**ОБОЗРЕНИЕ
ПРИКЛАДНОЙ И ПРОМЫШЛЕННОЙ
Том 13 МАТЕМАТИКИ Выпуск 1
2006**

ГОЛЬДШТЕЙН Л., РАЙНЕРТ Г.

ПРЕОБРАЗОВАНИЯ РАСПРЕДЕЛЕНИЙ, ОРТОГОНАЛЬНЫЕ МНОГОЧЛЕНЫ И ХАРАКТЕРИЗАЦИИ СТЕЙНА

(Перевод с английского *Д. А. Жданова*)

Аннотация

Вводится новый класс преобразований распределений, который характеризуется уравнениями, связывающими взвешенные математические ожидания тестовых функций случайной величины с заданным распределением и взвешенные математические ожидания производных более высокого порядка тестовых функций случайной величины с преобразованным распределением. Этот класс включает в себя преобразования нулевого и объемного смещений. Когда в качестве тестовых функций рассматриваются многочлены, получается связь между семействами распределений, замкнутыми относительно взятия операции суммы независимых слагаемых (и в частности, бесконечно делимыми распределениями), и семейством преобразований, индуцированным связанными с ним ортогональными системами многочленов. Для семейств этих преобразований, обобщающих хорошо известное свойство объемного смещения, преобразование суммы независимых случайных величин получается заменой слагаемых, выбранных согласно многомерному распределению на индексном множестве, на независимые величины, распределениями которых являются распределения, полученные в ходе преобразования тем же самым семейством распределений. Ряд преобразований, связанных с классическими ортогональными системами многочленов, имеет в качестве неподвижной точки первоначальное распределение, либо член того же самого семейства, но уже с другим параметром.

Goldstein L., *Reinert G.* Distributional transformations, orthogonal polynomials, and Stein characterizations. — Journal of Theoretical Probability, 2005, v. 18, № 1, pp. 237–260.

© Larry Goldstein, Gesine Reinert, 2005 г.

© Перевод на русский язык с разрешения владельца прав. Редакция журнала «ОПиПМ», 2006 г.

[1]Department of Mathematics, University of Southern California, Los Angeles CA 90089-2532, USA

[2]Department of Statistics, University of Oxford, Oxford OX1 3TG, UK, supported in part by EPSRC grant no. GR/R52183/01

§ 1. Введение

Преобразование нулевого смещения было введено в [13]. Это отображение обладает свойствами, схожими со свойствами хорошо известного преобразования объемного смещения (см., например, [15]), но в отличие от преобразования объемного смещения, которое задается на неотрицательных величинах, его можно применять и к случайным величинам с нулевым средним. Отличительной особенностью преобразования нулевого смещения является то, что его единственной неподвижной точкой является нормальное распределение с нулевым средним. Именно по этой причине оно используется в методе Стейна для нормальной аппроксимации ([13], [11], [12] и [14]). Преобразование объемного смещения связано также с функциями K_i, рассматриваемыми в работах [16], [5], [22], [6] и других; хороший обзор приведен в [8].

Мы покажем, что классическое преобразование объемного смещения и преобразование нулевого смещения являются частными случаями преобразования распределений случайной величины X в $X^{(P)}$ при помощи измеримой функции «смещения» P. Поясним сказанное более точно. Для данных X и P обозначим $\mathcal{C}^m$ семейство функций, у которых существуют измеримые производные m-го порядка, и положим

$$\mathcal{F}^m(P) = \{F \in \mathcal{C}^m \colon \mathbf{E}\,|P(X)F(X)| < \infty\}.$$

Рассмотрим преобразования, которые обладают следующим свойством:

$$\mathbf{E}\,P(X)F(X) = \alpha\,\mathbf{E}\,F^{(m)}(X^{(P)}) \quad \text{для любой} \quad F \in \mathcal{F}^m(P), \tag{1}$$

где $\alpha = (m!)^{-1}\mathbf{E}\,P(X)X^m$, когда $X^m \in \mathcal{F}(P)$; потребуем $\alpha > 0$. Назовем данное распределение *смещенным распределением* X – P. В дискретном случае оператор дифференцирования заменяется на сеточный оператор (см. разделы 4.3 и 4.4).

В теореме 2.1 устанавливаются общие условия существования преобразований, удовлетворяющих (1). В этих условиях требуется только, чтобы функция «смещения» P m раз изменяла знак и удовлетворяла некоторым условиям ортогональности и положительности; назовем число m *порядком* получающегося преобразования. В частном случае, когда P есть многочлен, условие изменения знака можно записать в терминах корней и степени многочлена P, а свойства ортогональности при помощи моментов. Например, для каждого $m = 0, 1, \ldots$ существует преобразование распределений, которое определяется при помощи многочленов Эрмита порядка m как функция смещения. Область определения данного преобразования состоит из тех распределений, у которых первые $2m$ моментов совпадают с моментами стандартного нормального распределения; случай $m = 1$ соответствует преобразованию нулевого смещения [13].

В теореме 2.1 параграфа 2 доказано существование преобразований распределений при достаточно общих условиях. В § 3 показано, что для семейств преобразований, индуцированных ортогональной системой многочленов, существует некоторая дополнительная структура, в особенности для тех семейств распределений, которые замкнуты относительно операции суммирования независимых величин. Соответственно, для нормального, гамма, пуассоновского, биномиального и бета распределений, в § 4 изучаются семейства преобразований, определенные при помощи многочленов Эрмита, Лагерра, Шарлье, Кравчука и Гегенбауэра. Получены характеризационные уравнения стейновского типа более высокого порядка.

Данная статья написана в стиле работы [9], где впервые были описаны другие фундаментальные взаимосвязи между уравнениями Стейна и ортогональными многочленами. Подход, впервые изложенный в [9], был применен в [24], при этом была использована хорошо известная связь между ортогональными многочленами и процессами гибели и размножения. Этот же подход использовался (например, в [17]) для описания решений уравнений Стейна.

Сначала в статье дается обзор некоторых хорошо известных результатов, касающихся преобразования объемного смещения. Это преобразование является самым простым и хорошо изученным из всех преобразований распределений. Для неотрицательной случайной величины X такой, что $0 < \mathbf{E}\, X = \mu < \infty$, распределение случайной величины X^s, полученной X-объемным смещением, определяется следующим характеризационным уравнением:

$$\mathbf{E}\, XF(X) = \mu\, \mathbf{E}\, F(X^s) \quad \text{для любой} \quad F \in \mathcal{F}^0(X). \tag{2}$$

Ключевой особенностью распределения объемного смещения является следующая. Если $X_1, \ldots, X_n$ есть неотрицательные независимые случайные величины с конечными положительными ожиданиями $\mathbf{E}\, X_i = \mu_i$,

$$W = \sum_{i=1}^{n} X_i,$$

то случайную величину с распределением W-объемного смещения можно построить, заменяя величину X_i (выбранную с вероятностью, пропорциональной μ_i) независимой случайной величиной X_i^s, имеющей распределение X_i-объемного смещения. Другими словами, пусть I — случайная величина с распределением

$$\mathbf{P}\,\{I = i\} = \frac{\mu_i}{\sum_{j=1}^{n} \mu_j},$$

не зависящая от $X_1, \ldots, X_n$, тогда величина

$$W^s = W - X_I + X_I^s \tag{3}$$

имеет распределение W-объемного смещения. Обозначим $x^+ = \max\{0, x\}$, тогда объемное смещение есть случай (1) с функцией смещения $P(x) = x^+$. Это преобразование является преобразованием нулевого порядка, поскольку функция x^+ имеет $m = 0$ изменений знака на $\mathbf{R}$ и $\alpha = \mathbf{E}\,X^+$; если $X \geqslant 0$, то $X^+ = X$, и мы получаем обычную характеристику (2).

Введение преобразования нулевого смещения [13] было продиктовано схожестью между преобразованием объемного смещения и уравнением Стейна [21] для нормального распределения с нулевым средним. В частности, из тождества Стейна следует, что $Z \sim \mathcal{N}(0, \lambda)$ тогда и только тогда, когда

$$\mathbf{E}\,ZF(Z) = \lambda\,\mathbf{E}\,F'(Z) \quad \text{для любой} \quad F \in \mathcal{F}^1(Z). \tag{4}$$

Сравнивая (4) и (2), когда случайная величина X имеет нулевое среднее и положительную дисперсию λ, приходим к следующему определению. Будем говорить, что X^z имеет распределение X-объемного смещения, если

$$\mathbf{E}\,XF(X) = \lambda\,\mathbf{E}\,F'(X^z) \quad \text{для любой} \quad F \in \mathcal{F}^1(X). \tag{5}$$

Заметим, что в случае нулевого смещения, (5) есть то же самое, что и (2) в случае объемного смещения, только дисперсия заменяется средним, а F' заменяется F. Из характеризационного уравнения (4) следует, что нормальное распределение с дисперсией λ является неподвижной точкой преобразования объемного смещения. В [13] было показано, что преобразование объемного смещения X^z существует для любой случайной величины X, имеющей нулевое среднее и положительную дисперсию. Существование также следует из 2.1 как специальный случай (1) для функции $P(x) = x$, имеющей $m = 1$ изменений знака на $\mathbf{R}$, и α равняется дисперсии λ случайной величины X.

