(t) - g(t)}{t^2}\Big|^2 dt\bigg)^{1/2} + \frac{14}{T}\, \zeta_1(F,G).$$

Заметим, что интегрируемость b_2 обеспечивается условиями

$$\int_{-\infty}^{\infty} x\, dF(x) = \int_{-\infty}^{\infty} x\, dG(x), \qquad \int_{-\infty}^{\infty} |x|^2\ |d(F(x) - G(x))| < \infty,$$

где, как обычно, $|d(F(x) - G(x))|$ обозначает вариацию $F - G$ (представляющую собой конечную положительную меру на вещественной прямой).

Продолжая этот процесс на основе рекуррентной формулы (9.2), приходим к аналогичным оценкам для произвольных целых $s \ge 1$.

Теорема 9.1. *Пусть F и G – функции ограниченной вариации с одинаковыми моментами порядка до $s-1$ включительно* (*условие* (9.3)) *и такие, что*

$$\int_{-\infty}^{\infty} |x|^s\ |d(F(x) - G(x)| < \infty.$$

Тогда для всех $T \geq 3$

$$\zeta_s(F,G) \, \leq \, \bigg(\int_{-T}^{T} \left|\frac{f(t)-g(t)}{t^s}\right|^2 dt\bigg)^{1/2} + \bigg(\int_{-T}^{T} \left|\frac{d}{dt}\,\frac{f(t)-g(t)}{t^s}\right|^2 dt\bigg)^{1/2} + \frac{c}{T}\,\zeta_{s-1}(F,G)$$

с некоторой абсолютной постоянной $c > 0$ (можно положить $c = 14$).

В классе функций распределений это неравенство доказано В. М. Золотаревым [25] (с точностью до абсолютной постоянной); без доказательства оно также приводится в книге [23], с. 80.

10 Транспортные метрики

Пусть (M,d) – полное сепарабельное метрическое пространство. Обозначим через $\mathcal{F}_p(M)$ пространство всех (борелевских) вероятностных мер F на M с конечным моментом порядка $p \geq 1$, то есть, удовлетворяющих условию

$$\int d(x,x_0)^p \, dF(x) < \infty$$

для некоторого (и, следовательно, для всех) $x_0 \in M$. Положим

$$W_p(F,G) = \inf \bigg(\iint d(x,y)^p \, d\mu(x,y)\bigg)^{1/p},$$

где инфимум берется по всем вероятностным мерам μ на $M \times M$ с маргинальными распределениями F и G, то есть, таким мерам, что

$$\mu(A \times M) = F(A), \qquad \mu(M \times A) = G(A)$$

для всех борелевских $A \subset M$. Функционал W_p превращает $\mathcal{F}_p(M)$ в метрическое пространство.

Согласно историческому исследованию А. М. Вершика [26], расстояние W_1 было введено Л. В. Канторовичем в конце 1930-х гг. Им рассматривались и более общие функционалы вида

$$W = \inf \iint c(x,y) \, d\mu(x,y)$$

со следующей интерпретацией: если стоимость перемещения “частицы"из точки x в точку y равна $c(x,y)$, то стоимость оптимального перемещения “массы"F в G будет равна W (см. [27-28]). Поэтому расстояния W_p также называются транспортными или минимальными. Мы отсылаем к работам [29-32], где подробно обсуждаются многие важные свойства и приложения этих метрик, и здесь отметим только некоторые.

Как непосредственно следует из определения, функция $p \rightarrow W_p(F,G)$ не убывает, и значит метрика W_p становится сильнее с ростом p. При этом расстояние Канторовича связано с расстоянием Леви-Прохорова неравенством

$$\pi(F,G) \leq \Big(W_p(F,G)\Big)^{p/(p+1)}. \tag{10.1}$$

Оно представляет собой аналог соотношения (7.1) между метрикой Леви и метрикой в пространстве L^p. Приведем аналогичное доказательство; предположим, что $F(A) \geq G(A_h) + h$ для некоторого $h > 0$ и некоторого борелевского множества A в M. Тогда для любой вероятностной меры μ на $M \times M$ с маргинальными распределениями F и G

$$\begin{aligned}\int_M \int_M d(x,y)^p \, d\mu(x,y) \ &\geq \ \int_A \int_{M \setminus A_h} d(x,y)^p \, d\mu(x,y) \\ &\geq \ h^p \, \mu\Big(A \times M \, \cap \, M \times (M \setminus A_h)\Big) \ \geq \ h^p \, (F(A) - G(A_h)) \ \geq \ h^{p+1}.\end{aligned}$$

Следовательно, $W_p^p(F,G) \geq h^{p+1}$, что и доказывает (10.1).

Из этого неравенства вытекает, что топология, порождаемая W_p на $\mathcal{F}_p(M)$ – более сильная по сравнению с топологией слабой сходимости. На самом деле имеется такая характеризация ([30]): $W_p(F_n, F) \to 0$ при $n \to \infty$ тогда и только тогда, когда $\pi(F_n, F) \to 0$, и при этом для некоторого (эквивалентно – для всех) $x_0 \in M$

$$\int d(x,x_0)^p \, dF_n(x) \ \to \ \int d(x,x_0)^p \, dF(x).$$

Поэтому на многих подпространствах пространства $\mathcal{F}_p(M)$ сходимость в метрике W_p равносильна слабой сходимости.

При $p = 1$ знаменитая теорема Канторовича-Рубинштейна дает дуальное описание метрики W_1 ([33], [34]): для всех $F, G \in \mathcal{F}_1(M)$

$$W_1(F,G) = \sup \bigg| \int u \, dF - \int u \, dG \bigg|,$$

где супремум берется по всем функциям $u : M \to \mathbf{R}$, удовлетворяющим условию Липшица $|u(x) - u(y)| \leq d(x,y)$, $x, y \in M$.

Следовательно, в случае вещественной прямой $M = \mathbf{R}$ с каноническим расстоянием $d(x,y) = |x - y|$ мы возвращаемся к метрике в пространстве L^1:

$$W_1(F,G) = \zeta_1(F,G) = \|F - G\|_1 = \int_{-\infty}^{\infty} |F(x) - G(x)| \, dx. \tag{10.2}$$

Здесь справа F и G – функции распределения, ассоциированные с соответствующими вероятностными мерами. Однако W_p при $p > 1$ не сводится к расстоянию в пространстве L^p; при этом имеется похожее описание

$$W_p(F,G) \ = \ \bigg(\int_0^1 |F^{-1}(t) - G^{-1}(t)|^p \, dt \bigg)^{1/p}$$

в терминах обобщенных обратных функций $F^{-1}(t) \ = \ \inf \{x \in \mathbf{R} : F(x) \geq t\}$.

Тождество (10.2) подсказывает, что возможна тесная связь между расстояниями Канторовича и Золотарева. Интересные результаты в этом направлении были недавно получены Э. Рио [35]. В частности, им установлено, что для любых вероятностных распределений F и G на прямой

$$W_p(F,G) \ \leq \ cp \, \zeta_p(F,G)^{1/p}, \qquad p \geq 1, \tag{10.3}$$

где c – абсолютная постоянная.

11 Квадратичное расстояние Канторовича

В иерархии метрик W_p с $p > 1$ частный случай $p = 2$ наиболее популярный; часто расстояние W_2 рассматривается как аналог евклидова расстояния в пространстве $\mathcal{F}_2(M)$. Поэтому вопросы, связанные с оцениванием этого расстояния, – предмет многих исследований. При этом задача о близости F и G в метрике W_2 в терминах преобразований Фурье-Стилтьеса остается неисследованной. Это замечание относится и к другим значениям $p \geq 1$ за исключением случая $p = 1$ и $M = \mathbf{R}$ (ввиду соотношения (10.2)).

Приведем один важный результат М. Талаграна [36], связывающий квадратичное расстояние Канторовича с расстоянием Кульбака-Лейблера (см. также [37]): для всех $F \in \mathcal{F}_2(\mathbf{R}^k)$

$$W_2^2(F, \Phi) \leq 2\, D(F||\Phi), \tag{11.1}$$

где Φ – гауссовская мера на евклидовом пространстве $\mathbf{R}^k$ с плотностью

$$\varphi(x) = \frac{1}{(2\pi)^{k/2}}\, e^{-||x||^2/2}.$$

В более развернутом виде, когда распределение F имеет плотность $p(x) = \frac{dF(x)}{dx}$ по мере Лебега, мы имеем соотношение

$$W_2^2(F, \Phi) \leq 2 \int_{\mathbf{R}^k} p(x)\, \log \frac{p(x)}{\varphi(x)}\, dx.$$

Неравенство (11.1) называют транспортно-энтропийным; мы отсылаем читателя к работе [38], где приведены и другие интересные соотношения для W_2.

Применяя (11.1) к сглаженным распределениям и используя оценку (5.1) для расстояния Кульбака-Лейблера (теорема 5.1), можно оценить $W_2(F, \Phi)$ в терминах характеристической функции f распределения F. Будем рассматривать одномерную ситуацию; в итоге получим следующее утверждение с участием преобразования Фурье-Стилтьеса

$$g_\alpha(t) = e^{-t^2/2}\Big(1 + \alpha\, \frac{(it)^3}{3!}\Big)$$

"поправленного" гауссовского распределения.

Теорема 11.1. *Пусть F – вероятностное распределение на числовой прямой с характеристической функцией f, причем $\int_{-\infty}^{\infty} |x|^3\, dF(x) < \infty$. Для любых $T \geq 1$ и $\alpha \in \mathbf{R}$*

$$\begin{aligned} W_2(F, \Phi) \ &\leq \ 4 \Big(\int_{-T}^{T} |f(t) - g_\alpha(t)|^2\, dt\Big)^{1/4} \\ &\quad + 4 \Big(\int_{-T}^{T} |f'''(t) - g_\alpha'''(t)|^2\, dt\Big)^{1/4} + c\,\Big(\frac{1 + Q_T^{1/4}}{T} + |\alpha|\Big), \end{aligned}$$

где $c > 0$ – абсолютная постоянная, и

$$Q_T = \int_{-T}^{T} \Big(|f''(t)| + |f'(t)| + |f(t)|\Big)^2\, (1 + t^4)\, dt.$$

В приложениях T естественно выбирать так, чтобы интеграл Q_T оставался ограниченным, и при этом α следует выбирать малым порядка $\frac{1}{T}$, но необязательно равным нулю; это может существенно уменьшить величину первых двух интегралов.

Для доказательства теоремы будем использовать обозначения $W_2(X,Y)$ для $W_2(F,G)$ и $D(X||Y)$ для $D(F||G)$, когда случайные величины X и Y имеют распределения F и G, соответственно.

Положим $X_\sigma = X + \sigma Y$, $\sigma = \frac{1}{T}$, предполагая, что Y не зависит от X и имеет симметричное распределение с характеристической функцией h, такой что $h(t) = 0$ при $|t| \geq 1$, причем $\mathbf{E}\,|Y|^3 \leq C^3$ $(C \geq 1)$. Можно взять нормализованную сверточную степень треугольной характеристической функции $h_0(t) = (1-|t|)_+$, например,

$$h(t) = h_0^{*6}(6t)/h_0^{*6}(0).$$

Заметим, что после сворачивания h_0 с собой достаточное число раз, мы получим положительно определенную функцию, дифференцируемую нужное число раз. Следовательно, после нормализации будем иметь дело с характеристической функцией случайной величины, у которой конечны абсолютные моменты нужного порядка, и при этом можно контролировать носитель характеристической функции.

По определению квадратичного расстояния Канторовича,

$$W_2^2(X_\sigma, X) \leq \mathbf{E}\,(X_\sigma - X)^2 = \sigma^2\,\mathbf{E}Y^2 \leq \frac{C^2}{T^2}. \tag{11.2}$$

По неравенству треугольника для метрики W_2, имеем

$$W_2(X,Z) \leq W_2(X_\sigma, Z) + W_2(X_\sigma, X),$$

где будем предполагать, что случайная величина Z имеет стандартное нормальное распределение Φ. Следовательно,

$$W_2^2(X,Z) \leq 2W_2^2(X_\sigma, Z) + 2W_2^2(X_\sigma, X),$$

и применяя неравенства (11.1)–(11.2), получаем неравенство типа сглаживания

$$W_2^2(X,Z) \,\leq\, 4\,D(X_\sigma||Z) + \frac{2C^2}{T^2}.$$

Чтобы оценить энтропийный член в этом неравенстве, применим (5.1) к распределению случайной величины X_σ. Она имеет характеристическую функцию $f_\sigma(t) = f(t)h(\sigma t)$, так что

$$D(X_\sigma||Z) \,\leq\, 4\left(\|f_\sigma - g_\alpha\|_2 + \|f_\sigma''' - g_\alpha'''\|_2\right) + \alpha^2,$$

и, следовательно,

$$W_2^2(X,Z) \,\leq\, 16\left(\|f_\sigma - g_\alpha\|_2 + \|f_\sigma''' - g_\alpha'''\|_2\right) + \frac{2C^2}{T^2} + 4\alpha^2. \tag{11.3}$$

Поскольку $f_\sigma(t) = 0$ при $|t| \geq T$,

$$\|f_\sigma - g_\alpha\|_2^2 \,=\, \int_{-T}^{T} |f_\sigma(t) - g_\alpha(t)|^2\,dt + \int_{|t|\geq T} |g_\alpha(t)|^2\,dt,$$

при этом последний интеграл убывает экспоненциально по T (и даже быстрее). То же самое относится к производным, так что

$$\int_{|t|\geq T} |g_\alpha(t)|^2\,dt \leq \frac{C_1(1+|\alpha|)^2}{T^4}, \qquad \int_{|t|\geq T} |g_\alpha'''(t)|^2\,dt \leq \frac{C_1(1+|\alpha|)^2}{T^4}$$

