Московский государственный университет им. М. В. Ломоносова
Факультет вычислительной математики и кибернетики

В. Ю. Королев

ВЕРОЯТНОСТНЫЕ МОДЕЛИ:

ВВЕДЕНИЕ В АСИМПТОТИЧЕСКУЮ ТЕОРИЮ СЛУЧАЙНОГО СУММИРОВАНИЯ

Диалог МГУ
Москва
1997

УДК 519.2
ББК 22.17
К68

Рецензенты:
кафедра исследования операций Тверского государственного университета, доктор физико-математических наук профессор В. М. Золотарев,
доктор физико-математических наук профессор В. В. Калашников.

В.Ю.Королев

К68 **Вероятностные модели: введение в асимптотическую теорию случайного суммирования.** -
М.: Диалог-МГУ, 1997. - 72 с.
ISBN 5-89209-166-X

Данное учебное пособие содержит материал, являющийся основой курса "Вероятностные модели", читаемого автором в течение нескольких последних лет студентам третьего курса факультета вычислительной математики и кибернетики МГУ. На примере методов асимптотической теории случайного суммирования показывается, как, исходя из конкретных практических задач, следует модифицировать классические предельные теоремы теории вероятностей, чтобы получить более адекватные модели реальных процессов. Наряду с изложением общих теоретических результатов описываются некоторые модели из страховой и финансовой математики. Обсуждаются информационные принципы построения вероятностных моделей.

УДК 519.2
ББК 22.17

Работа над этим учебным пособием поддерживалась Российским фондом фундаментальных исследований, проекты 96-01-01919 и 97-01-00271.

ISBN 5-89209-166-X

Введение

Данное пособие содержит материал, являющийся основой курса "вероятностные модели", читаемого автором в течение нескольких последних лет студентам третьего курса факультета вычислительной математики и кибернетики МГУ.

Этот курс посвящен применению предельных теорем теории вероятностей к решению прикладных задач. Предельные теоремы традиционно являются ядром теории вероятностей. Как отмечено в знаменитой книге (Гнеденко и Колмогоров, 1949), познавательная ценность теории вероятностей раскрывается только предельными теоремами. Однако зачастую практика выдвигает новые задачи, которые не могут найти адекватного решения или математического описания в рамках классических предельных теорем. Это, в свою очередь, требует модификации самих постановок задач в предельных теоремах, чтобы сделать их более гибкими.

Замечательным примером такого взамодействия между теорией и практикой является асимптотическая теория суммирования независимых случайных величин. В последнее время появилось много примеров из теории надежности, математической теории страхования или финансовой математики, когда классические предельные теоремы для сумм независимых случайных величин, такие как закон больших чисел или центральная предельная теорема, при их применении к анализу реальных ситуаций приводят к существенному расхождению получаемых математических моделей с экспериментальными данными. Это обстоятельство привело к расширению схемы, традиционно изучаемой в теории предельных теорем для сумм независимых случайных величин, а именно, объектом интенсивного изучения стали так назваемые случайные суммы, в которых число слагаемых само является случайной величиной. Именно предельные теоремы для случайных сумм и их применение к решению некоторых практических задач и изучаются в данном курсе лекций.

Из-за относительно небольшого объема (курс рассчитан на один семестр) мы ограничимся простейшей предельной схемой и будем рассматривать только "нарастающие" суммы, затронув схему серий лишь в заключительном разделе.

Предполагается, что читатель знаком с теорией вероятностей в рамках университетского курса для математических и физических специальностей.

Мы будем использовать традиционные обозначения. В качестве основных вероятностных моделей мы будем рассмативать распределения случайных величин. Поэтому мы сосредоточимся на изучении слабой сходимости случайных величин и их распределений. Напомним некоторые определения и утверждения из теории вероятностей, которые в дальнейшем будут интенсивно использоваться.

Будем считать, что все рассматриваемые случайные величины и векторы заданы на одном вероятностном пространстве $(\Omega, \mathcal{A}, \mathsf{P})$.