Преобразование объемного смещения было введено и использовалось в [13] с целью получить границы порядка n^{-1} в нормальной аппроксимации для гладких тестовых функций при условиях на момент третьего порядка в случае, когда имеется зависимость, порожденная процедурой простой случайной выборки. В [11] оно использовалось для установления границ нормального распределения для иерархических последовательностей, порожденных итерацией так называемой усредняющей функцией; в [12] для нормальной аппроксимации в комбинаторных центральных предельных теоремах со случайными перестановками, имеющими распределение, постоянное на всем цикле; в [14] рассматривалось существование преобразования объемного смещения более высокого порядка.

Преобразования нулевого смещения обладают свойством, аналогичным (3) для объемного смещения. В частности, в [13] было показано, что сумма независимых случайных величин с нулевым средним и конечной дисперсией может быть подвергнута преобразованию нулевого смещения, если заменить одну величину (выбранную с вероятностью, пропорциональной ее дисперсии) некой случайной величиной, независимой от тех слагаемых, которые составляют распределение нулевого смещения. Более точно, пусть $X_1, \ldots, X_n$ независимые случайные величины с нулевым средним и дисперсией $\lambda_i = \mathbf{E}\, X_i^2 > 0$,

$$W = X_1 + \cdots + X_n,$$

и I есть случайный индекс, независимый от $X_1, \ldots, X_n$, имеющий распределение

$$\mathbf{P}\,\{I = i\} = \frac{\lambda_i}{\sum_{j=1}^n \lambda_j}. \tag{6}$$

Тогда

$$W^z = W - X_I + X_I^z$$

имеет распределение W-объемного смещения, где X_i^z есть случайная величина, независимая от X_j, $j \neq i$, и имеющая распределение X_i-объемного смещения. Эта конструкция распространяется на семейства преобразований, связанных с ортогональными многочленами в теореме 3.1. В частности, в § 3 мы увидим, что для преобразований более высокого порядка суммы независимых случайных величин преобразуются заменой множественных величин, взятых из некоторого распределения (например, мультиномиальное, многомерное гипергеометрическое), на независимые случайные величины, имеющие распределения, которые получаются преобразованием того же самого семейства.

В § 2 даются условия на моменты случайной величины X и изменения знаков многочлена P, которые гарантируют существование распределения $X - P$ и приводится явное построение. В § 3 рассматривается специальный случай, когда P принадлежит семейству ортогональных многочленов. Обобщение на случай нескольких случайных величин конструкций нулевого и объемного смещений дается в § 3 и основывается на тождестве (25), выражающем ортогональный многочлен суммы как сумму подобных многочленов, причем степень слагаемых не превосходит степень многочлена. В разделах 4.1–4.5 рассматриваются многочлены Эрмита, Лагерра, Шарлье, Кравчука и Гегенбауэра, соответствующие нормальному, гамма, пуассоновскому, биномиальному и бета распределениям, соответственно. В качестве специальных примеров бета распределений мы рассмотрим равномерное распределение $\mathcal{U}[-1, 1]$, распределение арксинуса и полукруговое распределение.

§ 2. Преобразования. Общие сведения

Начнем наше рассмотрение с теоремы существования и единственности для рассматриваемых типов преобразований распределений. Будем говорить, что измеримая функция P, заданная на $\mathbf{R}$, *положительна* на интервале I, если $P(x) \geqslant 0$ для любого $x \in I$, причем строгое неравенство верно хотя бы для одной точки x. Точно также определяется *отрицательная* P на интервале I. Будем говорить, что P ровно $m = 0, 1, \ldots$ раз *изменяет знак*, если $\mathbf{R}$ можно разбить на $m+1$ таких непустых непересекающихся подынтервалов, что P чередует знак на последовательных интервалах. Хотя выбрать конечные точки для интервалов, на которых P равняется нулю, можно произвольным образом, тем не менее будем считать, что изменение знака происходит на границах интервала. Из единственности, установленной в теореме 2.1, следует, что смещенное распределение $X - P$, построенное в ходе доказательства теоремы 2.1, будет тем же самым для любого выбора границы интервала. В примере 2.1 дается некоторое дополнительное объяснение этого факта для некоторого частного случая. Заметим, что для существования искомого распределения, несмотря на некоторые ограничения, сформулированные в теореме 2.1 условия ортогональности касаются только лишь P и требуются для конечных порядков; такие условия не обязательно накладывают ограничения на моменты случайной величины X, как показано в примере 2.1.

Теорема 2.1. *Пусть X есть случайная величина, $m \in \{0, 1, 2, \ldots\}$, измеримая функция P имеет ровно m изменений знака, причем на самом правом интервале имеет положительный знак и*

$$\frac{1}{m!}\mathbf{E}\, X^k P(X) = \alpha \delta_{k,m}, \qquad k = 0, \ldots, m, \tag{8}$$

где $\alpha > 0$. Тогда для такой случайной величины $X^{(P)}$, что

$$\mathbf{E}\, P(X)F(X) = \alpha\, \mathbf{E}\, F^{(m)}(X^{(P)}), \qquad \text{для любой } F \in \mathcal{F}^m(P), \tag{9}$$

существует единственное распределение.

В теореме 2.1 утверждается, что если смещающая функция P имеет m изменений знака, и все степени X, меньшие m-й степени, ортогональны $P(X)$ в смысле $\mathbf{L}^2(X)$ (а именно, когда $P(X) \in \{1, X, \ldots, X^{m-1}\}^{\perp}$ и $\mathbf{E}\, X^m P(X) > 0$), тогда случайная величина X принадлежит области определения преобразования распределений порядка m, определяемого при помощи этой смещающей функции. Как уже упоминалось выше, специального рассмотрения заслуживают преобразование растяжения и преобразование с нулевым сдвигом.

Доказательство. Приведем явное построение случайной величины $X^{(P)}$. Сделаем замену P на P/α, тогда теорему будет достаточно доказать для случая $\alpha = 1$. Обозначим $r_1, \ldots, r_m$ точки изменения знака P, и пусть

$$Q(x) = \prod_{i=1}^{m} (x - r_i), \tag{10}$$

причем пустое произведение равно 1. Пусть μ_X есть распределение X. По построению Q и P одного знака, поэтому

$$d\mu_Y(y) = \frac{1}{m!} Q(y)P(y)\, d\mu_X(y) \tag{11}$$

есть мера. Поскольку из (8) при $k = m$ следует, что $\mathbf{E}\, Q(X)P(X) = m!$, (11) будет и вероятностной мерой. Далее, Y и $\{U_i\}_{i\geqslant 1}$ взаимно независимы, причем Y имеет распределение μ_Y, а U_j имеет функцию распределения u^i на $[0,1]$, где $r_0 = Y$ и $r_{m+1} = 0$. Покажем, что

$$X^{(P)} = \sum_{k=1}^{m+1} \left(\prod_{i=k}^{m} U_i \right) (r_{k-1} - r_k) \tag{12}$$

удовлетворяет (9), тем самым доказывая существование смещенного распределения $X - P$.

Для $k = 0, \ldots, m$ положим

$$R_k(x) = \prod_{i=k+1}^{m} (x - r_i),$$

многочлен степени $m - k$. Заметим сначала, что для некоторого F, для которого хотя бы одна из частей

$$\mathbf{E}\, F(Y) = \frac{1}{m!} \mathbf{E}\, F(X)Q(X)P(X) \tag{13}$$

существует, из (13) и (8) имеем

$$\mathbf{E} \left(1 / \prod_{i=1}^{k} (Y - r_i) \right) = \frac{1}{m!}\, \mathbf{E}\, R_k(X)P(X) = \delta_{m-k,m}. \tag{14}$$

Будем проводить доказательство по индукции. В частности, для $k \geqslant 1$, положив

$$V_k = \prod_{i=k}^{m} U_i, \quad W_k = \sum_{j=k}^{m+1} V_j (r_{j-1} - r_j), \tag{15}$$

и взяв $X^{(P)}$ вида (12), покажем, что для любой $F \in \mathcal{C}_c^\infty$ (семейство бесконечное раз дифференцируемых функций с компактным носителем) и $k = 0, \ldots, m$

$$\mathbf{E}\, F^{(m)}(X^{(P)}) = k!\, \mathbf{E} \left\{ \frac{F^{(m-k)}(V_{k+1}(Y - r_{k+1}) + W_{k+2})}{V_{k+1}^k \prod_{i=1}^{k} (Y - r_i)} \right\}. \tag{16}$$

Из (14) видно, что математическое ожидание, стоящее в правой части данного выражения, существует, поскольку F и все ее производные ограниченны, V_{k+1} не зависит от Y для любого k и $\mathbf{E}\, U_i^{-k} < \infty$ для $i \geqslant k+1$.