с некоторой постоянной $C_1 \geq 1$. Используя временное обозначение

$$\|u\|_T = \Big(\int_{-T}^{T} |u(t)|^2\,dt\Big)^{1/2},$$

из (11.3) получаем, что

$$\begin{aligned} W_2^2(X,Z) &\leq 16\,\|f_\sigma - g_\alpha\|_T + 16\,\|f_\sigma''' - g_\alpha'''\|_T + \frac{C_2(1+|\alpha|)}{T^2} + 4\alpha^2 \\ &\leq 16\,\|f_\sigma - g_\alpha\|_T + 16\,\|f_\sigma''' - g_\alpha'''\|_T + \frac{C_3}{T^2} + C_4\alpha^2 \end{aligned}$$

с некоторыми постоянными C_j. Более того, на интервале $[-T,T]$ можно аппроксимировать f_σ с помощью f в смысле L^2-нормы, и в силу неравенства треугольника для L^2-нормы,

$$\begin{aligned} W_2^2(X,Z) &\leq 16\,\|f - g_\alpha\|_T + 16\,\|f''' - g_\alpha'''\|_T + \frac{C_3}{T^2} + C_4\alpha^2 \\ &\quad + 16\,\|f_\sigma - f\|_T + 16\,\|f_\sigma''' - f'''\|_T. \end{aligned} \tag{11.4}$$

Остается оценить последние две нормы. Ввиду симметричности распределения случайной величины Y, имеем $h'(0) = i\,\mathbf{E}Y = 0$. Кроме того, $|h''(s)| \leq \mathbf{E}Y^2 \leq C^2$. Следовательно, по формуле Тейлора $|h(s)-1| \leq \frac{C^2}{2}\,s^2$ для всех s. Отсюда

$$|f_\sigma(t) - f(t)| \leq |f(t)|\,|h(\sigma t) - 1| \leq \frac{C^2\sigma^2}{2}\,t^2|f(t)|$$

и

$$\|f_\sigma - f\|_T \,\leq\, \frac{C^2\sigma^2}{2}\Big(\int_{-T}^{T} t^4|f(t)|^2\,dt\Big)^{1/2} \,\leq\, \frac{C^2}{2T^2}\,\sqrt{Q_T}. \tag{11.5}$$

Используя теперь $|h'(s)| \leq C|s|$, $|h^{(r)}(s)| \leq \mathbf{E}\,|Y|^r \leq C^r$ $(r = 2,3)$ и дифференцируя трижды $f_\sigma(t)$, получаем аналогичную поточечную оценку

$$\begin{aligned} |f_\sigma'''(t) - f'''(t)| &= |\,3\sigma f''(t)h'(\sigma t) + 3\sigma^2 f'(t)h''(\sigma t) + \sigma^3 f(t)h'''(\sigma t)| \\ &\leq 3C^3\sigma^2\Big(|tf''(t)| + |f'(t)| + |f(t)|\Big), \end{aligned}$$

откуда

$$\|f_\sigma''' - f'''\|_T \,\leq\, 3C^3\sigma^2\,\bigg(\int_{-T}^{T}\Big(|tf''(t)| + |f'(t)| + |f(t)|\Big)^2\,dt\bigg)^{1/2} \,\leq\, \frac{3C^3}{T^2}\,\sqrt{Q_T}.$$

Применяя это неравенство вместе с (11.5) в (11.4), приходим к требуемой оценке

$$W_2^2(X,Z) \,\leq\, 16\,\|f - g_\alpha\|_T + 16\,\|f''' - g_\alpha'''\|_T + \frac{C_3 + 4C^3\sqrt{Q_T}}{T^2} + C_4\alpha^2.$$

12 Сглаживающие меры с компактным носителем

Во всех предшествующих неравенствах, где использовались сглаживающие меры, преобразования Фурье-Стилтьеса этих мер имели компактный носитель. Но в некоторых задачах желательно, чтобы сами сглаживающие меры имели компактный носитель. В этом случае соответствующие преобразования Фурье-Стилтьеса могут быстро убывать (на бесконечности), но они уже не могут быть сосредоточены на конечном интервале. Вопрос о возможной скорости убывания был рассмотрен Ингманом, доказавшим следующую теорему ([39], см. также [40]). Мы ее сформулируем в несколько измененном виде.

Теорема 12.1. *Пусть задана неубывающая функция* $u : [1, \infty) \to [0, \infty)$*, такая что*

$$I = \int_1^\infty \frac{u(t)}{t}\, dt < \infty.$$

Для любого $c > 1$ *найдется симметричная вероятностная мера* H*, сосредоточенная на интервале* $[-cI, cI]$*, у которой характеристическая функция* f *удовлетворяет*

$$|f(t)| \le e^{-tu(t)\log c}, \qquad t \ge 4.$$

Полагая, например, $u(t) = \frac{\alpha}{et^\alpha}$ с параметром $0 < \alpha < 1$ и беря $c = e$, можно подобрать меру H на интервале $[-1, 1]$ с характеристической функцией, допускающей оценку

$$|f(t)| \le \exp\Big\{ -\frac{\alpha}{e}\, t^{1-\alpha}\Big\}, \qquad t \ge 4.$$

Другой пример $u(t) = \frac{\kappa}{\log^2(1+t)}$ с подходящим значением $\kappa > 0$ приводит к еще более быстрому – почти экспоненциальному убыванию

$$|f(t)| \le \exp\bigg\{ -\frac{\log 2}{2e}\, \frac{t}{\log^2(1+t)}\bigg\}, \qquad t \ge 4,$$

и при этом H будет снова сосредоточена на $[-1, 1]$.

Однако, как известно, при условии на компактность носителя неравенство вида $|f(t)| \le Ce^{-ct}$ $(t \ge t_0)$ получить невозможно (с какими-либо положительными постоянными c и C). Тем не менее, в интегральном смысле такая экспоненциальная оценка возможна ввиду следующего элементарного утверждения.

Теорема 12.2. *Для любого* $T \ge 0$ *найдется симметричная вероятностная мера* H*, сосредоточенная на интервале* $[-1, 1]$*, у которой характеристическая функция* f *удовлетворяет*

$$\int_{|t| \ge T} |f(t)|\, dt \ \le\ 2\pi e^2\, e^{-\frac{T}{e}}. \tag{12.1}$$

При этом, для $T \ge 1$

$$\int_{|t| \ge T} |f(t)|\, \frac{dt}{|t|} \ \le\ 3\, e^{-\frac{T}{e}}. \tag{12.2}$$

Как было предложено Ингманом, в качестве меры H в теореме 12.1 можно взять распределение сходящегося случайного ряда

$$S = c \sum_{n=2}^{\infty} \frac{u(n)}{n} X_n,$$

в котором случайные величины X_n независимы и равномерно распределены на $(-1,1)$.

Для доказательства теоремы 12.2 достаточно взять конечную сумму

$$S_n = \frac{1}{n} \sum_{k=1}^{n} X_k \qquad (n \geq 2)$$

с $n = [\frac{T}{e}]$ при $T \geq 2e$. В этом случае $2 \leq n \leq \frac{T}{e}$, и S_n имеет характеристическую функцию $f(t) = \mathbf{E}\, e^{itS_n} = (\frac{\sin(t/n)}{t/n})^n$. Поэтому

$$\begin{aligned} \int_T^{\infty} |f(t)|\, dt \ &\leq \ \int_T^{\infty} \left(\frac{n}{t}\right)^n dt \ = \ \frac{T}{n-1}\, e^{-n \log \frac{T}{n}} \\ &\leq \ \frac{T}{[\frac{T}{e}]-1}\, e^{-[\frac{T}{e}]} \ \leq \ \frac{Te}{[\frac{T}{e}]-1}\, e^{-\frac{T}{e}} \ \leq \ 2e^2\, e^{-\frac{T}{e}}. \end{aligned}$$

В случае $0 \leq T \leq 2e$ можно взять $n = 2$ и применить обратное преобразование Фурье к функции $f(t) = (\frac{\sin(t/2)}{t/2})^2$. Так как S_2 имеет треугольную плотность $p(x) = (1 - |x|)_+$, получаем $\int_0^\infty f(t)\, dt = \pi p(0) = \pi$. Следовательно,

$$\int_0^{\infty} f(t)\, dt = \pi \leq \pi e^2\, e^{-\frac{T}{e}}.$$

Объединяя оба случая, приходим к неравенству (12.1).

Аналогично при $T \geq 2e$

$$\int_T^{\infty} |f(t)|\, \frac{dt}{t} \ \leq \ \int_T^{\infty} \left(\frac{n}{t}\right)^n \frac{dt}{t} \ = \ \frac{1}{n}\, e^{-n \log \frac{T}{n}} \ \leq \ \frac{1}{2}\, e^{-[\frac{T}{e}]} \ \leq \ \frac{e}{2}\, e^{-\frac{T}{e}}.$$

В случае $1 \leq T \leq 2e$ можно взять $n = 1$, что дает

$$\int_T^{\infty} |f(t)|\, \frac{dt}{t} \ \leq \ \int_T^{\infty} \frac{1}{t^2}\, d \ = \ \frac{1}{T} \ \leq \ \frac{3}{2}\, e^{-\frac{T}{e}}$$

(используем $e^{1/e} < 1.5$). Объединяя оба случая, приходим к (12.2).

Распределения H из теоремы 12.2 применялись В. М. Золотаревым в доказательстве неравенства для расстояния Леви (теорема 3.2).

13 Знакопеременные сглаживающие меры

Перейдем к сглаживающим мерам с дополнительными свойствами, хотя и без сохранения свойства положительности.

Теорема 13.1. *Пусть задано целое число $s \geq 1$. Для любого $T \geq 1$ найдется симметричная знакопеременная мера R на интервале $[-1,1]$ с полной вариацией $\|R\|_{\mathrm{TV}} \leq c_s$, такая что*

$$R([-1,1]) = 1, \qquad \int_{-1}^{1} x^p \, dR(x) = 0 \quad (p = 1, \dots, s-1), \tag{13.1}$$

причем с преобразованием Фурье-Стилтьеса f, удовлетворяющим неравенству

$$\int_{|t| \geq T} |f(t)| \, \frac{dt}{|t|} \, \leq \, 3c_s \, e^{-\frac{T}{e}}. \tag{13.2}$$

Можно положить $c_1 = c_2 = 1$, $c_3 = 3$ и $c_s = s\, C_{2s}^{s-1}$ при $s \geq 4$.

В случаях $s = 1$ и $s = 2$ условие (13.1) выполняется автоматически для вероятностной меры $R = H$ из теоремы 12.2 (в силу симметричности H; поэтому $c_1 = c_2 = 1$). При $s = 3$ это условие выполняется для $p = 1$, но нарушается для $p = 2$. Чтобы оно выполнялось, необходимо отказаться от свойства положительности R как меры, то есть, от монотонности ассоциированной с ней функции $R(x) = R((-\infty, x])$ и взять, например,

$$R(x) = 2H(x\sqrt{2}) - H(x).$$

В этом случае $\|R\|_{\mathrm{TV}} \leq 3$, и, следовательно, можно положить $c_3 = 3$.

В общей ситуации пусть

$$R(x) = w_1 H\Big(\frac{x}{b_1}\Big) + \ldots + w_s H\Big(\frac{x}{b_s}\Big),$$

где H – функция распределения из теоремы 12.2 (которую мы отождествляем с мерой H), причем $b_i \neq b_j$ $(i \neq j)$. Если все $b_i \in (0,1]$, то R как мера будет сосредоточена на интервале $[-1,1]$ с полной вариацией

$$\|R\|_{\mathrm{TV}} \leq \sum_{i=1}^{s} |w_i|. \tag{13.3}$$

Преобразование Фурье-Стилтьеса f меры R выражается через характеристическую функцию h меры H равенством $f(t) = \sum_{i=1}^{s} w_i \, h(b_i t)$. Поэтому, по теореме 12.2

$$\int_{|t| \geq T} |f(t)| \, \frac{dt}{|t|} \, = \, \sum_{i=1}^{s} |w_i| \int_{|t| \geq T/b_i} |h(t)| \, \frac{dt}{|t|} \, \leq \, 3 \sum_{i=1}^{s} |w_i| \, e^{-\frac{T}{e}}. \tag{13.4}$$

Перейдем теперь к условию (13.1). Оно сводится к линейному уравнению с s неизвестными $w = (w_1, \dots, w_s)$, которые можно записать в матричной форме как $Vw = e$, где V – матрица Вандермонда

$$V \, = \, \begin{pmatrix} 1 & 1 & \cdots & 1 \\ b_1 & b_2 & \cdots & b_s \\ \vdots & \vdots & \ddots & \vdots \\ b_1^{s-1} & b_2^{s-1} & \cdots & b_s^{s-1} \end{pmatrix},$$

и $e = (1, 0, \dots, 0)$ (как столбец). Отсюда находим $w = V^{-1}e$, и можно оценить правую часть (13.3) в терминах b_i. Для этого воспользуемся следующим результатом, полученным в [41], о норме матрицы, обратной к матрице Вандермонда. Именно, если норму $s \times s$–матрицы $A = (a_{ij})$ определить равенством

$$\|A\| = \max_{1 \le i \le s} \sum_{j=1}^{s} |a_{ij}|,$$

то для $A = V^{-1}$ имеем

$$\|V^{-1}\| \,\le\, \max_{1 \le i \le s} \prod_{j \ne i} \frac{1 + |b_j|}{|b_i - b_j|}. \tag{13.5}$$

Например, при выборе $b_i = \frac{i}{s}$ получаем верхнюю оценку

$$\|V^{-1}\| \,\le\, \prod_{j=2}^{s} \frac{1 + \frac{j}{s}}{\frac{j}{s} - \frac{1}{s}} \,=\, C_{2s}^{s-1}.$$

Так как $w_i = (Ae)_i = a_{i1}$, то

$$\sum_{i=1}^{s} |w_i| \,=\, \sum_{i=1}^{s} |a_{i1}| \,\le\, \sum_{i=1}^{s} \sum_{j=1}^{s} |a_{ij}| \,\le\, s\|A\| \,\le\, sC_{2s}^{s-1}.$$

С учетом (13.3)-(13.4) приходим к (13.2) вместе с неравенством $\|R\|_{\mathrm{TV}} \le s\, C_{2s}^{s-1}$, что и доказывает теорему.