Последовательность функций распределения $F_1, F_2, \ldots$ слабо сходится к функции распределения F (что обозначается $F_n \Rightarrow F$) при $n \to \infty$, если $F_n(x) \to F(x)$ при

$n \to \infty$ в каждой точке x, в которой предельная функция распределения непрерывна. Если $X, X_1, X_2, \ldots$ – случайные величины с функциями распределения $F, F_1, F_2, \ldots$ соответственно, то будем говорить, что последовательность $\{X_n\}$ слабо сходится к X (обозначая это $X_n \Rightarrow X$), если $F_n \Rightarrow F$. Как известно, $X_n \Rightarrow X$, если и только если $\mathsf{E}\varphi(X_n) \to \mathsf{E}\varphi(X)$ для любой непрерывной и ограниченной функции φ. Слабая сходимость случайных величин эквивалентна поточечной сходимости их характеристических функций.

Семейство случайных величин $\{X_n\}$ называется слабо компактным, если из любой последовательности его элементов можно выделить слабо сходящуюся подпоследовательность. Как известно, семейство случайных величин $\{X_n\}$ слабо компактно тогда и только тогда, когда

$$\lim_{R\to\infty} \sup_n \mathsf{P}(|X_n| > R) = 0$$

(см., например, (Гнеденко и Колмогоров, 1949), с. 43).

Расстоянием (метрикой) Леви между функциями распределения F_1 и F_2 называется величина

$$L(F_1, F_2) = \inf\{h > 0 : \ F_1(x-h) - h \le F_2(x) \le F_1(x+h) + h \text{ для всех } x\}.$$

Расстояние Леви между функциями распределения F_1 и F_2 равно длине стороны наибольшего квадрата, который можно вписать между их графиками. Если X_1 и X_2 – случайные величины с функциями распределения F_1 и F_2 соответственно, то будем считать, что $L(X_1, X_2) = L(F_1, F_2)$. Сходимость в смысле расстояния Леви эквивалентна слабой сходимости (см., например, (Гнеденко и Колмогоров, 1949), с. 38).

Рассмотрим многомерное обобщение метрики Леви. Пусть $(E, \mathcal{E}, \rho)$ – метрическое пространство и $\mathcal{P}(E)$ – семейство вероятностных мер на измеримом пространстве $(E, \mathcal{E})$. Пусть $A \subseteq E$. Положим $\rho(x, A) = \inf\{\rho(x,y) : \ y \in A\}$. Пусть $\varepsilon > 0$. Обозначим $A^\varepsilon = \{x \in E : \ \rho(x, A) < \varepsilon\}$, $\ A \in \mathcal{E}$. Пусть P_1 и P_2 – произвольные вероятностные меры из $\mathcal{P}(E)$. Положим

$$\sigma(\mathsf{P}_1, \mathsf{P}_2) = \inf\{\varepsilon > 0 : \ \mathsf{P}_1(A) \le \mathsf{P}_2(A^\varepsilon) + \varepsilon \text{ для любого замкнутого } A \in \mathcal{E}\}.$$

Расстоянием Леви-Прохорова между распределениями P_1 и P_2 называется величина

$$\pi(\mathsf{P}_1, \mathsf{P}_2) = \max\{\sigma(\mathsf{P}_1, \mathsf{P}_2), \sigma(\mathsf{P}_2, \mathsf{P}_1)\}.$$

Расстоянием Леви-Прохорова между случайными векторами X и Y будем считать расстояние Леви-Прохорова между индуцированными ими распределениями: $\pi(X, Y) = \pi(\mathsf{P}_X, \mathsf{P}_Y)$. Последовательность случайных векторов $X_1, X_2, \ldots$ слабо сходится к случайному вектору X, если $\mathsf{E}\varphi(X_n) \to \mathsf{E}\varphi(X)$ для любой непрерывной ограниченной функции φ. Как известно, слабая сходимость случайных векторов эквивалентна их сходимости в смысле метрики Леви-Прохорова (см., например, (Ширяев, 1989), с. 375).

Совпадение распределений и сходимость по вероятности будут соответственно обозначаться символами $\stackrel{d}{=}$ и $\stackrel{P}{\to}$.