Случай $k = 0$ представляет собой утверждение $X^{(P)} = V_1(Y - r_1) + W_2$, которое следует из определений (12) и (15). Предположим теперь, что (16) выполняется для некоторого $0 \leqslant k < m$. Подставив $V_{k+1} = U_{k+1}V_{k+2}$ в (16) и взяв математическое ожидание по U_{k+1}, которое имеет плотность $(k+1)u_{k+1}^k$, получаем

$$\mathbf{E}\,F^{(m)}(X^{(P)})$$
$$= (k+1)!\,\mathbf{E}\int_0^1 \left\{ \frac{F^{(m-k)}(u_{k+1}V_{k+2}(Y - r_{k+1}) + W_{k+2})}{u_{k+1}^k V_{k+2}^k \prod_{i=1}^k (Y - r_i)} \right\} u_{k+1}^k\, du_{k+1}.$$

Сокращая на u_{k+1}^k и интегрируя, получаем

$$(k+1)!\,\mathbf{E}\left\{ \frac{F^{(m-(k+1))}(V_{k+2}(Y - r_{k+1}) + W_{k+2}) - F^{(m-(k+1))}(W_{k+2})}{V_{k+2}^{k+1}\prod_{i=1}^{k+1}(Y - r_i)} \right\}. \tag{17}$$

Так как $k+1 \geqslant 1$, то используя независимость V_{k+2} и Y для любого k, независимость W_{k+2} и Y для любого $k \geqslant 0$, и (14), имеем, что второй член под знаком математического ожидания в (17) равен нулю. Доказательство по индукции завершается, если заметить, что из обозначений (15) следует $V_{k+2}(Y - r_{k+1}) + W_{k+2} = V_{k+2}(Y - r_{k+2}) + W_{k+3}$.

Подставим теперь в (16) $k = m$. Поскольку $V_{m+1} = 1, W_{m+2} = 0$ и $r_{m+1} = 0$, с учетом (13) получаем

$$\mathbf{E}\,F^{(m)}(X^{(P)}) = m!\mathbf{E}\left\{\frac{F(Y)}{Q(Y)}\right\} = \mathbf{E}\,P(X)F(X),$$

значит, равенство в (9) выполняется для любой $F \in \mathcal{C}_c^\infty$.

В случае же $F \in \mathcal{F}^m(X)$ заменим F на

$$F(x) - \sum_{j=0}^{m-1} \frac{F^{(j)}(0)}{j!} x^j.$$

В силу (8), не ограничивая общности, можно считать, что $F^{(j)}(0) = 0$, $j = 0, \ldots, m-1$. И поэтому для любой измеримой функции f

$$F(x) = I^m f, \quad \text{где} \quad If = \int_0^x f.$$

Поскольку $F = F_1 - F_2$, где $F_1(x) = I^m f^+$ и $F_2(x) = I^m f^-$, в силу линейности достаточно рассматривать $f \geqslant 0$. Пусть $0 \leqslant f_n \uparrow f$, тогда $I^m f_n = F_n \uparrow F$. Используя теоремы о монотонной и мажорируемой сходимостях, соответственно, для левой и правой частей (9), получаем, что равенство в (9) выполняется для $F \in \mathcal{F}^m(X)$.

Распределение $X^{(P)}$ единственно, поскольку (9) выполняется для любой $F \in \mathcal{C}_c^\infty$, что и требовалось доказать. ■

Существование распределения $X^{(P)}$ следует из теоремы представления Рисса, если показать линейность оператора T, заданного на пространстве непрерывных функций с компактным носителем $\mathcal{C}_c^0$, определенного следующим образом:

$$Tf = \mathbf{E}\, P(X)F(X), \quad \text{где} \quad F(x) = I^m f.$$

Заряд $d\mu = P d\mu_X$ обладает свойством $\int x^j\, d\mu = \mathbf{E}\, X^j P(X) = 0$, $j = 0, 1, \ldots, m-1$. Поэтому в силу свойства изменения знака функции P, чтобы показать положительность T (а значит, и $Tf = \int_a^b f\, d\mu^{(m)}$ для любой меры $\mu^{(m)}$, которая является и вероятностной мерой, поскольку $\mathbf{E}\, X^m P(X) = m!$), можно воспользоваться на конечном интервале $[a, b]$ теоремой 5.4 из главы XI [19] (см. также пример 1.4 в главе XI). Аналогичными рассуждениями пользовались мы в [13] при доказательстве существования распределения с нулевым смещением для нулевого среднего и конечной вариации X: заметим, что $F = If$ не уменьшается, если $f \geqslant 0$, и поэтому X и $F(X)$ положительно коррелированы, откуда оператор

$$Tf = \mathbf{E}\, XF(X) \geqslant \mathbf{E}\, X\, \mathbf{E}\, F(X) = 0$$

положителен.

П р и м е р 2.1. Рассмотрим применение теоремы 2.1 в случае, когда $P(x)$ имеет ровно $m = 1$ изменение знака в точке $r_1 = 0$. Потребуем $\mathbf{E}\, P(X) = 0$ и $\alpha = \mathbf{E}\, XP(X) > 0$ для того, чтобы невырожденное X принадлежало области преобразования, обладающего свойством

$$\mathbf{E}\, P(X)F(X) = \alpha \mathbf{E}\, F'(X^{(P)}). \tag{18}$$

Вспомним, что при доказательстве теоремы мы промасштабировали случайную величину X так, чтобы $\alpha = 1$. Имеем $Q(x) = x$ в (10) и, значит, распределение Y в (11) имеет вид

$$d\mu_Y(y) = xP(x)\, d\mu_X(y)/\alpha.$$

Подставляя в (12) $m = 1$, $r_0 = Y, r_2 = 0$ и U_j с функцией распределения u^j на $[0, 1]$, получаем

$$X^{(P)} = \sum_{k=1}^{m+1} \left(\prod_{i=k}^{m} U_i \right) (r_{k-1} - r_k) = U_1(r_0 - r_1) + (r_1 - r_2) = U_1 Y.$$

Поскольку $X^{(P)}$ абсолютно непрерывно, можно непосредственно проверить, что плотность задается формулой

$$f^{(P)}(x) = \alpha^{-1}\, \mathbf{E}\, (P(X) \,|\, X > x). \tag{19}$$

Если $\int_0^x P(u)\, du$ конечен для любого x и $c = \int \exp\{-\alpha^{-1} \int_0^x P(u)\, du\}\, dx < \infty$, то преобразование (18) имеет неподвижную точку в распределении с плотностью

$$f(x) = c^{-1} \exp\left\{-\frac{1}{\alpha}\int_0^x P(u)\,du\right\}.$$

Например, когда $P(x) = x$, то f есть плотность нормального распределения с нулевым средним и дисперсией α.

Взяв P в виде функции знака:

$$P(x) = \mathbf{1}\{x > 0\} - \mathbf{1}\{x < 0\},$$

получаем пример преобразования, задаваемого дискретной функцией P. Следовательно, в общем случае условия ортогональности не могут быть сведены к ограничениям на моменты X, в частности, в (8) при $k = 0$ необходимо, чтобы X имело медиану, равную 0. Если к тому же $\alpha = \mathbf{E}\,|X|$ конечно, что получается из (8) при $k = 1$, то из теоремы 2.1 следует, что X принадлежит области определения отображения, удовлетворяющего (18). Из (19) получаем, что плотность преобразованного распределения есть

$$f^{(P)}(x) = \begin{cases} \mathbf{P}\{X > x\}/\mathbf{E}\,|X|, & x > 0, \\ \mathbf{P}\{X < x\}/\mathbf{E}\,|X|, & x < 0. \end{cases} \tag{20}$$

При таком выборе P для распределения Y в (11) получаем

$$d\mu_Y(y) = |y|\,d\mu_X(y)/\mathbf{E}\,|X|,$$

которое есть распределение $|X|$-объемного смещения. Поэтому смещенное распределение $X - P$ получается умножением $Y \sim \mu_Y$ на независимую величину $\mathcal{U}[0,1]$. Это преобразование имеет неподвижную точку в распределении Лапласа с плотностью

$$f(x) = \frac{1}{2\alpha}\exp\left\{-\frac{1}{\alpha}|x|\right\}.$$

Полагая $P(x) = \mathbf{1}\{x > 1\} - \mathbf{1}\{x < -1\}$, получаем преобразование, имеющее своей областью определения те величины X, для которых $\alpha = \mathbf{E}\,(|X|\mathbf{1}\{|X| > 1\}) < \infty$, и удовлетворяющее

$$\mathbf{P}\{X > 1\} = \mathbf{P}\{X < -1\}. \tag{21}$$

Поскольку $P(x) = 0$ на множестве $[-1, 1]$, можно сказать, что изменение знака появляется в точке из интервала $(-1, 1)$, и многочлен Q в доказательстве теоремы 2.1 можно взять в виде

$$Q(x) = x - r_1 \quad \text{для любого} \quad r_1 \in (-1, 1).$$

Так как мы имеем единственность, то распределение, построенное при доказательстве теоремы 2.1, не зависит от выбора r_1; действительно, в этом случае из (21) следует, что распределение $d\mu_Y$ в (11) одинаково для любого $r_1 \in (-1, 1)$.