14 Аналог неравенства Эссеена для L^1-метрики

Приведем пример, где может быть использована теорема 12.2. В частности, нас интересует один вариант теоремы 8.1, в которой L^2-норма функции $a(t)/t$ на интервале $[-T, T]$ заменялась бы на L^1-норму и при этом удалялся интеграл с производной.

Докажем следующее утверждение. Пусть $A : \mathbf{R} \to \mathbf{R}$ – функция ограниченной вариации с преобразованием Фурье-Стилтьеса

$$a(t) = \int_{-\infty}^{\infty} e^{itx}\, dA(x) \qquad (t \in \mathbf{R}),$$

причем $A(-\infty) = 0$.

Теорема 14.1. *Если* $1 \le \beta - \alpha \le T \log T$ $(T \ge 1)$, *то*

$$\int_{\alpha}^{\beta} |A(x)|\, dx \,\le\, \frac{\beta - \alpha}{2\pi} \int_{-T}^{T} \Big|\frac{a(t)}{t}\Big|\, dt + c\|A\|_{\mathrm{TV}}\, \frac{\log T}{T}, \tag{14.1}$$

где $c > 0$ *– абсолютная постоянная.*

В несколько ослабленной форме, когда $A = F - G$ – разность двух функций распределения, такая оценка может быть получена с помощью неравенства Золотарева для расстояния Леви (теорема 3.2). Действительно, если $h > L(F, G)$, то для всех x

$$F(x) - G(x) \le (G(x + h) - G(x)) + h,$$

$$G(x) - F(x) \le (F(x+h) - F(x)) + h,$$

откуда

$$|F(x) - G(x)| \,\le\, (F(x+h) - F(x)) + (G(x+h) - G(x)) + h.$$

Интегрируя это неравенство по промежутку $[\alpha, \beta]$ и затем устремляя $h \to L(F, G)$, получаем

$$\int_\alpha^\beta |F(x) - G(x)|\, dx \,\le\, (1 + (\beta - \alpha))\, L(F, G).$$

Поэтому неравенство (3.3) дает при всех $T > 1.3$

$$\int_\alpha^\beta |F(x) - G(x)|\, dx \,\le\, \frac{1 + (\beta - \alpha)}{2\pi} \int_{-T}^{T} \Big|\frac{f(t) - g(t)}{t}\Big|\, dt + 2e(1 + (\beta - \alpha))\, \frac{\log T}{T},$$

где f и g – характеристические функции распределений F и G. Таким образом, преимущество (14.1) – в отсутствии множителя $\beta - \alpha$ в правой части.

Для доказательства теоремы 14.1 нам потребуется одно вспомогательное неравенство, имеющее самостоятельный интерес.

Лемма 14.2. *Пусть $A : [\alpha, \beta] \to \mathbf{R}$ – функция ограниченной вариации. Для любого целого числа $N \ge 1$*

$$\int_\alpha^\beta |A(x)|\, dx \,\le\, \sum_{k=1}^N \Big| \int_{x_{k-1}}^{x_k} A(x)\, dx \Big| + \frac{\beta - \alpha}{N}\, \|A\|_{\mathrm{TV}}, \tag{14.2}$$

где $x_k = \alpha + (\beta - \alpha)\,\frac{k}{N}$.

Для разностей функций распределения (с заменой $\|A\|_{\mathrm{TV}}$ на 2) это неравенство доказано в [42]. Общий случай аналогичен; обозначим через I совокупность индексов $k = 1, \dots, N$, таких что в k-ом интервале $\Delta_k = (x_{k-1}, x_k)$ функция $A(x)$ не меняет знак. Остальные индексы образуют дополнительное подмножество $J \subset \{1, \dots, N\}$. Тогда для всех $k \in I$

$$\int_{\Delta_k} |A(x)|\, dx = \Big| \int_{\Delta_k} A(x)\, dx \Big|.$$

Если же $k \in J$, то, очевидно,

$$\sup_{x \in \Delta_k} |A(x)| \,\le\, \sup_{x, y \in \Delta_k} |A(x) - A(y)| \,\le\, |A|(\Delta_k),$$

где $|A|$ – вариация функции A, рассматриваемая как положительная мера на интервале $[\alpha, \beta]$. В этом случае

$$\int_{\Delta_k} |A(x)|\, dx \le |A|(\Delta_k)\, |\Delta_k|, \qquad |\Delta_k| = \frac{\beta - \alpha}{N}.$$

Комбинируя обе оценки, заключаем, что интеграл $\int_\alpha^\beta |A(x)|\, dx$ не превосходит

$$\sum_{k \in I} \Big| \int_{\Delta_k} A(x)\, dx \Big| + \sum_{k \in J} |A|(\Delta_k) |\Delta_k| \,\le\, \sum_{k=1}^N \Big| \int_{\Delta_k} A(x)\, dx \Big| + \frac{\beta - \alpha}{N} \sum_{k=1}^N |A|(\Delta_k).$$

В силу аддитивности меры $|A|$, последняя сумма не превосходит $\|A\|_{\mathrm{TV}}$.

Лемма доказана.

Перейдем к выводу неравенства (14.1), для чего нам потребуется оценить интеграл в правой части (14.2). Можно предполагать, что $\int_{-T}^{T} |\frac{a(t)}{t}|\,dt < \infty$. Рассмотрим функцию вида

$$U(x) = \frac{1}{2h}\int_{x-h}^{x+h} A(y)\,dy \qquad (x \in \mathbf{R},\ h > 0),$$

представляющую собой свертку меры A с равномерным распределением на интервале $(-h, h)$ (параметр h будет выбран позднее). Она имеет преобразование Фурье-Стилтьеса $\frac{\sin(th)}{th}\,a(t)$.

Фиксируя еще один параметр $\sigma > 0$, который будет выбран в зависимости от T, рассмотрим свертку

$$U_\sigma(x) = \int_{-\infty}^{\infty} U(x - \sigma y)\,dH(y),$$

где H – вероятностная мера из теоремы 12.2. Так как эта мера сосредоточена на $[-1, 1]$, сразу получаем

$$|U_\sigma(x) - U(x)| \ \le\ \sup_{|z|\le\sigma} |U(x - z) - U(x)| \ \le\ \frac{\sigma}{2h}\,|A|(x - h - \sigma, x + h + \sigma), \tag{14.3}$$

где последний член обозначает вариацию функции A на $(x - h - \sigma, x + h + \sigma)$.

С другой стороны, U_σ имеет преобразование Фурье-Стилтьеса

$$g(t) = f(\sigma t)\frac{\sin(th)}{th}\,a(t),$$

где f – характеристическая функция меры H; при этом $U_\sigma(-\infty) = 0$. Функция g интегрируема, и, следовательно, можно применить обратное преобразование Фурье-Стилтьеса: для всех x

$$U_\sigma(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\,\frac{g(t)}{-it}\,dt.$$

Используя $\sup_t |a(t)| \le \|A\|_{\mathrm{TV}}$ за пределами интервала $[-T, T]$, отсюда получаем

$$|U_\sigma(x)| \le \frac{1}{2\pi}\int_{-T}^{T}\Big|\frac{a(t)}{t}\Big|\,dt + \frac{1}{2\pi}\,\|A\|_{\mathrm{TV}}\int_{|t|\ge\sigma T}\Big|\frac{f(t)}{t}\Big|\,dt.$$

Положим здесь $\sigma = \frac{2e\log T}{T}$ и применим неравенство (12.2) к $2e\log T$ (вместо T), чтобы оценить последний интеграл. Тогда приходим к оценке

$$|U_\sigma(x)| \le \frac{1}{2\pi}\int_{-T}^{T}\Big|\frac{a(t)}{t}\Big|\,dt + \frac{3}{2\pi}\ \frac{\|A\|_{\mathrm{TV}}}{T^2}$$

и из (14.3) получаем

$$|U(x)| \ \le\ \frac{1}{2\pi}\int_{-T}^{T}\Big|\frac{a(t)}{t}\Big|\,dt + \frac{3}{2\pi}\ \frac{\|A\|_{\mathrm{TV}}}{T^2} + \frac{\sigma}{2h}\,|A|(x - \sigma - h, x + \sigma + h).$$

Таким образом, по определению функции U и с учетом выбора σ, для всех $x \in \mathbf{R}$, $h > 0$ и $T \geq 1$ имеем

$$\Big|\int_{x-h}^{x+h} A(y)\,dy\Big| \ \leq \ \frac{h}{\pi}\int_{-T}^{T}\Big|\frac{a(t)}{t}\Big|\,dt + \frac{3}{\pi}\,\frac{h}{T^2}\,\|A\|_{\rm TV} + \frac{2e\log T}{T}\,\varepsilon\Big(x, h + \frac{2e\log T}{T}\Big), \qquad (14.4)$$

где используем обозначение

$$\varepsilon(x,r) = |A|(x-r, x+r).$$

Теперь вернемся к лемме 14.2, используя такое же разбиение $[\alpha,\beta]$ на интервалы $\Delta_k = (x_{k-1}, x_k)$ с концами $x_k = \alpha + (\beta-\alpha)\,\frac{k}{N}$. Применяя (14.4) к точкам $z_k = \frac{x_{k-1}+x_k}{2}$ с $h = \frac{\beta-\alpha}{2N}$ и производя суммирование по всем $k = 1,\dots,N$, неравенство (14.2) дает

$$\begin{aligned}\int_\alpha^\beta |A(x)|\,dx \ \ &\leq \ \ \frac{\beta-\alpha}{2\pi}\int_{-T}^{T}\Big|\frac{a(t)}{t}\Big|\,dt + \frac{3}{2\pi}\,\frac{\beta-\alpha}{T^2}\,\|A\|_{\rm TV} \\ &\quad + \frac{\beta-\alpha}{N}\,\|A\|_{\rm TV} + \frac{2e\log T}{T}\,\sum_{k=1}^N \varepsilon\Big(z_k, h + \frac{2e\log T}{T}\Big). \qquad (14.5)\end{aligned}$$

Заметим, что $\sum_{k=1}^N \varepsilon(z_k, lh) \leq l\,\|A\|_{\rm TV}$ для любого целого $l \geq 1$. Поэтому сумма в (14.5) не превосходит

$$\Big(\frac{2e\log T}{Th} + 2\Big)\,\|A\|_{\rm TV}.$$

Следовательно, с учетом того, что $\beta - \alpha \leq T\log T$, приходим к оценке

$$\begin{aligned}\int_\alpha^\beta |A(x)|\,dx \ \ &\leq \ \ \frac{\beta-\alpha}{2\pi}\int_{-T}^{T}\Big|\frac{a(t)}{t}\Big|\,dt \\ &\quad + \|A\|_{\rm TV}\Big[\frac{3}{2\pi}\,\frac{\log T}{T} + \frac{\beta-\alpha}{N} + \frac{2e\log T}{T}\,\Big(\frac{2e\log T}{Th} + 2\Big)\Big]. \qquad (14.6)\end{aligned}$$

Полагая $N = [(\beta-\alpha)\,T/\log T] + 1$, имеем $\frac{\beta-\alpha}{N} \leq \frac{\log T}{T}$ и, если $\beta - \alpha \geq 1$, то также

$$Th \ = \ T\,\frac{\beta-\alpha}{2N} \ \geq \ T\,\frac{\beta-\alpha}{2\,\frac{(\beta-\alpha)\,T}{\log T} + 2} \ \geq \ \frac{1}{3}\,\log T.$$

Таким образом, с точностью до множителя $\frac{\log T}{T}$ выражение в квадратных скобках из правой части (14.6) не превосходит $\frac{3}{2\pi} + 1 + 2e(6e+2) < 102$, и мы приходим к (14.1) с константой $c = 102$. Теорема 14.1 доказана.

15 Варианты неравенства Берри-Эссеена

Снова пусть $A : \mathbf{R} \rightarrow \mathbf{R}$ – функция ограниченной вариации, $A(-\infty) = 0$. Возвратимся к оценкам для L^∞-нормы

$$\|A\| = \sup_x |A(x)|$$

в терминах преобразования Фурье-Стилтьеса

$$a(t) = \int_{-\infty}^\infty e^{itx}\,dA(x).$$

Полагая $F = 0$, $G = A$, неравенство Берри-Эссеена (теорема 2.1) дает оценку

$$\|A\| \le c \int_{-T}^{T} \Big|\frac{a(t)}{t}\Big|\, dt + c'\, \frac{L_1}{T} \qquad (T > 0) \tag{15.1}$$

с абсолютными постоянными $c, c' > 0$ в предположении, что $|A'(x)| \le L_1$ для всех x. Интересно, что при усилении свойства гладкости в терминах производных функции A более высокого порядка неравенство (15.1) может быть существенно улучшено. В частности, справедливо такое утверждение.

Теорема 15.1. *Если A дважды дифференцируема, причем $\sup_x |A''(x)| \le L_2$, то для всех $T > 0$*

$$\|A\| \le c \int_{-T}^{T} \Big|\frac{a(t)}{t}\Big|\, dt + c'\, \frac{L_2}{T^2}, \tag{15.2}$$

где $c, c' > 0$ – абсолютные постоянные.

Можно также поинтересоваться, нельзя ли сделать эту оценку локальной, когда свойства гладкости функции A известны лишь в окрестности заданной точки x. В такой ситуации можно получить аналогичное утверждение.