1 Приближение распределений случайных сумм специальными смесями

Пусть $X_1, X_2, \ldots$ – независимые случайные величины. Обозначим $S_k = X_1 + \ldots + X_k$, $k \ge 1$. Пусть $\{a_k\}_{k\ge1}$ и $\{b_k\}_{k\ge1}$ – числовые последовательности, $b_k > 0$ $(k \ge 1)$. Обозначим

$$Y_k = \frac{S_k - a_k}{b_k}.$$

Характеристическую функцию случайной величины Y_k обозначим $h_k(t)$, $t \in \mathbb{R}$, $h_k(t) = \mathsf{E}e^{itY_k}$. Пусть $\{N_k\}_{k\ge1}$ – целочисленные положительные случайные величины, независимые от последовательности $\{X_j\}_{j\ge1}$ при каждом $k \ge 1$. Наша цель – изучить асимптотическое поведение случайных величин

$$Z_k = \frac{S_{N_k} - c_k}{d_k},$$

где $\{c_k\}_{k\ge1}$ и $\{d_k\}_{k\ge1}$ – числовые последовательности, $d_k > 0$, $k \ge 1$. Характеристическую функцию случайной величины Z_k обозначим $f_k(t)$, $f_k(t) = \mathsf{E}e^{itZ_k}$. Вообще говоря, последовательности $\{a_k\}, \{b_k\}$ и $\{c_k\}, \{d_k\}$, гарантирующие слабую сходимость или слабую компактность случайных величин $\{Y_k\}$ и $\{Z_k\}$, могут быть разными.

В этом разделе мы будем предполагать, что последовательности $\{a_k\}$ и $\{b_k\}$ обеспечивают слабую сходимость центрированных и нормированных сум

$$Y_k \Rightarrow Y \qquad (k \to \infty) \tag{1.1}$$

к некоторой случайной величине Y, характеристическую функцию которой мы обозначим $h(t)$, $t \in \mathbb{R}$, $h(t) = \mathsf{E}e^{itY}$. Положим

$$g_k(t) = \sum_{n=1}^{\infty} \mathsf{P}(N_k = n) \exp\left\{it\left(\frac{a_n - c_k}{d_k}\right)\right\} h\left(t\frac{b_n}{d_k}\right).$$

Лемма 1.1. *Пусть $b_k \to \infty$, $d_k \to \infty$ $(k \to \infty)$. Предположим, что имеет место сходимость (1.1) и семейство случайных величин $\{b_{N_k}/d_k\}_{k\ge1}$ слабо компактно. Тогда для любого $t \in \mathbb{R}$*

$$\lim_{k\to\infty} |f_k(t) - g_k(t)| = 0.$$

Доказательство. Несложно видеть, что по формуле полной вероятности

$$f_k(t) = \sum_{n=1}^{\infty} \mathsf{P}(N_k = n) \exp\left\{it\left(\frac{a_n - c_k}{d_k}\right)\right\} h_n\left(t\frac{b_n}{d_k}\right).$$

Пусть α и β – вещественные числа, $0 < \alpha < \beta < \infty$, выбором которых мы распорядимся позже. Введем множества

$$\mathcal{N}_1 = \mathcal{N}_1(k,\alpha) = \{n : b_n < \alpha d_k\},$$
$$\mathcal{N}_2 = \mathcal{N}_2(k,\alpha,\beta) = \{n : \alpha d_k \le b_n \le \beta d_k\},$$
$$\mathcal{N}_3 = \mathcal{N}_3(k,\beta) = \{n : b_n > \beta d_k\}.$$

Если $t = 0$, то утверждение выполняется тривиальным образом. Зафиксируем произвольное $t \ne 0$. Тогда

$$|f_k(t) - g_k(t)| \le$$
$$\sum_{n=1}^{\infty} \mathsf{P}(N_k = n)\left|\exp\left\{it\left(\frac{a_n - c_k}{d_k}\right)\right\}\left[h_n\left(t\frac{b_n}{d_k}\right) - h\left(t\frac{b_n}{d_k}\right)\right]\right| \le$$
$$\sum_{n=1}^{\infty} \mathsf{P}(N_k = n)\left|h_n\left(t\frac{b_n}{d_k}\right) - h\left(t\frac{b_n}{d_k}\right)\right| =$$
$$\sum_{n\in\mathcal{N}_1} \mathsf{P}(N_k = n)\left|h_n\left(t\frac{b_n}{d_k}\right) - h\left(t\frac{b_n}{d_k}\right)\right| +$$
$$\sum_{n\in\mathcal{N}_2} \mathsf{P}(N_k = n)\left|h_n\left(t\frac{b_n}{d_k}\right) - h\left(t\frac{b_n}{d_k}\right)\right| +$$
$$\sum_{n\in\mathcal{N}_3} \mathsf{P}(N_k = n)\left|h_n\left(t\frac{b_n}{d_k}\right) - h\left(t\frac{b_n}{d_k}\right)\right| \equiv$$
$$I_1(\alpha,k) + I_2(\alpha,\beta,k) + I_3(\beta,k).$$