§ 3. Преобразования, использующие ортогональные многочлены

Рассмотрим систему многочленов, ортогональных нетривиальному семейству распределений $Z_\lambda \sim \mathcal{L}_\lambda$, индексированному действительным параметром λ.

У с л о в и е 3.1. Для некоторого $m \geqslant 0$ многочлены $\{P_\lambda^k(x)\}_{0 \leqslant k \leqslant m}$ нормированны, имеют степень k, ортогональны семейству распределений $Z_\lambda \sim \mathcal{L}_\lambda$ и удовлетворяют неравенству $\mathbf{E}\,[P_\lambda^k(Z_\lambda)]^2 > 0$.

Заметим, что поскольку P_λ^k есть нормированный и ортогональный многочлен, он имеет k различных корней и положителен при $x \to \infty$ (см., например, [1]). Далее имеем

$$\mathbf{E}\, Z_\lambda^k P_\lambda^k(Z_\lambda) = \mathbf{E}\,[P_\lambda^k(Z_\lambda)]^2, \qquad k = 0, \ldots, m.$$

Когда мы изучаем преобразования, в которых используются семейства ортогональных многочленов, заданных в неявном виде, полученное распределение $X_\lambda^{(k)}$ индексируем параметром λ и степенью k многочлена.

Применяя теорему 2.1 в этом контексте, получаем следующее

Следствие 3.1. *Пусть выполнены условие* 3.1 *и* $\mathbf{E}\, Z_\lambda^{2m} < \infty$. *Для* $0 \leqslant k \leqslant m$ *положим*

$$\alpha_\lambda^{(k)} = \frac{1}{k!}\,\mathbf{E}\, Z_\lambda^k P_\lambda^k(Z_\lambda). \tag{22}$$

Тогда для всех $X \in \mathcal{M}_\lambda^k$, *где*

$$\mathcal{M}_\lambda^k = \{X\colon\, \mathbf{E}\, X^j = \mathbf{E}\, Z_\lambda^j,\ 0 \leqslant j \leqslant 2k\},$$

существует такая случайная величина $X_\lambda^{(k)}$, *что для любой функции* $F \in \mathcal{F}^k(P_\lambda^k)$

$$\mathbf{E}\, P_\lambda^k(X) F(X) = \alpha_\lambda^{(k)} \mathbf{E}\, F^{(k)}(X_\lambda^{(k)}). \tag{23}$$

Д о к а з а т е л ь с т в о. Поскольку $X \in \mathcal{M}_\lambda^k$, из условия 3.1 и ортогональности для $0 \leqslant j \leqslant k \leqslant m$ имеем

$$\frac{1}{k!}\mathbf{E}\, X^j P_\lambda^k(X) = \frac{1}{k!}\mathbf{E}\, Z_\lambda^j P_\lambda^k(Z_\lambda) = \frac{1}{k!}\mathbf{E}\, P_\lambda^j(Z_\lambda) P_\lambda^k(Z_\lambda) = \alpha_\lambda^{(k)} \delta_{j,k}.$$

Теперь остается воспользоваться теоремой 2.1. ■

Будем говорить, что семейство распределений Z_λ *замкнуто* относительно операции суммирования независимых слагаемых, если для независимых $Z_{\lambda_i} \sim \mathcal{L}_{\lambda_i}, i = 1, 2$, выполняется $Z_{\lambda_1} + Z_{\lambda_2} \sim \mathcal{L}_{\lambda_1 + \lambda_2}$. Существует специальная конструкция, когда функция преобразования в теореме 2.1 является членом системы многочленов, ортогональных данному семейству. В частности, следующая теорема 3.1 обобщает (3) и (7). В ней показано, как можно сумму независимых случайных величин P_λ^m-преобразовать, заменив случайно выбранную совокупность в сумме величинами, имеющими распределения, полученные с использованием той же самой ортогональной системы многочленов.

Пусть $n = 1, 2, \ldots$, рассмотрим мультииндекс $\mathbf{m} = (m_1, \ldots, m_n)$. Пусть

$$m = |\mathbf{m}| = \sum_{i=1}^n m_i, \quad \lambda = \sum_{i=1}^n \lambda_i,$$

и положим

$$\alpha_{\boldsymbol{\lambda}}^{(\mathbf{m})} = \prod_{i=1}^n \alpha_{\lambda_i}^{(m_i)}, \qquad P_{\boldsymbol{\lambda}}^{\mathbf{m}}(\mathbf{x}) = \prod_{i=1}^n P_{\lambda_i}^{m_i}(x_i), \tag{24}$$

где $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)$ и $\mathbf{x} = (x_1, \ldots, x_n)$.

Теорема 3.1. *Пусть Z_λ, $\lambda > 0$, есть семейство случайных величин, замкнутое относительно независимого сложения и $\mathbf{E}\, Z_\lambda^{2m} < \infty$. Предположим, что ассоциированные с этим семейством ортогональные многочлены $\{P_\lambda^k(x)\}_{0 \leqslant k \leqslant m}$ удовлетворяют условию 3.1 и, для некоторых весов $c_\mathbf{m}$, тождеству*

$$P_\lambda^m(w) = \sum_{\mathbf{m}\colon |\mathbf{m}|=m} c_\mathbf{m} P_{\boldsymbol{\lambda}}^{\mathbf{m}}(\mathbf{x}), \tag{25}$$

где $P_{\boldsymbol{\lambda}}^{\mathbf{m}}(\mathbf{x})$ задано в (24) и $w = x_1 + \cdots + x_n$. Тогда $\alpha_\lambda^{(m)}$ и $\alpha_{\boldsymbol{\lambda}}^{(\mathbf{m})}$, определенные в (22) и (24) соответственно, удовлетворяют

$$\alpha_\lambda^{(m)} = \sum_{\mathbf{m}\colon |\mathbf{m}|=m} c_\mathbf{m} \alpha_{\boldsymbol{\lambda}}^{(\mathbf{m})} \tag{26}$$

и существует величина $\mathbf{I}$, независимая от всех других величин, имеющая распределение

$$\mathbf{P}\{\mathbf{I} = \mathbf{m}\} = c_\mathbf{m} \frac{\alpha_{\boldsymbol{\lambda}}^{(\mathbf{m})}}{\alpha_\lambda^{(m)}}, \qquad |\mathbf{m}| = m. \tag{27}$$

Более того, для любых положительных $\lambda_1, \ldots, \lambda_n$ и независимых величин $X_1, \cdots, X_n$, где $X_i \in \mathcal{M}_{\lambda_i}^m$ и $W = \sum_{i=1}^n X_i$, величина

$$W_\lambda^{(m)} = \sum_{\mathbf{m}\colon |\mathbf{m}|=m} (X_i)_{\lambda_i}^{(I_i)}$$

имеет распределение $W - P_\lambda^m$.

Доказательство. Так как $X_i \in \mathcal{M}_{\lambda_i}^m$, то для $0 \leqslant k \leqslant 2m$, независимых $Z_{\lambda_i} \sim \mathcal{L}_{\lambda_i}$ и $Z_\lambda \sim \mathcal{L}_\lambda$ имеем

$$\mathbf{E}\, W^k = \mathbf{E}\left(\sum_{i=1}^n X_i\right)^k = \mathbf{E}\left(\sum_{i=1}^n Z_{\lambda_i}\right)^k = \mathbf{E}\, Z_\lambda^k.$$

Отсюда $W \in \mathcal{M}_\lambda^m$, и распределение $W^{(m)}$ существует из следствия 3.1. Равенство (26) получается, если (25) умножим на $W^m = (\sum_i X_i)^m$, потом возьмем математическое ожидание и воспользуемся независимостью и ортогональностью.