Теорема 15.2. *Предположим, что функция A дифференцируема s раз в окрестности $\Delta : |z - x| < se\, \frac{\log T}{T}$ данной точки x (где $T > 1$ и $s \ge 1$ – целое), причем*

$$\sup_{z \in \Delta} |A^{(s)}(z)| \le L_s(x, T). \tag{15.3}$$

Тогда с некоторой постоянной c_s, зависящей только от s,

$$|A(x)| \,\le\, \frac{1}{2\pi} \int_{-T}^{T} \Big|\frac{a(t)}{t}\Big|\, dt + c_s\, \|A\|_{\mathrm{TV}}\, \frac{1}{T^s} + c_s L_s(x, T) \left(\frac{\log T}{T}\right)^s. \tag{15.4}$$

Следует подчеркнуть, что подобные оценки вряд ли имеют ценность в таких задачах, как скорость сходимости в центральной предельной теореме, где рассматриваются функции $A = F - G$ с гладкими G, но, вообще говоря, с разрывными F. Тем не менее, для таких целей можно получить “дискретные"аналоги неравенств (15.2) и (15.4), причем условие (15.3) следует формулировать в терминах разностных операторов порядка s. На этом пути можно исследовать, например, разложения Эджворта для биномиального распределения. Ввиду ограниченности места, мы не затрагиваем такого рода обобщения.

Начнем с доказательства теоремы 15.2, предполагая, что $\int_{-T}^{T} |\frac{a(t)}{t}|\, dt < \infty$. Неравенство (15.4) основано на сглаживании A с помощью знакопеременной меры R из теоремы 13.1: рассмотрим свертку

$$A_\sigma(x) = \int_{-\infty}^{\infty} A(x - \sigma y)\, dR(y) \tag{15.5}$$

с параметром $\sigma = \frac{se \log T}{T}$. Поскольку R сосредоточена на интервале $[-1, 1]$, причем $R([-1, 1]) = 1$, имеем

$$A_\sigma(x) - A(x) = \int_{-1}^{1} (A(x - \sigma y) - A(x))\, dR(y).$$

Разлагая $A(x-\sigma y)$ по формуле Тейлора по степеням y и используя свойство (13.1), получаем

$$|A_\sigma(x) - A(x)| \,\le\, \frac{\sigma^s}{s!}\,\sup_{z\in\Delta}|A^{(s)}(z)| \,\le\, \frac{\sigma^s}{s!}\,L_s(x,T). \tag{15.6}$$

С другой стороны, A_σ имеет преобразование Фурье-Стилтьеса $g(t) = f(\sigma t)\,a(t)$, где f – преобразование Фурье-Стилтьеса меры R; при этом $A_\sigma(-\infty) = 0$. Следовательно, можно применить обратное преобразование Фурье-Стилтьеса:

$$A_\sigma(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\,\frac{g(t)}{-it}\,dt. \tag{15.7}$$

Используя $\sup_t |a(t)| \,\le\, \|A\|_{\rm TV}$ за пределами $[-T,T]$ и затем применяя неравенство (13.2), получаем

$$\begin{aligned} |A_\sigma(x)| \ &\le\ \frac{1}{2\pi}\int_{-T}^{T}\Big|\frac{a(t)}{t}\Big|\,dt + \frac{\|A\|_{\rm TV}}{2\pi}\int_{|t|\ge\sigma T}\Big|\frac{f(t)}{t}\Big|\,dt \\ &\le\ \frac{1}{2\pi}\int_{-T}^{T}\Big|\frac{a(t)}{t}\Big|\,dt + \frac{\|A\|_{\rm TV}}{2\pi}\,\frac{c_s}{T^s} \end{aligned}$$

с постоянной c_s из теоремы 13.1.

Остается скомбинировать это неравенство с (15.6):

$$|A(x)| \,\le\, \frac{1}{2\pi}\int_{-T}^{T}\Big|\frac{a(t)}{t}\Big|\,dt + \frac{\|A\|_{\rm TV}}{2\pi}\,\frac{c_s}{T^s} + \frac{(se)^s}{s!}\,L_s(x,T)\,\Big(\frac{\log T}{T}\Big)^s.$$

Так как $c_1 = c_2 = 1$, получаем, в частности,

$$\begin{aligned} |A(x)| \ &\le\ \frac{1}{2\pi}\int_{-T}^{T}\Big|\frac{a(t)}{t}\Big|\,dt + \frac{\|A\|_{\rm TV}}{2\pi}\,\frac{1}{T} + e\,L_1(x,T)\,\frac{\log T}{T}, \\ |A(x)| \ &\le\ \frac{1}{2\pi}\int_{-T}^{T}\Big|\frac{a(t)}{t}\Big|\,dt + \frac{\|A\|_{\rm TV}}{2\pi}\,\frac{1}{T^2} + 2e^2 L_2(x,T)\,\Big(\frac{\log T}{T}\Big)^2. \end{aligned}$$

В общем случае $c_s \le s\,C_{2s}^{s-1}$, и мы приходим к требуемому неравенству (15.4) с $c_s = c^s$, где c – абсолютная постоянная.

Доказательство (15.2) несколько проще; можно следовать стандартным рассуждениям, используемым при выводе неравенства Берри-Эссеена. В качестве сглаживающей меры в (15.5) возьмем вероятностную меру R с треугольной характеристической функцией $f(t) = (1-|t|)_+$, то есть, с плотностью

$$\psi(x) = \frac{1}{2\pi}\left(\frac{\sin\frac{x}{2}}{\frac{x}{2}}\right)^2$$

и положим $\sigma = 1/T$. Тогда снова имеем тождество (15.7), откуда сразу следует, что

$$\|A_\sigma\| \,\le\, \frac{1}{2\pi}\int_{-\infty}^{\infty}\Big|\frac{a(t)}{t}\Big|\,f(\sigma t)\,dt \,\le\, \frac{1}{2\pi}\int_{-T}^{T}\Big|\frac{a(t)}{t}\Big|\,dt. \tag{15.8}$$

С другой стороны, отталкиваясь от (15.5), для фиксированного $l>0$ запишем

$$A_\sigma(x) - A(x) \,=\, \int_{-l}^{l}(A(x-\sigma y) - A(x))\,dR(y) + \int_{|y|>l}(A(x-\sigma y) - A(x))\,dR(y).$$

Разлагая $A(x-\sigma y)$ по формуле Тейлора до квадратичного члена и используя симметричность меры R, для первого интеграла легко видеть, что

$$\Big|\int_{-l}^{l}(A(x-\sigma y)-A(x))\,dR(y)\Big| \ \le\ \int_{-l}^{l}\frac{(\sigma y)^2}{2}\,L_2\,dR(y)\ =\ \frac{l^3}{3}\,\frac{L_2}{T^2}.$$

Модуль второго интеграла просто оценим величиной $2\gamma\,\|A\|$, где $\gamma = 1-R[-l,l]$. Применяя (15.8), получаем неравенство

$$\|A\|\ \le\ \frac{1}{2\pi}\int_{-T}^{T}\Big|\frac{a(t)}{t}\Big|\,dt+\frac{l^3}{3}\,\frac{L_2}{T^2}+2\gamma\,\|A\|,$$

которое в случае $\gamma<\frac{1}{2}$ дает искомую оценку

$$\|A\|\ \le\ \frac{1}{2\pi\,(1-2\gamma)}\int_{-T}^{T}\Big|\frac{a(t)}{t}\Big|\,dt+\frac{l^3}{3\,(1-2\gamma)}\,\frac{L_2}{T^2}.$$

Чтобы зафиксировать численные значения постоянных, положим, например, $l=\frac{3\pi}{2}$. В этом случае

$$\begin{aligned}2\gamma\ &=\ 4\int_{\frac{3\pi}{2}}^{\infty}\psi(x)\,dx\ =\ \frac{4}{\pi}\int_{\frac{3\pi}{2}}^{\infty}\frac{1-\cos x}{x^2}\,dx\\ &\le\ \frac{4}{\pi}\int_{\frac{3\pi}{2}}^{\frac{5\pi}{2}}\frac{1}{x^2}\,dx+\frac{4}{\pi}\int_{\frac{5\pi}{2}}^{\infty}\frac{2}{x^2}\,dx\ =\ \frac{64}{15\pi^2}\ =\ 0.4323...\end{aligned}$$

Следовательно, (15.2) имеет место с постоянными $c=\frac{1}{2\pi(1-2\gamma)}=0.2804...$ и $c'=\frac{l^3}{3\,(1-2\gamma)}=\frac{9\pi^3}{8\,(1-2\gamma)}=61.44...$

16 Сглаживание с полиномиальным весом

Вернемся теперь к функциям ограниченной вариации вида $A=F-G$ с неубывающими функциями F, причем допускаются условия типа гладкости на G. Нас теперь будет интересовать обобщение теоремы 2.1 для расстояния Колмогорова с полиномиальным весом, а именно, для

$$\rho_s(F,G)\ =\ \sup_x\,(1+|x|^s)\,|F(x)-G(x)|,$$

где $s\ge 0$ – заданное целое число (так что $\rho_0=\rho$).

Необходимость изучения таких расстояний объясняется важностью неравномерных оценок

$$|F(x)-G(x)|\le\frac{c}{1+|x|^s},\qquad x\in\mathbf{R}, \tag{16.1}$$

в которых можно положить $c=\rho_s(F,G)$. Например, частный случай $s=2$ позволяет оценить расстояние в метрике L^1: интегрируя неравенство (16.1) по всей прямой, приходим к соотношению

$$W_1(F,G)\le\pi\rho_2(F,G).$$

Оно сохраняется и для расстояния в метрике L^p при всех $p \geq 1$. Заметим, что при $p > 1$ можно применить (16.1) и с $s = 1$ и тогда получим

$$\|F - G\|_p \leq \left(\frac{2}{p-1}\right)^{1/p} \rho_1(F, G).$$

Прежде, чем оценивать ρ_s в терминах соответствующих преобразований Фурье-Стилтьеса функций F и G, в этом разделе мы зафиксируем одно общее соотношение типа (2.2) между расстоянием $\rho_s(F, G)$ и L^∞-нормой сглаженной функции

$$A_s(x) = x^s (F(x) - G(x)).$$

Для этого естественно потребовать, чтобы F и G имели конечные абсолютные моменты порядка s, то есть,

$$\int_{-\infty}^{\infty} |x|^s \, dF(x) < \infty, \qquad \int_{-\infty}^{\infty} |x|^s \, |dG(x)| < \infty,$$

где $|G|$ обозначает вариацию G как положительную меру на прямой. В этом случае A_s представляет собой функцию ограниченной вариации, причем $A_s(-\infty) = A_s(\infty) = 0$.

Зафиксируем функцию распределения H и положим $H_T(x) = H(Tx)$, $T > 0$, а также выберем $l > 0$ так, чтобы выполнялось условие $1 - H[-l, l] \leq \frac{1}{4}$.

Лемма 16.1. *Предположим, что функция распределения F и функция ограниченной вариации G имеют конечные абсолютные моменты целого порядка $s \geq 1$, причем $G(-\infty) = 0$, $G(\infty) = 1$. Если G дифференцируема, и ее производная удовлетворяет*

$$\sup_x \, (1 + |x|^s) \, |G'(x)| \leq L, \tag{16.2}$$

то для любого $T \geq 1$

$$\|A_0\| + \|A_s\| \leq c\Big(\|A_0 * H_T\| + \|A_s * H_T\| + \frac{L}{T}\Big), \tag{16.3}$$

где постоянная $c = c(s, l)$ зависит только от s и l.

Приведем доказательство, следуя стандартным, хотя и рутинным рассуждениям (см. [3], лемма 8, гл. VI). Сначала рассмотрим неполиномиальный случай $s = 0$ и выведем неравенство (2.2). Полагая $\sigma = 1/T$, рассмотрим свертку

$$I(x) \equiv (A * H_T)(x) = \int_{-\infty}^{\infty} A(x - \sigma y) \, dH(y) = I_0(x) + I_1(x),$$

где $A = A_0 = F - G$, с разбиением последнего интеграла на

$$I_0(x) = \int_{|y| \leq l} A(x - \sigma y) \, dH(y), \quad I_1(x) = \int_{|y| > l} A(x - \sigma y) \, dH(y).$$

Имеем $|I_1(x)| \leq \gamma \, \|A\|$ с коэффициентом $\gamma = 1 - H[-l, l]$, в то время как для оценивания первого интеграла, используя монотонность F и липшицевость G, можно применить неравенства

$$A(x - \sigma y) \geq A(x - \sigma l) - 2\sigma l L, \qquad -A(x - \sigma y) \geq -A(x + \sigma l) - 2\sigma l L,$$

которые влекут

$$|I_0(x)| \geq (1-\gamma)\,\max\{A(x-\sigma l), -A(x+\sigma l)\} - 2(1-\gamma)\,\sigma l L.$$

Поскольку $|I(x)| \geq |I_0(x)| - |I_1(x)|$, получаем

$$\|I\| \geq (1-\gamma)\,\max\{A(x-\sigma l), -A(x+\sigma l)\} - 2(1-\gamma)\,\sigma l L - \gamma\,\|A\|.$$

Беря супремум по всем x, приходим к оценке $\|I\| \geq (1-2\gamma)\,\|A\| - 2(1-\gamma)\,\sigma l L$, то есть,

$$\|A\| \,\leq\, \frac{1}{1-2\gamma}\,\|A * H_T\| + 2l\,\frac{1-\gamma}{1-2\gamma}\,\frac{L}{T},$$

что в точности совпадает с (2.2). В частности, если $\gamma \leq \frac{1}{4}$, эта оценка дает

$$\|A\| \,\leq\, 2\,\|A * H_T\| + 3l\,\frac{L}{T}. \tag{16.4}$$

Пусть теперь $s \geq 1$. Фиксируя произвольное значение $\varepsilon \in (0, \frac{1}{2}]$, выберем точку x_0, так чтобы

$$|A_s(x_0)| \,\geq\, (1-\varepsilon)\,\|A_s\| \,\geq\, \frac{1}{2}\,\|A_s\|. \tag{16.5}$$

Не умаляя общности, можно сделать два предположения:

$$|x_0| \geq 4sl, \qquad |A_s(x_0)| \,\geq\, 2^{s+3}\,\sigma l L. \tag{16.6}$$

Действительно, если первое условие нарушается, то в силу (16.4)-(16.5),

$$\begin{aligned}\|A_s\| \,\leq\, 2\,|A_s(x_0)| \;&\leq\; 2\,(4sl)^s\,|A(x_0)| \\ &\leq\; 2\,(4sl)^s\,\|A\| \,\leq\, 4\,(4sl)^s\,\|A * H_T\| + 6l\,(4sl)^s\,\frac{L}{T},\end{aligned}$$

что автоматически влечет (16.3), а именно

$$\|A\| + \|A_s\| \,\leq\, (2 + 4\,(4sl)^s)\,\|A * H_T\| + (3l + 6l\,(4sl)^s)\,\frac{L}{T}. \tag{16.7}$$

Если же нарушается второе условие, то $\|A_s\| \leq 2\,|A_s(x_0)| \leq 2^{s+4}\,l\,\frac{L}{T}$, что дает аналогичную оценку

$$\|A\| + \|A_s\| \,\leq\, 2\,\|A * H_T\| + (3 + 2^{s+4})\,l\,\frac{L}{T}. \tag{16.8}$$

Как и в начале доказательства, рассмотрим аналогичную свертку

$$I(x) \equiv (A_s * H_T)(x) \,=\, \int_{-\infty}^{\infty} A_s(x-\sigma y)\,dH(y) = I_0(x) + I_1(x)$$

с разбиением последнего интеграла на

$$I_0(x) = \int_{|y|\leq l} A_s(x-\sigma y)\,dH(y), \qquad I_1(x) = \int_{|y|>l} A_s(x-\sigma y)\,dH(y).$$

Снова $|I_1(x)| \leq \gamma\,\|A_s\|$, $\gamma = 1 - H[-l,l]$, так что для всех x

$$\|I\| \,\geq\, |I(x)| \,\geq\, |I_0(x)| - \gamma\,\|A_s\|. \tag{16.9}$$

Далее будем оценивать $|I_0(x)|$ снизу либо в точке $x = x_0 - \sigma l$, либо в точке $x = x_0 + \sigma l$ в зависимости от знака $A(x_0)$. Пусть для определенности $A(x_0) < 0$.