Пусть $\varepsilon > 0$ произвольно. Рассмотрим $I_1(\alpha,k)$. Поскольку функция h, будучи характеристической, непрерывна в нуле, а t фиксировано, можно выбрать $\alpha_1 = \alpha_1(\varepsilon,t)$ так, чтобы

$$\sup_k \sup_{n\in\mathcal{N}_1(k,\alpha_1)} \left|1 - h\left(t\frac{b_n}{d_k}\right)\right| \le \sup_{|\tau|\le\alpha_1|t|} |1 - h(\tau)| < \varepsilon. \tag{1.2}$$

Так как имеет место (1.1), семейство случайных величин $\{Y_k\}$ слабо компактно. Поэтому по критерию слабой компактности (см. (Лоэв, 1962), с. 206) семейство характеристических функций $\{h_k\}_{k\ge1}$ равностепенно непрерывно в нуле, то есть для любого $\varepsilon > 0$ существует $\delta = \delta(\varepsilon) > 0$ такое, что

$$\sup_k \sup_{|\tau|\le\delta} |1 - h_k(\tau)| < \varepsilon. \tag{1.3}$$

Для такого δ выберем $\alpha_2 = \alpha_2(\varepsilon,t)$ так, чтобы

$$\alpha_2 = \alpha_2(\varepsilon,t) \le \frac{\delta}{|t|}.$$

Тогда из (1.3) следует, что

$$\sup_k \sup_{n\in\mathcal{N}_1(k,\alpha_2)} \left|1 - h_n\left(t\frac{b_n}{d_k}\right)\right| \le \sup_n \sup_{|\tau|\le\delta} |1 - h_n(t)| < \varepsilon. \tag{1.4}$$

Положим $\alpha = \min(\alpha_1,\alpha_2)$. Тогда из соотношений (1.2) и (1.4) мы получим

$$I_1(\alpha,k) \le \sup_k \sup_{n\in\mathcal{N}_1(k,\alpha)} \left|1 - h\left(t\frac{b_n}{d_k}\right)\right| + \sup_k \sup_{n\in\mathcal{N}_1(k,\alpha)} \left|1 - h_n\left(t\frac{b_n}{d_k}\right)\right| < 2\varepsilon. \tag{1.5}$$

Рассмотрим $I_3(\beta,k)$. Поскольку по предположению семейство $\{b_{N_k}/d_k\}_{k\ge1}$ слабо компактно, для любого $\varepsilon > 0$ можно указать такое $\beta = \beta(\varepsilon)$, которое обеспечивает справедливость неравенства

$$I_3(\beta,k) \le 2\sum_{n\in\mathcal{N}_3(k,\beta)} \mathsf{P}(N_k = n) = 2\mathsf{P}\left(\frac{b_{N_k}}{d_k} > \beta\right) \le 2\sup_k \mathsf{P}\left(\frac{b_{N_k}}{d_k} > \beta\right) < \varepsilon. \tag{1.6}$$

Наконец, рассмотрим $I_2(k,\alpha,\beta)$ с α и β, определенными выше. Имеем

$$I_2(k,\alpha,\beta) \le \sum_{n\in\mathcal{N}_2} \mathsf{P}(N_k = n) \sup_{|\tau|\le\beta|t|} |h_n(\tau) - h(\tau)| \le$$
$$\sup_{n\in\mathcal{N}_2(k,\alpha,\beta)} \sup_{|\tau|\le\beta|t|} |h_n(t) - h(t)|. \tag{1.7}$$

Условие (1.1) влечет равномерную на каждом конечном интервале сходимость характеристических функций h_k к h при $k \to \infty$. Так как $b_k \to \infty$ и $a_k \to \infty$ $(k \to \infty)$, мы имеем

$$\inf \mathcal{N}_2(k,\alpha,\beta) \to \infty \qquad (k \to \infty).$$

Это означает, что при $k \to \infty$ правая часть соотношения (1.7) стремится к нулю, то есть для любого ε можно указать такое $k_0 = k_0(\varepsilon)$, что для всех $k \ge k_0$

$$I_2(k,\alpha,\beta) < \varepsilon. \tag{1.8}$$

Объединяя (1.5), (1.6) и (1.8), мы получаем, что при $k \ge k_0$

$$|f_k(t) - g_k(t)| < 4\varepsilon.$$

Произвольность ε в этом неравенстве доказывает лемму.