Из (26) для любого $F \in \mathcal{C}_c^\infty$ получаем

$$\alpha_\lambda^{(m)} \mathbf{E}\, F^{(m)}(W_\lambda^{(m)}) = \mathbf{E} \sum_{\mathbf{m}} c_{\mathbf{m}} \alpha_{\boldsymbol{\lambda}}^{(\mathbf{m})} F^{(m)}(W_\lambda^{(m)}). \tag{28}$$

Используя независимость и последовательно применяя тождество

$$\alpha_{\lambda_i}^{(m_i)} \mathbf{E}\, F^{(q)}((X_i)_{\lambda_i}^{(m_i)} + y) = \mathbf{E}\, P_{\lambda_i}^{m_i}(X_i) F^{(q-m_i)}(X_i + y), \tag{29}$$

получим, что правая часть (28) равна

$$\mathbf{E} \sum_{\mathbf{m}} c_{\mathbf{m}} P_{\boldsymbol{\lambda}}^{\mathbf{m}}(\mathbf{X}) F(W) = \mathbf{E}\, P_\lambda^m(W) F(W), \tag{30}$$

здесь мы также воспользовались (25). Сравнивая (28) и (30), получаем

$$\alpha_\lambda^{(m)} \mathbf{E}\, F^{(m)}(W_\lambda^{(m)}) = \mathbf{E}\, P_\lambda^m(W) F(W),$$

для любой $F \in \mathcal{C}_c^\infty$. Следовательно, $W_\lambda^{(m)}$ имеет смещенное распределение W-P_λ^m. ■

Для системы (возможно бесконечной) нормированных многочленов $\{P_\lambda^m(x)\}$, ортогональных $\mathcal{L}_\lambda$, определим производящую функцию

$$\phi_t(x, \lambda) = \sum_{m \geqslant 0} P_\lambda^m(x) \frac{t^m}{m!}. \tag{31}$$

Константы $\alpha_\lambda^{(m)}$ можно найти, если подставить $F(x) = x^m$ в (23); другой способ заключается в возведении (31) в квадрат, взятии математического ожидания и использовании ортогональности:

$$\mathbf{E}\,[\phi_t(Z_\lambda, \lambda)]^2 = \sum_{m \geqslant 0} \alpha_\lambda^{(m)} \frac{t^{2m}}{m!}. \tag{32}$$

Следующая теорема 3.2 используется в специальных случаях, рассматриваемых в разделах 4.1–4.4.

Теорема 3.2. *Если производящая функция* (31) *удовлетворяет условию*

$$\phi_t(w, \lambda) = \prod_{i=1}^n \phi_t(x_i, \lambda_i), \tag{33}$$

где $w = x_1 + \dots + x_n$ *и* $\lambda = \lambda_1 + \dots + \lambda_n$, *тогда* (25), *а значит и* (27), *в теореме* 3.1 *имеют вид, соответственно,*

$$c_{\mathbf{m}} = \binom{m}{\mathbf{m}} \quad \text{и} \quad \mathbf{P}\,\{\mathbf{I} = \mathbf{m}\} = \binom{m}{\mathbf{m}} \frac{\alpha_{\boldsymbol{\lambda}}^{(\mathbf{m})}}{\alpha_\lambda^{(m)}}, \qquad |\mathbf{m}| = m.$$

Доказательство. Перепишем (33) в следующем виде:

$$\begin{aligned}\sum_{m\geqslant 0}\frac{t^m}{m!}P_\lambda^m(w) &= \prod_{i=1}^n\sum_{m_i\geqslant 0}P_{\lambda_i}^{m_i}(x_i)\frac{t^{m_i}}{m_i!}\\ &= \sum_{m_1,\cdots,m_n}P_{\boldsymbol\lambda}^{\mathbf m}(\mathbf x)\frac{t^{m_1+\cdots+m_n}}{m_1!\cdots m_n!}\\ &= \sum_{m\geqslant 0}^{\infty}\frac{t^m}{m!}\sum_{\mathbf m=m}\binom{m}{\mathbf m}P_{\boldsymbol\lambda}^{\mathbf m}(\mathbf x).\end{aligned}$$

Отсюда имеем (25) со значениями, указанными в теореме. ■

Заметим также, что возведение в квадрат (25), взятие ожидания и использование независимости и ортогональности приводит к следующей формуле:

$$\alpha_\lambda^{(m)} = \sum_{|\mathbf m|=m}\binom{m}{\mathbf m}^{-1}c_{\mathbf m}^2\alpha_{\boldsymbol\lambda}^{\mathbf m}. \tag{34}$$

Так что результат теоремы 3.2 можно получить, если приравнять коэффициенты в (26) и (34), когда $\alpha_{\boldsymbol\lambda}^{\mathbf m}$ принимает достаточно много значений.

Завершим этот раздел приведением результата о потенциале повторного смещения.

Теорема 3.3. *Пусть выполняется условие* 3.1. *Предположим, что семейство распределений Z_λ замкнуто относительно преобразования многочленов $P_\lambda^k(x)$, т. е. существует такое $\mu(\lambda,k)$, что*

$$(Z_\lambda)_\lambda^{(k)} = Z_{\mu(\lambda,k)}.$$

Тогда, если $X\in\mathcal M_\lambda^m$, то $X_\lambda^{(k)}\in\mathcal M_{\mu(k,\lambda)}^{m-k}$ для $k\leqslant m$. В частности, для такого неотрицательного j, что $0\leqslant k+j\leqslant m$, существует распределение $(X_\lambda^{(k)})_{\mu(\lambda,k)}^{(j)}$.

Доказательство. Пусть $0\leqslant j\leqslant 2(m-k)$ и $F(x)=x^{k+j}/(k+j)_k$, где $(x)_k=x(x-1)\cdots(x-k+1)$. Тогда

$$\begin{aligned}\alpha_\lambda^{(k)}\,\mathbf E\,(X_\lambda^{(k)})^j &= \alpha_\lambda^{(k)}\,\mathbf E\,F^{(k)}(X_\lambda^{(k)}) = \mathbf E\,P_\lambda^k(X)F(X)\\ &= \mathbf E\,P_\lambda^k(Z_\lambda)F(Z_\lambda) = \alpha_\lambda^{(k)}\,\mathbf E\,F^{(k)}((Z_\lambda)_\lambda^{(k)}) = \alpha_\lambda^{(k)}\,\mathbf E\,(Z_{\mu(\lambda,k)})^j.\end{aligned}$$

Таким образом, первые $2(m-k)$ моментов $X_\lambda^{(k)}$ совпадают с соответствующими моментами $Z_{\mu(\lambda,k)}$, и существование распределения $(X_\lambda^{(k)})_{\mu(\lambda,k)}^{(j)}$ следует из 3.1. ■

§ 4. Специальные системы ортогональных многочленов

В разделах 4.1–4.5 мы ограничимся рассмотрением классических семейств многочленов Эрмита, Лагерра, Шарлье, Кравчука и Гегенбауэра, ортогональных нормальному, гамма, пуассоновскому, биномиальному и бета распределениям, соответственно. Все эти семейства удовлетворяют условию 3.1 и, за исключением последнего случая, имеют производящую функцию, удовлетворяющую (33). Нормальное и пуассоновское распределения являются неподвижными точками соответствующих преобразований. В случаях гамма, биномиального и бета распределений соответствующие преобразования ставят в соответствие те же самые семейства, но со смещенным параметром. Дальнейшие сведения о связях между вероятностными распределениями и производящими функциями таких систем многочленов можно посмотреть в [2] и [3].

4.1. Многочлены Эрмита. Пусть $\sigma^2 = \lambda > 0$, определим совокупность многочленов Эрмита $\{H_\lambda^m(x)\}_{m\geqslant 0}$ через производящую функцию

$$e^{xt-\frac{1}{2}\lambda t^2} = \sum_{m=0}^{\infty} H_\lambda^m(x)\frac{t^m}{m!} \tag{35}$$

или, что эквивалентно, через формулу Родрига

$$H_\lambda^m(x) = (-\lambda)^m e^{\frac{x^2}{2\lambda}}\,\frac{d^m}{dx^m}e^{-\frac{x^2}{2\lambda}}. \tag{36}$$

Эти многочлены ортогональны нормальному распределению $\mathcal{N}(0,\lambda)$ с плотностью $(2\pi\lambda)^{-1/2}\exp\{-x^2/(2\lambda)\}$.

Пусть $F \in \mathcal{C}_c^\infty$ и $Z_\lambda \sim \mathcal{N}(0,\lambda)$, применяя формулу Родрига (36), получаем

$$\begin{aligned}\mathbf{E}\,H_\lambda^m(Z_\lambda)F(Z_\lambda) &= \int_{-\infty}^{\infty}(-\lambda)^m e^{\frac{x^2}{2\lambda}}\left(\frac{d^m}{dx^m}e^{-\frac{x^2}{2\lambda}}\right)F(x)\frac{e^{-\frac{x^2}{2\lambda}}}{\sqrt{\lambda 2\pi}}\,dx\\ &= \int_{-\infty}^{\infty}(-\lambda)^m\left(\frac{d^m}{dx^m}e^{-\frac{x^2}{2\lambda}}\right)F(x)\frac{1}{\sqrt{\lambda 2\pi}}\,dx\\ &= \lambda^m\int_{-\infty}^{\infty}F^{(m)}(x)\frac{e^{-\frac{x^2}{2\lambda}}}{\sqrt{\lambda 2\pi}}\,dx\\ &= \lambda^m\,\mathbf{E}\,F^{(m)}(Z_\lambda).\end{aligned} \tag{37}$$

Отсюда,

$$(Z_\lambda)_\lambda^{(m)} = Z_\lambda,$$

т. е. для каждого $m = 0, 1, \ldots,$ нормальная случайная величина $Z_\lambda \sim \mathcal{N}(0,\lambda)$ есть неподвижная точка преобразования m-го порядка, индуцированного $H_\lambda^m(x)$.