Для всех $z \in [x_0 - 2\sigma l, x_0]$ найдется точка z_0 в том же промежутке, такая что $G(z) = G(x_0) + (z - x_0)\, G'(z_0)$. При этом $|z_0| \geq |x_0|$, если $x_0 < 0$, и $z_0 \geq \frac{1}{2}\, x_0$, если $x_0 > 0$. Действительно, в последнем случае наихудший вариант достигается для $z_0 = x_0 - 2\sigma l$. Но $x_0 \geq 4\sigma l$ и $\sigma \leq 1$, что влечет $z_0 \geq \frac{1}{2}\, x_0$. Таким образом, с учетом условия (16.2)

$$|G(z) - G(x_0)| = |z - x_0|\, |G'(z_0)| \leq 2\sigma l\, \frac{L}{|z_0|^s} \leq 2^{s+1} \sigma l\, \frac{L}{|x_0|^s}.$$

Используя монотонность функции F, имеем

$$\begin{aligned} A(z) &= F(z) - G(z) \\ &\leq F(x_0) - G(x_0) + 2^{s+1}\sigma l\, \frac{L}{|x_0|^s} \ = \ A(x_0) + 2^{s+1}\sigma l\, \frac{L}{|x_0|^s} \ < \ \frac{3}{4}\, A(x_0), \end{aligned}$$

где на последнем шаге мы воспользовались вторым условием в (16.6). Кроме того, $|z|^s \geq |x_0|^s$ в случае $x_0 < 0$, а в случае $x_0 > 0$, используя первое условие в (16.6) вместе с предположением $\sigma \leq 1$, имеем

$$z^s \geq (x_0 - 2\sigma l)^s \geq (x_0 - 2l)^s \geq \Big(1 - \frac{1}{2s}\Big)^s\, x_0^s \geq \frac{1}{\sqrt{e}}\, x_0^s.$$

В обоих случаях $|z|^s \geq \frac{1}{\sqrt{e}}\, |x_0|^s$, и следовательно функция $A_s(z) = z^s A(z)$ не меняет знак в интервале $x_0 - 2\sigma l \leq z \leq x_0$ и удовлетворяет на нем неравенству

$$|A_s(z)| \geq \frac{3}{4\sqrt{e}}\, |A_s(x_0)|.$$

Отсюда

$$|I_0(x)| = \int_{|y| \leq l} |A_s(x - \sigma y)|\, dH(y) \geq (1 - \gamma)\, \frac{3}{4\sqrt{e}}\, |A_s(x_0)| \geq \frac{1 - \gamma}{1 - \varepsilon}\, \frac{3}{4\sqrt{e}}\, \|A_s\|,$$

где на последнем шаге мы воспользовались (16.5).

С аналогичными рассуждениями эти неравенства сохраняются и в случае $A(x_0) > 0$ для точки $x = x_0 + \sigma l$ и интервала $x_0 \leq z \leq x_0 + 2\sigma l$. Таким образом, устремляя $\varepsilon \to 0$, в обоих случаях приходим к одной и той же оценке

$$|I_0(x)| \ \geq \ \frac{3}{4\sqrt{e}}\, (1 - \gamma)\, \|A_s\|.$$

Применяя ее в (16.9), получаем

$$\|I\| \ \geq \ \Big(\frac{3}{4\sqrt{e}}\, (1 - \gamma) - \gamma\Big)\, \|A_s\| \ \geq \ 0.091\, \|A_s\|,$$

где в последнем неравенстве предполагается, что $\gamma \leq \frac{1}{4}$. Следовательно, в этом случае $\|A_s\| \leq 11\, \|I\| = 11\, \|A_s * H_T\|$, и с учетом (16.4) имеем

$$\|A\| + \|A_s\| \ \leq \ 2\, \|A * H_T\| + 11\, \|A_s * H_T\| + 3l\, \frac{L}{T}.$$

Принимая во внимание (16.7)-(16.8), приходим к (16.3). Лемма доказана.

17 Общие неравномерные оценки

Правая часть в неравенстве (16.3) из леммы 16.1 может быть далее оценена в терминах преобразований Фурье-Стилтьеса F и G. На этом шаге можно рассматривать более широкие классы функций.

Пусть A – функция ограниченной вариации с конечным абсолютным моментом целого порядка $s \geq 1$ (для вариации $|A|$, как меры на прямой), причем $A(-\infty) = A(\infty) = 0$. В этом случае соответствующее ей преобразование Фурье-Стилтьеса

$$a(t) = \int_{-\infty}^{\infty} e^{itx}\, dA(x) = -it \int_{-\infty}^{\infty} e^{itx}\, A(x)\, dx$$

имеет s непрерывных производных, а функция

$$A_s(x) = x^s A(x)$$

тоже имеет ограниченную вариацию и удовлетворяет $A_s(-\infty) = A_s(\infty) = 0$.

Теорема 17.1. *Для любой функции распределения H с характеристической функцией h имеем*

$$\|A_s * H\| \ \leq \ \frac{2}{\pi} \sup_{x \in \mathbf{R},\ \kappa \geq 1} \bigg| \int_{-\infty}^{\infty} e^{-itx}\, \frac{a^{(s)}(t)}{t}\, h(\kappa t)\, dt \bigg|. \tag{17.1}$$

В частности, без какого-либо сглаживания (или когда H – единичная нагрузка в нуле с $h(t) = 1$)

$$\|A_s\| \ \leq \ \frac{2}{\pi} \sup_{x} \bigg| \int_{-\infty}^{\infty} e^{-itx}\, \frac{a^{(s)}(t)}{t}\, dt \bigg|. \tag{17.2}$$

Однако, как и в обычном неравенстве Берри-Эссеена, сглаживание сводит задачу оценивания $\|A_s\|$ к оцениванию $a^{(s)}(t)$ на интервале $|t| \leq T$ с ошибкой порядка $1/T$.

Интегралы в (17.1)-(17.2) всегда конечны, но понимаются в смысле главного значения как пределы интегралов по множеству $\varepsilon < |t| < T$ с $\varepsilon \downarrow 0$ и $T \uparrow \infty$. Здесь можно использовать следующий вариант обращения для преобразования Фурье-Стилтьеса: если B – функция ограниченной вариации с преобразованием Фурье-Стилтьеса b, то для всех x

$$\frac{1}{\pi} \lim_{\varepsilon \downarrow 0,\ T \uparrow \infty} \int_{\varepsilon < |t| < T} \frac{e^{-itx}}{-it}\, b(t)\, dt \ = \ \int_{-\infty}^{\infty} \Big(1_{\{y<x\}} - 1_{\{y>x\}}\Big)\, dB(y). \tag{17.3}$$

При этом, если B непрерывна справа и $B(-\infty) = 0$ (что всегда и предполагается), то интегралы в правой части (17.3) связаны с L^∞-нормой функции B соотношениями

$$\|B\| \leq \sup_x \bigg| \int_{-\infty}^{\infty} \Big(1_{\{y<x\}} - 1_{\{y>x\}}\Big)\, dB(y) \bigg| \ \leq \ 3\, \|B\|. \tag{17.4}$$

Действительно, обозначим эти интегралы через $I(x)$ и положим $M = \sup_x |I(x)|$. Устремляя $x \to \infty$, приходим к $M \geq |B(\infty)|$. А так как $2B(x) = B(\infty) + I(x)$ в каждой точке x, в которой B непрерывна, моментально получаем $|B(x)| \leq M$, то есть, левое неравенство в (17.4). Кроме того, $|I(x)| \leq 2\,|B(x)| + |B(\infty)| \leq 3\,\|B\|$, что дает правое неравенство.

Для доказательства теоремы нам понадобятся две элементарные формулы.

Лемма 17.2. *При $s \geq 1$ функция A_s имеет преобразование Фурье-Стилтьеса*

$$a_s(t) = i^{-s}\, t \left(\frac{a(t)}{t}\right)^{(s)} = i^{-s} \int_0^1 \left(a^{(s)}(t) - a^{(s)}(\eta t)\right) d\eta^s, \qquad t \neq 0. \tag{17.5}$$

Первое соотношение (которое верно и при $s = 0$) получается, если продифференцировать s раз по переменной t равенство $\frac{a(t)}{it} = -\int_{-\infty}^{\infty} e^{itx} A(x)\, dx$ и затем проинтегрировать по частям:

$$\Big(\frac{a(t)}{it}\Big)^{(s)} \;=\; -i^s \int_{-\infty}^{\infty} e^{itx} A_s(x)\, dx \;=\; \frac{i^s}{it} \int_{-\infty}^{\infty} e^{itx} dA_s(x) \;=\; \frac{i^s}{it}\, a_s(t).$$

Однако это рассуждение не совсем стогое, так как A_s не обязательно интегрируема по Лебегу. Но интегрируема функция A_{s-1}, так что

$$\Big(\frac{a(t)}{it}\Big)^{(s-1)} \;=\; -i^{s-1} \int_{-\infty}^{\infty} e^{itx} A_{s-1}(x)\, dx, \qquad t \neq 0. \tag{17.6}$$

Так как $dA_s(x) = x^s dA(x) + sA_{s-1}(x)\, dx$, имеем еще одно тождество

$$\int_{-\infty}^{\infty} e^{itx}\, dA_s(x) = \int_{-\infty}^{\infty} e^{itx}\, x^s\, dA(x) + s \int_{-\infty}^{\infty} e^{itx}\, A_{s-1}(x)\, dx,$$

то есть, с учетом (17.6)

$$a_s(t) = i^{-s} \left[a^{(s)}(t) - s\Big(\frac{a(t)}{t}\Big)^{(s-1)}\right].$$

Поэтому остается установить равенство $a^{(s)}(t) - s\,(\frac{a(t)}{t})^{(s-1)} = t\,(\frac{a(t)}{t})^{(s)}$ или – что то же самое в терминах $v(t) = \frac{a(t)}{t}$ – равенство

$$(tv(t))^{(s)} = t\, v^{(s)}(t) + s\, v^{(s-1)}(t),$$

которое уже очевидно.

Второе соотношение в (17.5) вытекает из тождества

$$\left(\frac{a(t)}{t}\right)^{(s)} = \frac{1}{t} \int_0^1 \left(a^{(s)}(t) - a^{(s)}(\eta t)\right) d\eta^s, \qquad t \neq 0, \tag{17.7}$$

справедливого для любой s раз непрерывно дифференцируемой функции a, такой что $a(0) = 0$. Не умаляя общности, здесь можно предположить, что a имеет $s+1$ непрерывную производную. Тогда (17.7) получается, если продифференцировать s раз равенство $\frac{a(t)}{t} = \int_0^1 a'(\eta t)\, d\eta$ и затем проинтегрировать по частям. Лемма 17.2 доказана.