Лемма 1.1 открывает путь поисков приближений для распределений случайных сумм, если распределения индексов, то есть количеств слагаемых в суммах известны. Эта лемма описывает сближение распределений случайных сумм с аппроксимирующими законами, определяемыми характеристическими функциями $g_k(t)$.

Как мы видим, распределение, соответствующее характеристической функции $g_k(t)$, имеет довольно сложный вид и представляет собой смесь, вообще говоря, бесконечно большого числа распределений и потому малопригодно для рпактических применений. Однако, в некоторых частных случаях вид аппроксимирующего распределения существенно упрощается. Один из таких случаев связан с асимптотически неслучайными индексами, то есть такими, что

$$\frac{b_{N_k}}{d_k} \Rightarrow const. \tag{1.9}$$

Подобная ситуация имеет место во многих прикладных задачах, например, связанных с обобщенными пуассоновскими процессами, которые мы будем рассматривать позже, изучая асимптотическое поведение классического процесса риска. Сейчас же мы докажем два утверждения, вытекающих из Леммы 1.1. В каждом из них будет предполагаться сходимость (1.9).

Обозначим

$$w_k(t) = \mathsf{E}\exp\left\{it\left(\frac{a_{N_k} - c_k}{d_k}\right)\right\}, \quad r_k(t) = w_k(t) h\left(t\frac{b_k}{d_k}\right), \quad t \in \mathbb{R}.$$

Теорема 1.1. *Предположим, что $b_k \to \infty, d_k \to \infty$ $(k \to \infty)$ и имеет место (1.1). Пусть*

$$\frac{b_{N_k} - b_k}{d_k} \Rightarrow 0 \qquad (k \to \infty). \tag{1.10}$$

I. *Если*

$$\lim_{k\to\infty} \frac{b_k}{d_k} = 0, \tag{1.11}$$

то при любом $t \in \mathbb{R}$

$$\lim_{k\to\infty} |f_k(t) - w_k(t)| = 0. \tag{1.12}$$

II. *Если*

$$0 < c_0 \equiv \inf_k \frac{b_k}{d_k} \le \sup_k \frac{b_k}{d_k} \equiv c_1 < \infty, \tag{1.13}$$

то при любом $t \in \mathbb{R}$

$$\lim_{k\to\infty} |f_k(t) - r_k(t)| = 0. \tag{1.14}$$

Доказательство. Условие (1.10) с учетом (1.11) или (1.13) гарантирует слабую компактность семейства случайных величин $\{b_{N_k}/d_k\}_{k\ge1}$. Таким образом, в силу Леммы 1.1 в случае **I** необходимо показать, что

$$\lim_{k\to\infty} |g_k(t) - w_k(t)| = 0, \tag{1.15}$$

а в случае **II** надо убедиться, что

$$\lim_{k\to\infty} |g_k(t) - r_k(t)| = 0. \tag{1.16}$$

Начнем с доказательства (1.15). По ходу доказательства Леммы 1.1 мы ввели множество $\mathcal{N}_1 = \mathcal{N}_1(k,\alpha) = \{n : b_n < \alpha d_k\}$, $\alpha > 0$. Если $t = 0$, то (1.15) имеет месо тривиальным образом. Зафиксируем произвольное ненулевое t. Тогда

$$|g_k(t) - w_k(t)| =$$

$$\left|\sum_{n=1}^{\infty} \mathsf{P}(N_k = n)\exp\left\{it\left(\frac{a_n - c_k}{d_k}\right)\right\}\left[h\left(t\frac{b_n}{d_k}\right) - 1\right]\right| \le$$

$$\sum_{n=1}^{\infty} \mathsf{P}(N_k = n)\left|h\left(t\frac{b_n}{d_k}\right) - 1\right| \le$$

$$\sum_{n\in\mathcal{N}_1} \mathsf{P}(N_k = n)\left|h\left(t\frac{b_n}{d_k}\right) - 1\right| + \mathsf{P}(N_k \notin \mathcal{N}_1). \tag{1.17}$$