Из (37) видно, что $\alpha_\lambda^{(m)} = \lambda^m$. То же самое мы бы получили, если воспользовались (32) и

$$\mathbf{E}\,[e^{Z_\lambda t-\frac{1}{2}\lambda t^2}]^2 = e^{\lambda t^2} = \sum_{m\geqslant 0}\lambda^m\frac{t^{2m}}{m!}.$$

Далее, поскольку производящая функция (35) удовлетворяет условиям теоремы 3.2, распределение случайного индекса $\mathbf{I}$ из теоремы 3.1 является мультиномиальным $\mathrm{Mult}\,(m, \boldsymbol{\lambda})$. В случае нулевого смещения и $m = 1$ это мультиномиальное распределение сводится к распределению «выберите индекс, пропорциональный дисперсии», что отражено в (6).

В заключение покажем 2 направления (способа), в которых классическое уравнение Стейна может быть обобщено на эрмитов случай. Пусть $\mathcal{N}h = \mathbf{E}\,h(Z)$ есть стандартное нормальное ожидание h. Тогда уравнения

$$f'(x)H_1^{m-1}(x) - H_1^m(x)f(x) = h(x) - \mathcal{N}h \tag{38}$$

и

$$f^{(m)}(x) - H_1^m(x)f(x) = h(x) - \mathcal{N}h \tag{39}$$

сводятся к обычному уравнению Стейна для $m = 1$ (см. [21], [22]):

$$f'(x) - xf(x) = h(x) - \mathcal{N}h. \tag{40}$$

В частности, ожидания в левых частях, взятые для случайной величины W, равны нулю для любой $f \in C_c^\infty$ тогда и только тогда, когда W есть стандартная нормальная величина.

4.2. Многочлены Лагерра. Пусть $\lambda > 0$ и $\{L_\lambda^m(x)\}_{m \geqslant 0}$ есть совокупность многочленов Лагерра, определенная при помощи производящей функции

$$(1+t)^{-\lambda} \exp\left\{\frac{xt}{1+t}\right\} = \sum_{m=0}^{\infty} L_\lambda^m(x)\frac{t^m}{m!} \tag{41}$$

или, что эквивалентно, через формулу Родрига

$$L_\lambda^m(x) = (-1)^m x^{-\lambda+1} e^x \frac{d^m}{dx^m} x^{\lambda+m-1} e^{-x}. \tag{42}$$

Эти многочлены ортогональны гамма-распределению с параметром λ, имеющему плотность $x^{\lambda-1}e^{-x}/\Gamma(\lambda), x > 0$. Пусть $F \in \mathcal{C}_c^\infty$ и Z_λ имеет эту плотность. Применяя формулу Родрига (42), получаем

$$\begin{aligned}
\mathbf{E}\,L_\lambda^m(Z_\lambda)F(Z_\lambda) &= \int_0^\infty (-1)^m x^{-\lambda+1} e^x \left(\frac{d^m}{dx^m} x^{\lambda+m-1} e^{-x}\right) F(x) \frac{x^{\lambda-1}e^{-x}}{\Gamma(\lambda)}\,dx \\
&= \frac{(-1)^m}{\Gamma(\lambda)} \int_0^\infty \left(\frac{d^m}{dx^m} x^{\lambda+m-1} e^{-x}\right) F(x)\,dx \\
&= \frac{\Gamma(\lambda+m)}{\Gamma(\lambda)} \int_0^\infty F^{(m)}(x) \frac{x^{\lambda+m-1}e^{-x}}{\Gamma(\lambda+m)}\,dx \\
&= (\lambda)^m\,\mathbf{E}\,F^{(m)}(Z_{\lambda+m}), \qquad (43)
\end{aligned}$$

где $(\lambda)^m$ есть растущий факториал:

$$(\lambda)^m = \lambda(\lambda+1)\cdots(\lambda+m-1) = \frac{\Gamma(\lambda+m)}{\Gamma(\lambda)}.$$

Поэтому

$$(Z_\lambda)_\lambda^{(m)} = Z_{\lambda+m}.$$

Из (43) видно, что $\alpha_\lambda^{(m)} = (\lambda)^m$. То же самое можно получить, воспользовавшись (32) и

$$\mathbf{E}\left\{(1+t)^{-\lambda}\exp\left\{\frac{Z_\lambda t}{1+t}\right\}\right\}^2 = (1-t^2)^{-\lambda} = \sum_{m\geqslant 0}(\lambda)^m\frac{t^{2m}}{m!}.$$

Поскольку производящая функция (41) удовлетворяет условиям теоремы 3.2, случайный индекс $\mathbf{I}$ из теоремы 3.1 имеет распределение

$$\mathbf{P}\{\mathbf{I}=\mathbf{m}\} = \binom{m}{\mathbf{m}}\frac{\prod_{i=1}^n(\lambda_i)^{m_i}}{(\lambda)^m} = \frac{\prod_{i=1}^n\binom{\lambda_i+m_1-1}{m_i}}{\binom{\lambda+m-1}{m}}.$$

Это есть не что иное как многомерное гипергеометрическое распределение с параметрами m и $\lambda_1+m_1-1,\ldots,\lambda_n+m_n-1$ (см. [18], стр. 301).

Однако гамма-распределение не является неподвижной точкой для преобразований Лагерра, как нормальное для эрмитовых. Тем не менее, существует уравнение Стейна для гамма-распределения, соответсвующее (40) для нормальной случайной величины, которое может быть использовано для изучения аппроксимации гамма-семейства; более подробно см. [20]. В частности, получаем характеристику Стейна: $X\sim\Gamma(\lambda,1)$ тогда и только тогда, когда

$$\mathbf{E}\,(X-\lambda)f(X) = \mathbf{E}\,Xf'(X)$$

для любой гладкой функции f. Используя соотношение $L_\lambda^1(x) = x-\lambda$, получаем, что преобразование Лагерра первого порядка $X^{(1)}$ удовлетворяет

$$\mathbf{E}\,(X-\lambda)f(X) = \lambda\,\mathbf{E}\,f'(X^{(1)})$$

для любой гладкой функции f. Сравнивая эти два уравнения, видим, что $X\sim\Gamma(\lambda,1)$ тогда и только тогда, когда для любой гладкой функции f выполняется

$$\mathbf{E}\,Xf'(X) = \lambda Ef'(X^{(1)}).$$

Другими словами, $X\sim\Gamma(\lambda,1)$ тогда и только тогда, когда $X^{(1)}$, преобразование Лагерра первого порядка случайной величина X, равняется своему преобразованию объемного смещения X^s.

4.3. Многочлены Шарлье*. Пусть $\lambda > 0$ и $(x)_k = x(x-1)\cdots(x-k+1)$ есть убывающий факториал. Определим семейство многочленов Шарлье $\{C_\lambda^m(x)\}_{m\geqslant 0}$ через производящую функцию

$$e^{-\lambda t}(1+t)^x = \sum_{m=0}^{\infty} C_\lambda^m(x)\frac{t^m}{m!} \tag{44}$$

или, что эквивалентно, через

$$C_\lambda^m(x) = \sum_{k=0}^{m} \binom{m}{k}(x)_k(-\lambda)^{m-k}. \tag{45}$$

Это семейство ортогонально распределению Пуассона $\mathcal{P}(\lambda)$ с вероятностями $e^{-\lambda}\lambda^k/k!, k = 0, 1, \ldots$. Из (45) выводится формула Родрига

$$C_\lambda^m(x) = (-1)^m\Gamma(x+1)\lambda^{m-x}\nabla^m\left(\frac{\lambda^x}{\Gamma(x+1)}\right), \tag{46}$$

где $\nabla f(x) = f(x) - f(x-1)$ есть левосторонняя разность.

Поскольку при помощи преобразований из 2.1, определенных через производные тестовых функций, получаются абсолютно непрерывные распределения (когда $m \geqslant 1$), никакое дискретное распределение не будет неподвижной точкой. Однако, в соответствии с (9), для целочисленной случайной величины X можно определить дискретное смещенное распределение $X - P$ посредством

$$\mathbf{E}\,P(X)F(X) = \alpha\,\mathbf{E}\,\Delta^m F(X^{(m)}) \quad \text{для любой} \quad F \in \mathcal{F}_\Delta(P), \tag{47}$$

где $\Delta f(x) = f(x+1) - f(x)$ и

$$\mathcal{F}_\Delta(P) = \{F\colon \mathbf{R} \to \mathbf{R}\colon \mathbf{E}\,|P(X)F(X)| < \infty\}.$$

Так для любого $m = 0, 1\ldots$ распределение Пуассона (λ) является неподвижной точкой

$$(Z_\lambda)_\lambda^{(m)} = Z_\lambda$$

для дискретного преобразования (47), в котором P заменено на C_λ^m. Так как $Z_\lambda \sim \mathcal{P}(\lambda)$, то по формуле Родрига (46) и

$$\sum_{k=0}^{\infty}\nabla^m b_k a_k = (-1)^m\sum_{k=0}^{\infty} b_k\Delta^m a_k \tag{48}$$

*В отечественной литературе эти многочлены называют многочленами Пуассона–Шарлье. — Прим. ред.