Перейдем теперь к доказательству теоремы 17.1. Применяя (17.3)–(17.5) к функции $A_s * H$ с преобразованием Фурье-Стилтьеса $a_s(t)h(t)$, получаем

$$\|A_s * H\| \;\leq\; \frac{1}{\pi} \sup_x \left| \int_{-\infty}^{\infty} e^{-itx}\, h(t) \left[\int_0^1 \frac{a^{(s)}(t) - a^{(s)}(\eta t)}{t}\, d\eta^s \right] dt \right|. \tag{17.8}$$

Проверим, что в правой части можно поменять порядок интегрирования, что следовало бы из того, что

$$\delta_0(\varepsilon,\varepsilon') \equiv \int_{\varepsilon'<|t|<\varepsilon} e^{-itx}\, h(t) \left[\int_0^1 \frac{a^{(s)}(\eta t)}{t}\, d\eta^s \right] dt \to 0 \qquad (0<\varepsilon'<\varepsilon,\ \varepsilon\to 0)$$

и

$$\delta_1(T,T') \equiv \int_{T<|t|<T'} e^{-itx}\, h(t) \left[\int_0^1 \frac{a^{(s)}(\eta t)}{t}\, d\eta^s \right] dt \to 0 \qquad (T<T',\ T\to\infty).$$

Для каждого $\eta\in(0,1)$ с точностью до множителя i^s функция $t \to a^{(s)}(\eta t)h(t)$ является преобразованием Фурье-Стилтьеса для свертки $V_\eta(x) = B(x/\eta) * H(x)$, где $B(x) = \int_{-\infty}^x y^s\, dA(y)$. Поэтому, вводя функцию $\psi(t) = \int_0^t \frac{\sin u}{u}\, du$ и полагая $V = V_1$, имеем

$$\begin{aligned}
\delta_1(T,T') &= \int_0^1 \left[\int_{T<|t|<T'} e^{-itx}\, h(t)\, \frac{a^{(s)}(\eta t)}{t}\, dt \right] d\eta^s \\
&= \int_0^1 \left[\int_{-\infty}^{\infty} \int_{T<|t|<T'} \frac{e^{it(y-x)}}{t}\, dt\, dV_\eta(y) \right] d\eta^s \\
&= 2i \int_0^1 \int_{-\infty}^{\infty} \Big(\psi(T'(y-x)) - \psi(T(y-x)) \Big)\, dV_\eta(y)\, d\eta^s \\
&= 2i \int_0^1 \int_{-\infty}^{\infty} u(y,\eta)\, dV(y)\, d\eta^s,
\end{aligned}$$

где

$$u(y,\eta) = \psi(T'(\eta y - x)) - \psi(T(\eta y - x)).$$

Так как $|\psi| \le C$, причем $\psi(t) \to \frac{\pi}{2}$ при $t\to\infty$ и $\psi(t) \to -\frac{\pi}{2}$ при $t \to -\infty$, модуль функции u ограничен абсолютной постоянной $2C$ равномерно по всем $T' > T > 0$, и при этом $u(y,\eta) \to 0$ при $T\to\infty$ для любых y и η. Следовательно, в силу теоремы Лебега о мажорируемой сходимости, $\delta_1(T,T') \to 0$. С аналогичными рассуждениями также доказывается, что $\delta_0(\varepsilon,\varepsilon') \to 0$ при $\varepsilon\to 0$.

Наконец, меняя порядок интегрирования, можно заключить, что выражение под знаком супремума в правой части (17.8) не превосходит

$$\left| \int_{-\infty}^{\infty} e^{-itx}\, h(t)\, \frac{a^{(s)}(t)}{t}\, dt \right| + \sup_{0<\eta<1} \left| \int_{-\infty}^{\infty} e^{-itx}\, h(t)\, \frac{a^{(s)}(\eta t)}{t}\ dt \right|,$$

что и приводит к требуемому неравенству (17.1). Теорема 17.1 доказана.

Замечание 17.3. При $s=0$ теорема 17.1 остается в силе, причем без условия $A(\infty)=0$, и при этом неравенство (17.1) можно уточнить. Действительно, применяя (17.3) и затем левое неравенство в (17.4) с $B = A*H$ и $b = ah$, получаем

$$\|A*H\| \ \le\ \frac{1}{\pi} \sup_x \left| \int_{-\infty}^{\infty} e^{-itx}\, \frac{a(t)}{t}\, h(t)\, dt \right|. \tag{17.9}$$

Если $A(\infty)=0$, то множитель $\frac{1}{\pi}$ можно заменить на $\frac{1}{2\pi}$. В этом случае $B(\infty)=0$, и поэтому (17.3) влечет тождество

$$\frac{1}{\pi} \sup_x \left| \int_{-\infty}^{\infty} \frac{e^{-itx}}{t}\, b(t)\, dt \right| \ =\ 2\,\|B\|.$$

Сделаем еще одно замечание, возвращаясь к “несглаженному"варианту (17.2). Так как с точностью до множителя i^s производная $a^{(s)}$ является преобразованием Фурье-Стилтьеса для функции $B(x) = \int_{-\infty}^{x} y^s\,dA(y)$, модуль интеграла в правой части (17.2) совпадает, согласно тождеству (17.3), с

$$\pi \left| \int_{-\infty}^{\infty} \Big(1_{\{y<x\}} - 1_{\{y>x\}} \Big)\, y^s\,dA(y) \right|.$$

Поэтому, принимая во внимание правое неравенство в (17.4), из (17.2) получаем оценку

$$\sup_x |x^s A(x)| \,\leq\, 6 \sup_x \left| \int_{-\infty}^{x} y^s\,dA(y) \right|,$$

в которой преобразование Фурье-Стилтьеса уже не фигурирует. В виду (17.4), эта оценка эквивалентна (17.2) с точностью до абсолютного множителя.

18 Неравномерные оценки для функций распределения

Вернемся теперь к лемме 16.1 и применим теорему 17.1 и неравенство (17.9) с $A = F - G$ и со сглаживающей функцией распределения $H(Tx)$ (вместо H). Тогда мы приходим к неравномерной оценке, т.е., оценке для расстояния

$$\rho_s(F,G) \,=\, \sup_x\, (1 + |x|^s)\, |F(x) - G(x)|$$

в терминах преобразований Фурье-Стилтьеса

$$f(t) = \int_{-\infty}^{\infty} e^{itx}\,dF(x), \qquad g(t) = \int_{-\infty}^{\infty} e^{itx}\,dG(x) \qquad (t \in \mathbf{R}).$$

Будем предполагать, что функция распределения F и функция ограниченной вариации G имеют конечные абсолютные моменты целого порядка $s \geq 1$, причем $G(-\infty) = 0$, $G(\infty) = 1$.

Теорема 18.1. *Если G дифференцируема и удовлетворяет $\sup_x (1+|x|^s)|G'(x)| \leq L$, то для любой характеристической функции h и всех $T \geq 1$*

$$\begin{aligned} c_s\, \rho_s(F,G) \;\; &\leq \;\; \sup_{x \in \mathbf{R}} \left| \int_{-\infty}^{\infty} e^{-itx}\, \frac{f(t) - g(t)}{t}\, h(t/T)\,dt \right| \\ &\quad + \sup_{x \in \mathbf{R},\ \kappa \geq 1} \left| \int_{-\infty}^{\infty} e^{-itx}\, \frac{f^{(s)}(t) - g^{(s)}(t)}{t}\, h(\kappa t/T)\,dt \right| + \frac{L}{T} \end{aligned} \tag{18.1}$$

с некоторой постоянной $c_s > 0$, зависящей от s и h.

В частности, с каноническим ядром $h(t) = (1 - |t|)_+$ и учитывая, что $\kappa \geq 1$ под супремумом в (18.1), получаем

$$c_s\, \rho_s(F,G) \,\leq\, \int_{-T}^{T} \Big| \frac{f(t) - g(t)}{t} \Big|\,dt + \int_{-T}^{T} \Big| \frac{f^{(s)}(t) - g^{(s)}(t)}{t} \Big|\,dt + \frac{L}{T}. \tag{18.2}$$

Без использования второго тождества в лемме 17.2, мы бы пришли к похожей оценке

$$c_s\,\rho_s(F,G) \,\le\, \int_{-T}^{T} \Big|\frac{f(t)-g(t)}{t}\Big|\,dt + \int_{-T}^{T} \Big|\frac{d^s}{dt^s}\,\frac{f(t)-g(t)}{t}\Big|\,dt + \frac{L}{T}, \tag{18.3}$$

которую можно найти, например, в книге В. В. Петрова [3].

Неравенства (18.2)–(18.3) имеет смысл применять в ситуации, когда F и G имеют конечные абсолютные моменты порядка $s+1$, причем

$$\int_{-\infty}^{\infty} x^p\,dF(x) = \int_{-\infty}^{\infty} x^p\,dG(x), \qquad p = 1,\dots,s.$$

В этом случае $|f^{(s)}(t) - g^{(s)}(t)| = O(|t|)$ при $t \to 0$, так что интегралы в правой части этих неравенств конечны. Если же имеется равенство моментов F и G только порядков $\le s-1$, то следует прибегнуть к дополнительным аргументам, которые, впрочем, тоже могут быть основаны на сглаживающем неравенстве (18.1).

Допустим, например, что для производных f и g порядка s имеется разложение

$$f^{(s)}(t) - g^{(s)}(t) = u(t) + v(t), \tag{18.4}$$

в котором v представляет собой преобразование Фурье-Стилтьеса функции ограниченной вариации V, такой что $V(-\infty) = 0$, но необязательно $V(\infty) = 0$. Если известно, что норма $\|V\|$ мала (желательно порядка $1/T$), и при этом $u(t) = tw(t)\,O(1/T)$ на интервале $[-T,T]$ для некоторой суммируемой функции w, то может быть полезным следующее приложение теоремы 18.1.

Следствие 18.2. *В условиях теоремы* 18.1 *и с учетом разложения* (18.4) *имеем*

$$c_s\,\rho_s(F,G) \,\le\, \int_{-T}^{T} \Big|\frac{f(t)-g(t)}{t}\Big|\,dt + \int_{-T}^{T} \Big|\frac{u(t)}{t}\Big|\,dt + \|V\| + \frac{L}{T} \tag{18.5}$$

для всех $T \ge 1$ с постоянной $c_s > 0$, зависящей только от s.

Для вывода этого неравенства на основе (18.1) нужно лишь оценить интегралы

$$I = \int_{-\infty}^{\infty} e^{-itx}\,\frac{v(t)}{t}\,h(\kappa t/T)\,dt$$

равномерно по всем x и κ с $h(t) = (1-|t|)_+$. Обозначим через H соответствующую функцию распределения и положим $H_q(x) = H(qx)$. Тогда, в силу тождества (17.3), и применяя верхнюю оценку в (17.4), имеем

$$|I| \,=\, \pi\,\bigg|\int_{-\infty}^{\infty} \Big(1_{\{y<x\}} - 1_{\{y>x\}}\Big)\,d\,(V * H_{T/\kappa})(y)\bigg| \,\le\, 3\pi\,\|V * H_{T/\kappa}\| \,\le\, 3\pi\,\|V\|$$

(поскольку сворачивание с любой функцией распределения не увеличивает L^∞-норму заданной функции ограниченной вариации). В итоге получаем (18.5).

Для иллюстрации разложения (18.4), предположим, что F представляет собой свертку вероятностных распределений F_k ($1 \le k \le n$) с характеристическими функциями f_k, у которых нулевые средние, дисперсии σ_k^2 удовлетворяют $\sum_{k=1}^n \sigma_k^2 = 1$, причем конечны моменты $\beta_{s,k}$ порядка $s \geq 3$. Таким образом, F имеет характеристическую функцию $f(t) = f_1(t) \cdots f_n(t)$. В качестве G можно взять стандартную нормальную функцию распределения Φ или, что лучше, – "поправленную" нормальную функцию из разложения Эджворта порядка s (которая уже не обязана быть монотонной). Тогда можно положить

$$v(t) = \sum_{k=1}^n f_1(t) \cdots f_{k-1}(t)\, f_k^{(s)}(t)\, f_{k+1}(t) \cdots f_n(t).$$

Очевидно, что соответствующая ей функция ограниченной вариации V имеет норму

$$\|V\| \,\leq\, \sum_{k=1}^n \sup_x \bigg| \int_{-\infty}^x y^s \, dF_k(y) \bigg| \,\leq\, \sum_{k=1}^n \beta_{s,k} \,\equiv\, L_s,$$

т.е., эта норма оценивается дробью Ляпунова L_s порядка s. Можно показать, что при $s=3$ вторая компонента $u(t) = f^{(3)}(t) - g^{(3)}(t) - v(t)$ тоже оценивается такой же дробью, причем с множителем $O(|t|\, e^{-t^2/6})$ на интервале $|t| \leq T$ с $T = 1/L_3$. Следовательно, в этом случае применение (18.5) приводит к хорошо известной оценке вида

$$\rho_3(F, \Phi) \leq c L_3 \tag{18.6}$$

с абсолютной постоянной c. (При дополнительных предположениях о поведении $f_k(t)$ при больших t справедливы и аналогичные оценки $\rho_s(F,G) \leq c_s L_s$ для $s \geq 4$).

На возможность такого подхода к неравномерным оценкам в центральной предельной теореме, связанного с вычленением в разложении (18.4) "плохой" компоненты $v(t)$, недавно обратил внимание И. Пинелис [43] (см. также [44]). Им было предложено новое доказательство (18.6) для сверточных степеней (т.е., распределений сумм независимых одинаково распределенных случайных величин) на основе неравенств сглаживания типа Бомана-Правица-Ваалера

$$\begin{aligned} F(x) \;&\leq\; \frac{1}{2} + \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-itx}}{-it}\, f(t) h(t/T)\, dt \qquad (x \in \mathbf{R},\ T > 0), \\ F(x-) \;&\geq\; \frac{1}{2} + \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-itx}}{-it}\, f(t) h(-t/T)\, dt, \end{aligned} \tag{18.7}$$

справедливых в классе всех функций распределения F с характеристическими функциями f. В таких неравенствах функции h (т.н. сглаживающие ядра), представляя собой преобразование Фурье-Стилтьеса знакопеременных мер, должны иметь специальную структуру. В приложениях они выбираются с носителем в интервале $[-1,1]$, так что интегралы в (18.7) берутся по интервалу $[-T,T]$; например, ядро Правица задается формулой

$$h(t) = (1 - |t|)\, \pi t \cot(\pi t) + |t| - i(1 - |t|)\, \pi t, \qquad |t| < 1$$

(см. [45]). Следуя Боману [46], Пинелис описал широкий класс функций h, удовлетворяющих (18.7). Если F имеет конечный абсолютный момент порядка s, и функция h имеет s непрерывных производных, то, как показано в [43], для всех $x \geq 0$

$$\begin{aligned} x^s(1 - F(x)) \;&\leq\; \frac{(-i)^s}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \frac{d^s}{dt^s}\, \frac{f(t)h(-t/T) - 1}{it}\, dt, \\ x^s(1 - F(x-)) \;&\geq\; \frac{(-i)^s}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \frac{d^s}{dt^s}\, \frac{f(t)h(t/T) - 1}{it}\, dt, \end{aligned}$$

что можно рассматривать как аналог теоремы 18.1. Другие интересные приложения такого рода неравенств, как (18.7), и связанные с ними экстремальные задачи в анализе Фурье обсуждались Ваалером [47].