Функция h, будучи характеристической, непрерывна в нуле. Поэтому для любого $\varepsilon > 0$ существует $\alpha = \alpha(t,\varepsilon)$ такое, что

$$\sup_k \sup_{n\in\mathcal{N}_1(k,\alpha)} \left|h\left(t\frac{b_n}{d_k}\right) - 1\right| \le \sup_{|\tau|<\alpha|t|} |h(\tau) - 1| < \varepsilon. \tag{1.18}$$

Из условий (1.10) и (1.11) вытекает, что

$$\frac{b_{N_k}}{d_k} \Rightarrow 0 \qquad (k \to \infty).$$

Поэтому для указанных α и ε можно указать такое k_0, что для всех $k \ge k_0$

$$\mathsf{P}(N_k \notin \mathcal{N}_1(\alpha, k)) = \mathsf{P}\left(\frac{b_{N_k}}{d_k} \ge \alpha\right) < \varepsilon. \tag{1.19}$$

Подставляя (1.18) и (1.19) в (1.17), мы получим

$$|g_k(t) - w_k(t)| < 2\varepsilon,$$

что в силу произвольности ε означает, что имеет место (1.15), а это, в свою очередь, в силу Леммы 1.1 влечет (1.12).

Обратимся к пункту **II**. Введем множество

$$\mathcal{N} = \mathcal{N}(k,\delta) = \{n : |b_n - b_k| \le \delta d_k\}, \quad \delta > 0.$$

Если $t = 0$, то (1.16) тривиально. Пусть $t \ne 0$. Тогда

$$|g_k(t) - r_k(t)| =$$

$$\left|\sum_{n=1}^{\infty} \mathsf{P}(N_k = n)\exp\left\{it\left(\frac{a_n - c_k}{d_k}\right)\right\}\left[h\left(t\frac{b_n}{d_k}\right) - h\left(t\frac{b_k}{d_k}\right)\right]\right| \le$$

$$\sum_{n=1}^{\infty} \mathsf{P}(N_k = n)\left|h\left(t\frac{b_n}{d_k}\right) - h\left(t\frac{b_k}{d_k}\right)\right| \le$$

$$\sum_{n\in\mathcal{N}} \mathsf{P}(N_k = n)\left|h\left(t\frac{b_n}{d_k}\right) - h\left(t\frac{b_k}{d_k}\right)\right| + \mathsf{P}(N_k \notin \mathcal{N}). \tag{1.20}$$

Поскольку, будучи характеристической, функция h равномерно непрерывна, для любого $\varepsilon > 0$ существует $\delta = \delta(t,\varepsilon) \in (0, c_0)$ такое, что

$$\sup_k \sup_{n\in\mathcal{N}(k,\delta)} \left| h\left(t\frac{b_n}{d_k}\right) - h\left(t\frac{b_k}{d_k}\right)\right| \le$$
$$\sup\{|h(t_1) - h(t_2)| : |t_1 - t_2| < \delta|t|\} < \varepsilon. \tag{1.21}$$

Но из (1.10) следует, что существуют ε и δ такие, что для всех $k \ge k_0$

$$\mathsf{P}(N_k \notin \mathcal{N}(k,\delta)) = \mathsf{P}\left(\left|\frac{b_{N_k} - b_k}{d_k}\right| > \delta\right) < \varepsilon. \tag{1.22}$$

Подставляя (1.21) и (1.22) в (1.20), мы получаем

$$|g_k(t) - r_k(t)| < 2\varepsilon,$$

что в силу произвольности ε означает, что имеет место (1.16), а потому из Леммы 1.1 вытекает, что выполнено (1.14). Теорема доказана.

Из Теоремы 1.1 следует, что при асимптотически неслучайных индексах (в смысле (1.9) или (1.10)) функция распределения $F_k(x) = \mathsf{P}(Z_k < x)$ неслучайно центрированной и нормированной случайной суммы S_{N_k} может быть аппроксимирована функцией распределения

$$R_k(x) = W_k(x) * H\left(x\frac{d_k}{b_k}\right), \tag{1.23}$$

где $W_k(x) = \mathsf{P}(a_{N_k} - c_k < xd_k)$, $H(x) = \mathsf{P}(Y < x)$, а символ $*$ здесь и в дальнейшем обозначает свертку распределений. Несложно убедиться, что функция распределения R_k является дискретной сдвиговой смесью функций распределения H. Действительно, в соответствии с (1.23), $R_k(x)$ – это функция распределения суммы двух независимых случайных величин Yb_k/d_k и $V_k = (a_{N_k} - c_k)/d_k$. Тогда по формуле полной вероятности