имеем

$$\begin{aligned}
\mathbf{E}\, C_\lambda^m(Z_\lambda)F(Z_\lambda) &= \sum_{k=0}^{\infty} e^{-\lambda}\frac{\lambda^k}{k!}C_\lambda^m(k)F(k) \\
&= \lambda^m(-1)^m\sum_{k=0}^{\infty} e^{-\lambda}\nabla^m\left(\frac{\lambda^k}{k!}\right)F(k) \\
&= \lambda^m\sum_{k=0}^{\infty}\frac{e^{-\lambda}\lambda^k}{k!}\Delta^m F(k) \\
&= \lambda^m\mathbf{E}\,\Delta^m F(Z_\lambda). \qquad (49)
\end{aligned}$$

Из (49) видно, что $\alpha_\lambda^{(m)} = \lambda^m$. То же самое можно получить другим способом, воспользовавшись (32) и

$$\mathbf{E}\,[e^{-\lambda t}\,(1+t)^{Z_\lambda}]^2 = e^{\lambda t^2} = \sum_{m\geqslant 0}\lambda^m\frac{t^{2m}}{m!}.$$

Существование смещенных распределений Шарлье, а также тот факт, что (29) остается справедливым при замене производных на оператор разности, позволяют легко установить выполнение условия, а значит, и справедливость утверждения теоремы 3.1 в этом дискретном случае. Далее, поскольку производящая функция (44) удовлетворяет условиям теоремы 3.2, распределение случайного индекса $\mathbf{I}$ в теореме 3.1, как и в нормальном случае, является мультиномиальным $\mathrm{Mult}\,(m,\boldsymbol{\lambda})$.

Так как многочлен Шарлье первой степени есть $C_\lambda^1(x) = x-\lambda$, то характеристики Стейна формы (38) или (39) (где многочлены Эрмита заменены на многочлены Шарлье, и производные заменены на разности) обобщают уравнение Стейна для распределения Пуассона с параметром λ, которое приведено в [7] и широко изучено, например, в [4].

4.4. Многочлены Кравчука. Пусть $\lambda = 1,2,\ldots$ и $p\in(0,1)$ фиксированны. Определим семейство многочленов Кравчука $\{K_\lambda^m(x)\}_{0\leqslant m\leqslant\lambda}$ через производящую функцию

$$(1+qt)^x\,(1-pt)^{\lambda-x} = \sum_{m=0}^{\lambda}\frac{t^m}{m!}K_\lambda^m(x), \qquad (50)$$

где $p+q=1$. Получаем семейство многочленов, ортогональное биномиальному распределению $\mathcal{B}(\lambda,p)$. В противоположность предыдущим примерам биномиальное распределение не является безгранично делимым распределением, и носителем его является ограниченное множество.

Следуя методу, изложенному в [3], эти многочлены можно также получить при помощи формулы Родрига

$$K_\lambda^m(x) = \frac{(-1)^m m!\binom{\lambda}{m} p^{m-x} q^x}{\binom{\lambda}{x}} \nabla^m \left\{ \binom{\lambda - m}{x} \left(\frac{p}{q}\right)^x \right\}.$$

И поэтому для $Z_\lambda \sim \mathcal{B}(\lambda, p)$, $0 \leqslant m \leqslant \lambda$, и ограниченной F имеем

$$\begin{aligned}
&\mathbf{E}\, K_\lambda^m(Z_\lambda) F(Z_\lambda) \\
&\quad = \mathbf{E}\, F(Z_\lambda) \frac{(-1)^m m!\binom{\lambda}{m} p^{m-Z_\lambda} q^{Z_\lambda}}{\binom{\lambda}{Z_\lambda}} \nabla^m \left\{ \binom{\lambda - m}{Z_\lambda} \left(\frac{p}{q}\right)^{Z_\lambda} \right\} \\
&\quad = \sum_{k=0}^{\lambda} \binom{\lambda}{k} p^k q^{\lambda-k} F(k) \frac{(-1)^m m!\binom{\lambda}{m} p^{m-k} q^k}{\binom{\lambda}{k}} \nabla^m \left\{ \binom{\lambda - m}{k} \left(\frac{p}{q}\right)^k \right\} \\
&\quad = m!\binom{\lambda}{m} p^m q^\lambda (-1)^m \sum_{k=0}^{\lambda} F(k) \nabla^m \left\{ \binom{\lambda - m}{k} \left(\frac{p}{q}\right)^k \right\}.
\end{aligned}$$

Пусть, как и прежде, $(\lambda)_m$ обозначает убывающий факториал. Используя (48), запишем последнее выражение в виде

$$\begin{aligned}
(\lambda)_m p^m q^\lambda &\sum_{k=0}^{\lambda} \binom{\lambda - m}{k} \left(\frac{p}{q}\right)^k \Delta^m F(k) \\
&= (\lambda)_m (pq)^m \sum_{k=0}^{\lambda} \binom{\lambda - m}{k} p^k q^{\lambda - m - k} \Delta^m F(k) \\
&= \alpha_\lambda^{(m)}\, \mathbf{E}\, \Delta^m F(Z_\lambda^{(m)}),
\end{aligned}$$

что дает

$$\alpha_\lambda^{(m)} = (\lambda)_m (pq)^m \quad \text{и} \quad (Z_\lambda)_\lambda^{(m)} = Z_{\lambda - m}.$$

Следовательно, аналогично гамма-семейству, биномиальное распределение не является неподвижной точкой соответствующего распределения, но распределение, полученное в результате преобразования, принадлежит тому же самому семейству. Можно вычислить $\alpha_\lambda^{(m)}$ и другим способом, если воспользоваться (32), (50) и разложением в ряд

$$\mathbf{E}\, (1 + qt)^{2Z_\lambda} (1 - pt)^{2\lambda - 2Z_\lambda} = (1 + pqt^2)^\lambda.$$

Что касается 4.3, утверждение теоремы 3.1 выполняется, и так как производящая функция (50) удовлетворяет условиям теоремы 3.2, то распределение случайного индекса $\mathbf{I}$ в теореме 3.1 задается формулой

$$\mathbf{P}\{\mathbf{I} = \mathbf{m}\} = \frac{\binom{m}{\mathbf{m}} (pq)^{\sum m_i} \prod_{i=1}^n (\lambda_i)_{m_i}}{(\lambda)_m (pq)^m} = \binom{m}{\mathbf{m}} \frac{\prod_{i=1}^n (\lambda_i)_{m_i}}{(\lambda)_m} = \frac{\prod_{i=1}^n \binom{\lambda_i}{m_i}}{\binom{\lambda}{m}},$$

что есть не что иное как многомерное гипергеометрическое распределение с параметрами m и $\lambda_1, \ldots, \lambda_n$ (см. [18], стр. 301).

Из [10] получаем характеристику Стейна: $X \sim \mathcal{B}(\lambda, p)$ тогда и только тогда, когда

$$p\,\mathbf{E}\,(\lambda - X)f(X+1) = q\,\mathbf{E}\,Xf(X),$$

для любых функций f, для которых существуют ожидания. Используя многочлен Кравчука первой степени $K_\lambda^1(x) = qx - p(\lambda - x)$, получаем преобразование Кравчука, которое удовлетворяет выражению

$$q\,\mathbf{E}\,Xf(X) - p\,\mathbf{E}\,(\lambda - X)f(X) = \lambda pq\,\mathbf{E}\,\Delta f(X^{(1)}).$$

Комбинируя эти уравнения, получаем, что $X \sim \mathcal{B}(\lambda, p)$ тогда и только тогда, когда

$$\begin{aligned} p\,\mathbf{E}\,(\lambda - X)\Delta f(X) &= p\mathbf{E}\,(\lambda - X)(f(X+1) - f(X)) \\ &= q\,\mathbf{E}\,Xf(X) - p\,\mathbf{E}\,(\lambda - X)f(X) \\ &= \lambda pq\,\mathbf{E}\,\Delta f(X^{(1)}). \end{aligned}$$

Полагая $g(x) = \Delta f(\lambda - x)$, видим, что $X \sim \mathcal{B}(\lambda, p)$ тогда и только тогда, когда $\lambda - X^{(1)}$ имеет распределение $(\lambda - X)$-объемного смещения, а именно тогда и только тогда, когда

$$\lambda - X^{(1)} \sim \mathcal{B}(\lambda - 1, q) + 1, \quad \text{что эквивалентно} \quad X^{(1)} \sim \mathcal{B}(\lambda - 1, p).$$