Отметим, что использование неравенств типа (18.7) позволяет обойти задачу сравнения расстояний между исходными и сглаженными вероятностными распределениями, и при этом может привести к уточнению ряда результатов (например, о выборе абсолютных постоянных). С другой стороны, применение (18.7) ограничено классом монотонных функций G, выступающих как аппроксимация для F, в отличие от более общих неравенств, таких как (18.1) и (18.5).

19 Нижние оценки для расстояния Колмогорова

В завершение перейдем к противоположной задаче – оцениванию расстояния Колмогорова между функциями распределения снизу в терминах соответствующих характеристических функций. Так как в некоторых задачах интересуются приближениями заданной функции распределения не обязательно монотонными функциями, имеет смысл рассматривать более широкий класс функций A ограниченной вариации и оценивать снизу L^∞-норму $\|A\|$ в терминах преобразований Фурье-Стилтьеса

$$a(t) = \int_{-\infty}^{\infty} e^{itx}\, dA(x).$$

Будем предполагать, что $A(-\infty) = A(\infty) = 0$.

Приведем сначала одну популярную оценку (см. [13], [48], [3]).

Теорема 19.1. *Имеет место неравенство*

$$\|A\| \,\geq\, \frac{1}{2\sqrt{2\pi}}\,\Big|\int_{-\infty}^{\infty} a(t)\, e^{-t^2/2} dt\Big|. \tag{19.1}$$

Если поведение функции $a(t)$ известно только вблизи нуля, более предпочтительным может оказаться следующее утверждение (которое, как нам кажется, в литературе не упоминается).

Теорема 19.2. *Для любого* $T > 0$

$$\|A\| \,\geq\, \frac{1}{3T}\,\Big|\int_0^T a(t)\Big(1 - \frac{t}{T}\Big)\, dt\Big|. \tag{19.2}$$

Стандартный подход к такого рода оценкам основан на применении теоремы Планшереля, то есть, на тождестве

$$\int_{-\infty}^{\infty} v(t)w(t)\, dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathcal{P}v(x)\bar{\mathcal{P}}\bar{w}(x)\, dx, \qquad v, w \in L^2(\mathbf{R}), \tag{19.3}$$

где $\mathcal{P}v = \widehat{v}$ обозначает преобразование Фурье функции v. Если функция $a(t)/t$ интегрируема, то в силу формулы обращения,

$$A(x) = -\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\,\frac{a(t)}{it}\, dt.$$

То есть, функция $A(-x)$ представляет собой преобразование Фурье функции $v(t) = -\frac{1}{2\pi i}\,\frac{u(t)}{t}$. Если также $v \in L^2$, то (19.3) принимает вид

$$\int_{-\infty}^{\infty} \frac{a(t)}{t}\, w(t)\, dt = -i \int_{-\infty}^{\infty} A(-x) \bar{\mathcal{P}} \bar{w}(x)\, dx = -i \int_{-\infty}^{\infty} A(x) \mathcal{P} w(x)\, dx.$$

Отсюда сразу получаем общую оценку

$$\Big| \int_{-\infty}^{\infty} \frac{a(t)}{t}\, w(t)\, dt \Big| \le \|A\| \int_{-\infty}^{\infty} |\mathcal{P} w(x)|\, dx. \tag{19.4}$$

Например, частный случай $w(t) = te^{-t^2/2}$ приводит нас к (19.1). При этом дополнительные предположения на $a(t)$ в этом неравенстве легко снимаются, так же как и в неравенстве (19.2). Чтобы вывести (19.2), применим (19.4) к $w(t) = \frac{t}{T}\,(1 - \frac{t}{T})_+\, 1_{(0,\infty)}(t)$. Пусть сначала $T = 1$. Тогда для всех $x \neq 0$

$$\mathcal{P} w(x) \ = \ \int_0^1 e^{itx}\, t(1-t)\, dt \ = \ \frac{-e^{ix} - 1}{x^2} + \frac{2(e^{ix} - 1)}{ix^3} \ = \ \frac{q(x)}{x^3},$$

где $q(x) = -xe^{ix} - 2i\, e^{ix} - x + 2i$. Отсюда $|q(x)| \le 2|x| + 4$ и $|\mathcal{P} w(x)| \le \frac{2|x|+4}{|x|^3}$. С другой стороны,

$$|\mathcal{P} w(x)| \le \int_0^1 t(1-t)\, dt = \frac{1}{6},$$

и, следовательно,

$$\begin{aligned} \int_{-\infty}^{\infty} |\mathcal{P} w(x)|\, dx \ &= \ 2 \int_0^4 |\mathcal{P} w(x)|\, dx + 2 \int_4^{\infty} |\mathcal{P} w(x)|\, dx \\ &\le \ \frac{4}{3} + 2 \int_4^{\infty} \frac{2x+4}{x^3}\, dx \ = \ \frac{17}{6} \ < \ 3. \end{aligned}$$

В общем случае для $w_T(x) = w(t/T)$, имеем $\mathcal{P} w_T(x) = T \cdot (\mathcal{P} w)(Tx)$, так что

$$\int_{-\infty}^{\infty} |\mathcal{P} w_T(x)|\, dx = \int_{-\infty}^{\infty} |\mathcal{P} w(x)|\, dx < 3.$$

Таким образом, использование $w = w_T$ в (19.4) доказывает (19.2).

Отметим еще, что применяя (19.4) к другой функции $w(t) = \frac{t}{T}\,(1 - \frac{|t|}{T})_+$, мы бы получили аналогичное неравенство

$$\|A\| \ \ge \ \frac{1}{3.5\, T} \left| \int_{-T}^{T} a(t) \Big(1 - \frac{|t|}{T}\Big) dt \right|.$$

20 Оценки в центральной предельной теореме

Пусть $F_n(x) = \mathbf{P}\{Z_n \le x\}$ – функция распределения нормализованной суммы

$$Z_n = \frac{1}{\sqrt{n}}\,(X_1 + \ldots + X_n)$$

независимых одинаково распределенных случайных величин $X_1, \dots, X_n$, таких что $\mathbf{E}X_1 = 0$, $\mathbf{E}X_1^2 = 1$. Согласно центральной предельной теореме, $F_n(x)$ сходятся при возрастании n к стандартной нормальной функции распределения

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2}\, dy.$$

Насколько близко F_n к Φ в заданной метрике d зависит от исходного распределения выборки, то есть, от распределения X_1. Тем не менее, для классических метрик, отвечающих за слабую сходимость, расстояние $d(F_n, \Phi)$ стремится к нулю со скоростью по крайней мере $\frac{c}{\sqrt{n}}$ при достаточно широких предположениях.

Чтобы установить результаты такого типа, достаточно сравнить характеристические функции

$$f_n(t) = \int_{-\infty}^{\infty} e^{itx}\, dF_n(x) = f_1\Big(\frac{t}{\sqrt{n}}\Big)^n, \qquad g(t) = \int_{-\infty}^{\infty} e^{itx}\, d\Phi(x) = e^{-t^2/2}$$

и воспользоваться соответствующими неравенствами сглаживания. Если 3-й абсолютный момент $\beta_3 = \mathbf{E}\,|X_1|^3$ конечен, то нетрудно разложить $f_1(t)$ по формуле Тейлора в окрестности нуля (до кубического члена) и, как следствие, получить оценку вида

$$|f_n(t) - g(t)| \,\le\, c\,\frac{\beta_3}{\sqrt{n}}\,|t|^3 e^{-t^2/4}, \qquad |t| \le \frac{\sqrt{n}}{\beta_3} \tag{20.1}$$

с некоторой абсолютной постоянной $c > 0$. Аналогичное соотношение справедливо и для первых 3 производных; более точно, в том же интервале

$$|f_n^{(s)}(t) - g^{(s)}(t)| \,\le\, c\,\frac{\beta_3}{\sqrt{n}}\,|t|^{3-s} e^{-t^2/4}, \qquad s = 0, 1, 2, 3. \tag{20.2}$$

1. *Расстояния Колмогорова, Леви и в метрике L^p.*

Применение (20.1) в теореме 2.1 с $T = \frac{\sqrt{n}}{\beta_3}$ приводит к классической теореме Берри-Эссеена для расстояния Колмогорова.

Теорема 20.1. *С некоторой абсолютной постоянной $c > 0$*

$$\rho(F_n, \Phi) \,\le\, c\,\frac{\beta_3}{\sqrt{n}}. \tag{20.3}$$

Значение наилучшей постоянной c в этом неравенстве неизвестно, но известно, что $0.4097 < c < 0.4690$ (см. [49]). Такое же утверждение справедливо и для расстояния Леви, поскольку в общем случае $L(F, \Phi) \le \rho(F, \Phi) \le (1 + \frac{1}{\sqrt{2\pi}})\, L(F, \Phi)$.

Если же 3-й абсолютный момент β_3 бесконечен, то скорость сходимости к нормальному закону может быть сколь угодно медленной. Как показано В. К. Мацкявичюсом [48] с использованием нижней оценки в теореме 19.1, для любой числовой последовательности $\varepsilon_n \to 0$ $(n \to \infty)$ можно подобрать общее распределение для случайных величин $X_1, X_2, \dots$ с нулевым средним и единичной дисперсией, так что будет выполняться неравенство

$$\rho(F_n, \Phi) \,\ge\, \varepsilon_n$$

при всех достаточно больших n.

С другой стороны, неравенство (20.3) может быть усилено в терминах расстояния Колмогорова с полиномиальным весом. Применяя (20.2) с $s = 2$ и тем же параметром T в неравенстве типа Берри-Эссеена (18.2), приходим к оценке Л. Д. Мешалкина и Б. А. Рогозина [50]

$$|F_n(x) - \Phi(x)| \, \leq \, c\, \frac{\beta_3}{(1+x^2)\sqrt{n}}. \tag{20.4}$$

В действительности имеет место еще более сильное утверждение, доказанное С. В. Нагаевым с использованием экспоненциальных оценок и дополнительных конструкций, таких как усечение (см. [51-52], [2]).

Теорема 20.2. *С некоторой абсолютной постоянной* $c > 0$ *для всех* $x \in \mathbf{R}$

$$|F_n(x) - \Phi(x)| \, \leq \, c\, \frac{\beta_3}{(1+|x|^3)\sqrt{n}}. \tag{20.5}$$

Подход Нагаева получил дальнейшее развитие во многих исследованиях, включая работы [53-57], где также изучались пути уточнения постоянной c в (20.5). Как уже отмечалось выше при обсуждении альтернативного подхода Пинелиса [43-44], это неравенство можно получить и с помощью следствия 18.2 с функцией

$$v(t) = \frac{1}{\sqrt{n}}\, f_1(t/\sqrt{n})^{n-1}\, f_1'''(t/\sqrt{n}),$$

выступающей в качестве особой ("плохой") компоненты в разложении (18.4) для 3-й производной $f_n'''(t)$.

Отметим еще, что из (20.4)-(20.5) сразу получаются верхние оценки и для расстояния в метрике L^p:

$$\|F_n - \Phi\|_p \, \leq \, c\, \frac{\beta_3}{\sqrt{n}}, \qquad p \geq 1.$$

В частности, при $p = 1$ приходим к неравенству Эссеена для среднего расстояния, которое может выведено и на основе теоремы 8.1 (или следствия 8.2).

2. *Расстояние Леви-Прохорова.*

Неравномерная оценка (20.5) недостаточна, чтобы на ее основе изучать скорость сходимости в других метриках, поэтому нужны другие неравенства сглаживания. Применяя (20.2) с $s = 1$ в теореме 6.1 с параметром $T = \frac{\sqrt{n}}{\beta_3}$, приходим к теореме В. В. Юринского [16], усиливающей теорему 20.1.

Теорема 20.3. *С некоторой абсолютной постоянной* $c > 0$

$$\pi(F_n, \Phi) \, \leq \, c\, \frac{\beta_3}{\sqrt{n}}.$$

Это утверждение сохраняется и для сумм одинаково распределенных случайных векторов в $\mathbf{R}^k$ с постоянной c, зависящей только от размерности k.

3. *Расстояния Золотарева.*

Для удобства будем писать $\zeta_s(X,Y)$ вместо $\zeta_s(F,G)$, когда случайные величины X и Y имеют распределения F и G сооответственно ($s = 1, 2, \dots$) Отметим два простых, но важных свойства этих идеальных метрик:

1) $\zeta_s(\lambda X, \lambda Y) = \lambda^s \zeta_s(X,Y)$ для всех $\lambda \ge 0$ (однородность порядка s);

2) $\zeta_s(X+X', Y+Y') \le \zeta_s(X,Y) + \zeta_s(X',Y')$, если случайные величины X, X' независимы, и также Y, Y' независимы (полуаддитивность).

В условиях центральной предельной теоремы с конечным 3-м моментом сразу получаем следующую теорему В. М. Золотарева.

Теорема 20.4. *Справедливо неравенство*

$$\zeta_3(F_n, \Phi) \,\le\, \frac{1}{\sqrt{n}}\,\zeta_3(F_1, \Phi).$$

Весьма интересно, что искомую скорость порядка $\frac{1}{\sqrt{n}}$ обеспечили вышеупомянутые свойства 1)-2) метрики ζ_3. В более общей ситуации справедливо неравенство

$$\zeta_s(F_n, \Phi) \,\le\, \frac{1}{n^{(s-2)/2}}\,\zeta_s(F_1, \Phi),$$

которое, правда, при $s \ge 4$ имеет смысл при дополнительных предположениях $\mathbf{E}\,|X_1|^s < \infty$, $\mathbf{E}X_1^p = \mathbf{E}Z^p$ $(p = 3, \dots, s-1)$, где Z – случайная величина со стандартным нормальным распределением. В частности, если $\beta_4 = \mathbf{E}X_1^4 < \infty$ и $\mathbf{E}X_1^3 = 0$ (например, когда исходное распределение F_1 симметрично относительно начала координат), то

$$\zeta_4(F_n, \Phi) \,=\, O\Big(\frac{1}{n}\Big).$$

Однако такое утверждение может быть существенно усилено в терминах метрики ζ_2. При конечности 4-го момента можно разложить $f_1(t)$ по формуле Тейлора в окрестности нуля по степеням t с остаточным членом вида $\beta_4 t^4$ и, как следствие, получить уточнение оценки (20.2), именно

$$|f_n^{(s)}(t) - \tilde g^{(s)}(t)| \,\le\, c\,\frac{\beta_4}{n}\,|t|^{4-s} e^{-t^2/4}, \qquad |t| \le \frac{\sqrt{n}}{\beta_3} \qquad (s = 0,1,2,3,4). \tag{20.6}$$

Здесь

$$\tilde g(t) = e^{-t^2/2}\Big(1 + \mathbf{E}X_1^3\,\frac{(it)^3}{3!}\,\frac{1}{\sqrt{n}}\Big)$$

– преобразование Фурье-Стилтьеса “поправленного"гауссовского распределения. В частности, $\tilde g(t) = g(t)$, если $\mathbf{E}X_1^3 = 0$.