4.5. Многочлены Гегенбауэра. В этом последнем разделе мы рассмотрим систему многочленов, ортогональную непрерывному распределению с компактным носителем. Для $\lambda > -1/2$ положим (см. [3])

$$\alpha_\lambda^{(m)} = \frac{\Gamma(\lambda)\Gamma(2\lambda + m)\Gamma(\lambda + 1)}{2^{2m}\Gamma(\lambda + m + 1)\Gamma(\lambda + m)\Gamma(2\lambda)}$$

и определим семейство многочленов Гегенбауэра $G_\lambda^m(x)$ через формулу Родрига:

$$G_\lambda^m(x) = (-1)^m \alpha_\lambda^{(m)} \frac{\Gamma(\lambda + \frac{1}{2})\Gamma(\lambda + m + 1)}{\Gamma(\lambda + 1)\Gamma(\lambda + m + \frac{1}{2})} (1 - x^2)^{\frac{1}{2} - \lambda} \frac{d^m}{dx^m} \{(1 - x^2)^{\lambda + m - \frac{1}{2}}\}.$$

Тогда $G_\lambda^m(x)$ нормированны, имеют степень m и удовлетворяют условию ортогональности

$$\frac{1}{m!}\,\mathbf{E}\,G_\lambda^k(Z_\lambda)G_\lambda^m(Z_\lambda) = \alpha_\lambda^{(m)}\delta_{k,m}, \qquad k = 0, \ldots, m,$$

где $Z_\lambda \sim g_\lambda$,

$$g_\lambda(x) = \frac{1}{\sqrt{\pi}} \frac{\Gamma(\lambda + 1)}{\Gamma(\lambda + \frac{1}{2})} (1 - x^2)^{\lambda - \frac{1}{2}}, \quad |x| \leqslant 1.$$

В частности, получаем следствие 3.1, что доказывает существование семейства преобразований Гегенбауэра. Это семейство распределений является специальным случаем I-распределений центрированного пирсоновского типа, которое иногда также называется I-распределениями типа бета-распределений (см. [23], стр. 150). Заметим, что для $\lambda = 1/2$ получаем равномерное распределение $\mathcal{U}[-1,1]$, для $\lambda = 0$ — распределение арксинуса, и для $\lambda = 1$ — полукруговое распределение [25], [26].

Пусть $F \in \mathcal{C}_c^\infty$. Рассмотрим результат G_λ^m-преобразования на $Z_\lambda \sim g_\lambda$. Имеем

$$\begin{aligned}
&\mathbf{E}\, G_\lambda^m(Z_\lambda)F(Z_\lambda)\\
&\quad= \frac{1}{\sqrt{\pi}}\frac{\Gamma(\lambda+1)}{\Gamma(\lambda+\frac12)}(-1)^m\alpha_\lambda^{(m)}\frac{\Gamma(\lambda+\frac12)\Gamma(\lambda+m+1)}{\Gamma(\lambda+1)\Gamma(\lambda+m+\frac12)}\\
&\quad\quad\times\int_{-1}^1 F(x)\frac{d^m}{dx^m}\{(1-x^2)^{\lambda+m-\frac12}\}\\
&\quad= \frac{1}{\sqrt{\pi}}\alpha_\lambda^{(m)}\frac{\Gamma(\lambda+m+1)}{\Gamma(\lambda+m+\frac12)}\int_{-1}^1 F^{(m)}(x)(1-x^2)^{\lambda+m-\frac12}\\
&\quad= \alpha_\lambda^{(m)}\,\mathbf{E}\,F^{(m)}(Z_{\lambda+m}),
\end{aligned}$$

что дает

$$(Z_\lambda)_\lambda^{(m)} = Z_{\lambda+m}.$$

Поэтому в случае $\lambda = 0$ получаем, что преобразование Гегенбауэра первого порядка распределения арксинуса является полукруговым законом.

В заключение заметим, что вышеупомянутые распределения типа бета-распределений не являются инвариантными относительно сложения, поэтому теорема 3.1 и ее конструкция не применимы. Однако так как $G_\lambda^1(x) = x$, то преобразование Гегенбауэра первого порядка есть не что иное как преобразование объемного смещения. Поэтому для сумм независимых случайных величин можно применить конструкцию, приведенную в (7).

Слова благодарности.[3] Авторы хотели бы поблагодарить организаторов программы «Метод Стейна и его применения: программа в честь Чарльза Стейна» и Институт Математических Наук в Сингапуре за их щедрое гостеприимство и превосходную встречу, где и была закончена данная работа. Также мы хотели бы поблагодарить рецензента за очень полезные комментарии.

[3]Эта работа была частично выполнена во время визита авторов в 2003 году в Национальный Университет Сингапура и Институт Математических Наук, который и организовал данный визит.

СПИСОК ЛИТЕРАТУРЫ

1. *Abramowitz*, *Stegun A. A., eds.* Handbook of Mathematical Functions. New York: M Dover, 1972.
2. *Asai N.*, *Kubo I.*, *Kuo H.* Multiplicative renormalization and generating functions I. — Taiwanese Journal of Mathematics, 2003, v. 7, № 1, p. 89–101.
3. *Asai N.*, *Kubo I.*, *Kuo H.* Multiplicative renormalization and generating functions II. — Taiwanese Journal of Mathematics, 2004, v. 8, № 4, p. 593–628.
4. *Barbour A.D.*, *Holst L.*, *Janson S.* Poisson Approximation. Oxford University Press, 1992.
5. *Bolthausen E.* An estimate of the remainder in a combinatorial central limit theorem. — Z. Wahrscheinlichkeitstheor. Verw. Geb., 1984, v. 66, p. 379–386.
6. *Cacoullos T.*, *Papathanasiou V.*, *Utev S.* Variational inequalities with examples and an application to the central limit theorem. — Ann. Probab., 1994, v. 22, p. 1607–1618.
7. *Chen L. H. Y.* Poisson approximation for dependent trials. — Ann. Probab., 1975, v. 3, p. 534–545.
8. *Chen L. H. Y.*, *Shao Q-M.* Stein's method and normal approximation. — Tutorial notes for the workshop on Stein's method and Applications. http://www.ims.nus.edu.sg/Programs/stein/files/tut_shao.
9. *Diaconis P.*, *Zabell S.* Closed form summation for classical distributions: variations on a theme of de Moivre. — Statist. Sci., 1991, v. 6, p. 284–302.
10. *Ehm W.* Binomial approximation to the Poisson binomial distribution. — Statist. Probab. Lett., 1991, v. 11, p. 7–16.
11. *Goldstein L.* Normal approximation for hierarchical sequences. — Ann. Appl. Probab., 2003. (In Press.)
12. *Goldstein L.* Berry Esseen bounds for combinatorial central limit theorems and pattern occurrences, using zero and size biasing. Preprint, 2004.
13. *Goldstein L.*, *Reinert G.* Stein's method and the zero bias transformation with application to simple random sampling. — Ann. Appl. Probab., 1997, v. 7, p. 935–952.
14. *Goldstein L.*, *Reinert G.* Zero biasing in one and higher dimensions, and applications. Preprint, 2004.
15. *Goldstein L.*, *Rinott Y.* On multivariate normal approximations by Stein's method and size bias couplings. — J. Appl. Prob., 1996, v. 33, p. 1–17.
16. *Ho S.-T., and Chen L. H. Y.* An L_p bound for the remainder in a combinatorial central limit theorem. — Ann. Probab., 1978, v. 6, p. 231–249.
17. *Holmes S.* Stein's Method for Birth and Death Processes. Preprint, 1998.
18. *Johnson N. L.*, *Kotz S.* Discrete Distributions. New York etc.: Wiley, 1969.
19. *Karlin S.*, *Studden W. J.* Tchebycheff Systems, with Applications in Analysis and Statistics. New York: Interscience, 1966.
20. *Luk H. M.* Stein's method for the gamma distribution and related statistical applications. — Ph.D. thesis. Los Angeles: University of Southern California, 1994.
21. *Stein C.* A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. — Proc. Sixth Berkeley Symp. Math. Statist. Probab., Berkeley: Univ. California Press., 1972, v. 2, p. 583–602.
22. *Stein C.* Approximate Computation of Expectations. California: IMS, Hayward, 1986.
23. *Kendall M. G.*, *Stuart S.* The Advanced Theory of Statistics. V. 1: Distribution Theory. London: Griffin, 1969.
24. *Schoutens W.* Orthogonal polynomials in Stein's method. — J. Math. Anal. Appl., 2001, v. 253, p. 515–531.

25. *Wigner E. P.* Characteristic vectors of bordered matrices with infinite dimension. — Ann. Math., 1955, v. 62, p. 548–564.
26. *Wigner E. P.* On the distribution of the roots of certain symmetric matrices. — Ann. Math., 1958, v. 67, p. 325–328.