Применяя (20.6) с $s = 1$ в теореме 9.1, можно доказать следующее неравенство, которое, по-видимому, не упоминается в литературе.

Теорема 20.5. *Если* $\mathbf{E}X_1^3 = 0$, $\beta_4 = \mathbf{E}X_1^4 < \infty$, *то с некоторой абсолютной постоянной* $c > 0$

$$\zeta_2(F_n, \Phi) \,\le\, c\,\frac{\beta_4}{n}.$$

Поскольку $\zeta_1^2 \,\le\, c'\zeta_2$, данная оценка согласуется с оценкой $\zeta_1(F_n, \Phi) \,\le\, c\,\frac{\beta_3}{\sqrt{n}}$ для среднего расстояния (асимптотически по n).

4. *Расстояния Канторовича.*

Для транспортных метрик Канторовича наилучшим известным результатом в задаче о скорости сходимости в центральной предельной теореме является следующее утверждение, доказанное Э. Рио [35] с использованием соотношения (10.3). Мы его приводим для случая одинаково распределенных слагаемых, как и в предыдущих теоремах (см. также [58]).

Теорема 20.6. *Если* $\beta_{p+2} = \mathbf{E}\,|X_1|^{p+2} < \infty$, $1 \le p \le 2$, *то*

$$W_p(F_n, \Phi) \,\le\, c_p\, \frac{\beta_{p+2}^{1/p}}{\sqrt{n}}, \tag{20.7}$$

где $c_p > 0$ *– постоянная, зависящая только от* p.

При $p = 1$ мы возвращаемся к известной оценке для среднего расстояния, а при $p = 2$ приходим к оценке для квадратичного расстояния Канторовича:

$$W_2(F_n, \Phi) \,\le\, c\, \frac{\sqrt{\beta_4}}{\sqrt{n}}. \tag{20.8}$$

Интересно, что конечность 4-й момента существенна для достижения стандартной скорости $\frac{1}{\sqrt{n}}$ в метрике W_2. Другой (т.н. энтропийный) подход к (20.8) предложен в [59]. Неравенство (20.8) можно также вывести с использованием оценки (20.2), если применить неравенство сглаживания теоремы 11.1 с $g_\alpha = \tilde{g}$ и $T = \frac{\sqrt{n}}{\beta_3}$. Что касается значений $p > 2$, (20.7) тоже сохраняет силу. С использованием разложений Эджворта такое неравенство недавно доказано в работе автора [60].

Наконец, отметим, что все перечисленные оценки сохраняют силу для сумм неодинаково распределенных слагаемых после соответствующих модификаций в терминах дробей Ляпунова. Имеется множество аналогичных результатов и для “сильных"метрик, таких как расстояние по вариации, расстояние Кульбака-Лейблера и других. Однако только моментных условий не достаточно для сходимости F_n к Φ с какой-либо скоростью в сильных метриках (см. [61]).

Автор выражает благодарность Ирине Шевцовой и Андрею Зайцеву за ценные замечания и стимулирующие обсуждения.

ЛИТЕРАТУРА

1. Феллер В. Введение в теорию вероятностей и ее приложения. Том 2. Изд. второе, М., Мир, 1967. Пер. с англ.

2. Петров В. В. Суммы независимых случайных величин. Москва, Физматлит, 1972, 416 с.

3. Петров В. В. Предельные теоремы для сумм независимых случайных величин. Москва, Наука, 1987.

4. Esseen C-G. Fourier analysis of distribution functions. A mathematical study of the Laplace-Gaussian law. Acta Math. 77 (1945), 1–125.

5. Berry A. C. The accuracy of the Gaussian approximation to the sum of independent variates. Trans. Amer. Math. Soc. 49 (1941), 122–136.

6. Файнлейб А. С. Обобщение неравенства Эссеена и его применение в вероятностной теории чисел. Изв. АН СССР. Сер. матем., 32:4 (1968), 859–879.

7. Bentkus V., Götze F. Optimal rates of convergence in the CLT for quadratic forms. Ann. Probab. 24 (1996), no. 1, 466–490.

8. Bohman H. Approximate Fourier analysis of distribution functions. Ark. Mat. 4 (1961), 99–157.

9. Золотарев В. М. О свойствах и связях некоторых типов метрик. Зап. научн. семин. ЛОМИ, 87 (1979), 18–35.

10. Линник Ю. В., Островский И. В. Разложения случайных величин и векторов. Москва, Наука, 1972, 480 с.

11. Золотарев В. М. Оценки различия распределений в метрике Леви. Труды МИАН СССР, 112 (1971), 224–231.

12. Зайцев А. Ю. О логарифмическом множителе в неравенствах сглаживания для расстояний Леви и Леви-Прохорова. Теория вероятн. и ее применен. 31:4 (1986), 782–784.

13. Ибрагимов И. А., Линник Ю. В. Независимые и стационарно связанные величины. Москва, Наука, 1965, 524 с.

14. Bobkov S. G., Chistyakov, G. P., Götze F. Berry-Esseen bounds in the entropic central limit theorem. Probab. Theory Related Fields, 159 (2014), no. 3–4, 435–478.

15. Прохоров Ю. В. Сходимость случайных процессов и предельные теоремы в теории вероятностей. Теория вероятн. и ее применен. 1 (1956), 177–238.

16. Юринский В. В. Неравенство сглаживания для оценок расстояния Леви-Прохорова. Теория вероятн. и ее применен. 20:1 (1975), 3–12.

17. Абрамов В. А. Оценка расстояния Леви-Прохорова. Теория вероятн. и ее применен. 21:2 (1976), 406–410.

18. Зайцев А. Ю. Оценки расстояния Леви-Прохорова в терминах характеристических функций и некоторые их применения. Проблемы теории вероятн. распределений.

VII, Зап. научн. семин. ЛОМИ, 119 (1982), 108–127.

19. Арак Т. В., Зайцев А. Ю. Равномерные предельные теоремы для сумм независимых случайных величин. Труды МИАН СССР, 174 (1986), 3–214.

20. Beckner W. Inequalities in Fourier analysis. Ann. of Math. (2) 102 (1975), no. 1, 159–182.

21. Esseen C-G. On mean central limit theorems. Kungl. Tekn. Hцgsk. Handl. Stockholm, no. 121, 1958, 30 pp.

22. Золотарев В. М. Идеальные метрики в проблеме аппроксимации распределений сумм независимых случайных величин. Теория вероятн. и ее применен. 22:3 (1977), 449–465.

23. Zolotarev V. M. Modern theory of summation of random variables. Modern Probability and Statistics. VSP, Utrecht, 1997, x+412 pp.

24. Сенатов В. В. Одна оценка метрики Леви-Прохорова. Теория вероятн. и ее применен. 29:1 (1984), 108–113.

25. Zolotarev V. M. Ideal metrics in the problems of probability theory. The Australian J. Statist. (21) 1979, no. 3, 193–208.

26. Vershik A. M. Long history of the Monge-Kantorovich transportation problem. Math. Intelligencer, 35 (2013), no. 4, 1–9.

27. Канторович Л. В. О перемещении масс. ДАН СССР, 37 (1942), №7–8, 227–229. См. также: Зап. научн. семин. ПОМИ, 312 (2004), 11–14.

28. Канторович Л. В. Об одной проблеме Монжа. Успехи матем. наук, 3 (1948), №2, 225–226. См. также: Зап. научн. семин. ПОМИ, 312 (2004), 15–16.

29. Рачев С. Т. Задача Монжа-Канторовича о перемещении масс и ее применение в стохастике. Теория вероятн. и ее применен. 29:4 (1984), 625–653.

30. Villani C. Topics in optimal transportation. Graduate Studies in Mathematics, vol. 58. American Mathematical Society, 2003.

31. Богачев В. И., Колесников, А. В. Задача Монжа–Канторовича: достижения, связи и перспективы. Успехи математ. наук, 407 (2012), том. 67:5, 3–110.

32. Bobkov S. G., Ledoux M. One-dimensional empirical measures, order statistics and Kantorovich transport distances. Preprint (2014).

33. Канторович Л. В., Акилов Г. П. Функциональный анализ. М., Наука, 3 изд., 1984, 752 с.

34. Dudley R. M. Real analysis and probability. The Wadsworth & Brooks/Cole, Mathematics Series, 1989.

35. Rio E. Upper bounds for minimal distances in the central limit theorem. Ann. Inst. Henri Poincaré, Probab. Stat. 45 (2009), no. 3, 802–817.

36. Talagrand M. Transportation cost for Gaussian and other product measures. Geom. Funct. Anal. 6 (1996), 587–600.

37. Bobkov S. G., Götze F. Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. J. Funct. Anal. 163 (1999), no. 1, 1–28.

38. Bobkov S. G., Gozlan, N., Roberto, C., Samson, P.-M. Bounds on the deficit in the logarithmic Sobolev inequality. J. Funct. Anal. 267 (2014), no. 11, 4110–4138.

39. Ingham A. E. A note on Fourier transforms. J. London Math. Soc. 9 (1934), 29–32.

40. Bhattacharya R. N., Ranga Rao R. Normal approximation and asymptotic expansions. John Wiley & Sons, Inc. 1976. New ed.: Soc. Industr. Appl. Math., Philadelphia, 2010.

41. Gautschi W. On inverses of Vandermonde and confluent Vandermonde matrices. Numer. Math. 4 (1962), 117–123.

42. Бобков С. Г. К одной теореме В. Н. Судакова о типичных распределениях. Зап. научн. семин. ПОМИ, 368 (2009), 59–74.

43. Pinelis I. On the nonuniform Berry–Esseen bound. Preprint (2013), arXiv:1301.2828.

44. Pinelis I. More on the nonuniform Berry–Esseen bound. Preprint (2013), arXiv:1302.0516.

45. Prawitz H. Limits for a distribution, if the characteristic function is given in a finite domain. Skand. Aktuarietidskr. 1972, 138–154 (1973).

46. Bohman H. To compute the distribution function when the characteristic function is known. Skand. Aktuarietidskr. 1963:41–46 (1964).

47. Vaaler J. D. Some extremal functions in Fourier analysis. Bull. Amer. Math. Soc. (N.S.) 12 (1985), no. 2, 183–216.

48. Мацкявичюс В. К. О нижней оценке скорости сходимости в центральной предельной теореме. Теория вероятн. и ее применен. 28:3 (1983), 565–569.

49. Шевцова И. Г. Об абсолютных константах в неравенстве типа Берри-Эссеена. Доклады РАН, 456 (2014), №6, 650–654.

50. Мешалкин Л. Д., Рогозин Б. А. Оценка расстояния между функциями распределений по близости их характеристических функций и ее применение в центральной предельной теореме. Предельные теоремы теории вероятностей, Ташкент: АН УзССР, 40–55, 1963.

51. Нагаев С. В. Некоторые предельные теоремы для больших уклонений. Теория вероятн. и ее применен. 10:2 (1965), 231–254.

52. Бикялис А. Оценки остаточного члена в центральной предельной теореме. Литовский математ. сборник, 6 (1966), №3, 323–346.

53. Paditz L. Abschätzungen der Konvergenzgeschwindigkeit zur Normalverteilung unter Voraussetzung einseitiger Momente. (German) Math. Nachr. 82 (1978), 131–156.

54. Paditz L. On the analytical structure of the constant in the nonuniform version of the Esseen inequality. Statistics 20 (1989), no. 3, 453–464.

55. Michel R. On the constant in the nonuniform version of the Berry–Esseen theorem. Z. Wahrsch. Verw. Gebiete 55 (1981), no. 1, 109–117.

56. Мирахмедов Ш. А. Об абсолютной константе в неравномерной оценке скорости

сходимости в центральной предельной теореме. Изв. Акад. Наук УзССР, сер. физ–мат. наук, №4, 26–30, 78, 1984.

56. Нефедова Ю. С., Шевцова И. Г. О точности нормальной аппроксимации для распределений пуассоновских случайных сумм. Информ. и ее применен. 6 (2011), вып. 1, 39–45.

57. Нефедова Ю. С., Шевцова И. Г. О неравномерных оценках скорости сходимости в центральной предельной теореме. Теория вероятн. и ее применен. 57:1 (2012), 62–97.

58. Rio E. Asymptotic constants for minimal distance in the central limit theorem. Electron. Commun. Probab. 16 (2011), 96–103.

59. Bobkov S. G. Entropic approach to E. Rio's central limit theorem for W_2-transport distance. Statist. Probab. Lett. 83 (2013), no. 7, 1644–1648.

60. Bobkov S. G. Berry-Esseen bounds and Edgeworth expansions in the central limit theorem for transport distances. Preprint (2016). To appear in: Probab. Theory Related Fields.

61. Сенатов В. В. Центральная предельная теорема. Точность аппроксимации и асимптотические разложения. М., Книжный дом ЛИБРОКОМ, 2008, 350 с.

Сергей Г. Бобков (Sergey G. Bobkov)

School of Mathematics
University of Minnesota
Vincent Hall, 206 Church St. S.E.
Minneapolis, MN 55455 USA

Email: bobkov@math.umn.edu