$$R_k(x) = \mathsf{P}(V_k + Yb_k/d_k < x) =$$
$$\sum_{n=1}^{\infty} \mathsf{P}(N_k = n)\mathsf{P}\left(\frac{a_n - c_k}{d_k} + Y\frac{b_k}{d_k} < x\right) =$$
$$\sum_{n=1}^{\infty} \mathsf{P}(N_k = n)\mathsf{P}\left(Y < \frac{xd_k + c_k - a_n}{b_k}\right) =$$
$$\sum_{n=1}^{\infty} \mathsf{P}(N_k = n)H\left(\frac{xd_k + c_k - a_n}{b_k}\right).$$

Подобные распределения впервые появились в качестве аппроксимаций для распределений неслучайно центрированных случайных сумм в работе Г. Роббинса (Robbins, 1948), который рассмотрел суммы одинаково распределенных слагаемых с конечной дисперсией. Позднее результаты Роббинса были обобщены З. Рыхликом и Д. Шиналем на суммы необязательно одинаково распределенных слагаемых (также с конечными дисперсиями) (Rychlik and Szynal, 1972). Теорема 1.1 обобщает упомянутые

результаты, так как в ее формулировке нет никаких предположений о существовании моментов.

Лемма 1.1 описывает *асимптотическое сближение* распределений случайных сумм с аппроксимирующими законами, которые также зависят от номера суммы. В следующем разделе мы рассмотрим *сходимость* распределений случайных сумм.

2 Теорема переноса. Предельные законы

Утверждения, описывающие перенос асимптотических свойств сумм независимых случайных величин в неслучайном числе на случайные суммы, принято называть теоремами переноса. Впервые этот термин был использован в работе (Гнеденко и Фахим, 1969).

Пусть, как и ранее, $H(x) = \mathsf{P}(Y < x)$.

Теорема 2.1. *Пусть числовые последовательности $\{a_k\}, \{b_k\}, \{c_k\}$ и $\{d_k\}$ таковы, что $b_k \to \infty, d_k \to \infty$ $(k \to \infty)$ и выполнено условие (1.1). Дополнительно предположим, что имеет место слабая сходимость*

$$\left(\frac{b_{N_k}}{d_k}, \frac{a_{N_k} - c_k}{d_k}\right) \Rightarrow (U, V) \qquad (k \to \infty) \tag{2.1}$$

где U и V – некоторые случайные величины. Тогда

$$\mathsf{P}(Z_k < x) \Rightarrow \mathsf{E}H\left(\frac{x - V}{U}\right) \qquad (k \to \infty). \tag{2.2}$$

Доказательство. Легко видеть, что распределению, стоящему в правой части соотношения (2.2), соответствует характеристическая функция

$$f(t) = \mathsf{E}h(Ut)e^{itV}, \quad t \in \mathbb{R},$$

где h – характеристическая функция случайной величины Y из (1.1). Поэтому мы докажем, что имеет место (2.2), если мы убедимся, что для любого $t \in \mathbb{R}$

$$\lim_{k\to\infty} f_k(t) = f(t). \tag{2.3}$$

Обозначим $U_k = b_{N_k}/d_k, V_k = (a_{N_k} - c_k)/d_k$. Тогда, очевидно, характеристическую функцию f_k можно записать в виде

$$f_k(t) = \mathsf{E}h_k(U_k t)e^{itV_k},$$

где, как и ранее, h_k – характеристическая функция случайной величины $Y_k = (S_k - a_k)/b_k$. Поскольку

$$|f_k(t) - f(t)| \le |f_k(t) - g_k(t)| + |g_k(t) - f(t)|,$$

с учетом Леммы 1.1 достаточно показать, что

$$\lim_{k\to\infty} |g_k(t) - f(t)| = 0.$$

Но $g_k(t) = \mathsf{E}h(U_k t)e^{itV_k}$. Заметим, что функция $\phi_t(x, y) = h(tx)e^{ity}$ при каждом фиксированном $t \in \mathbb{R}$ непрерывна и ограничена по x и y. В соответствии с определением слабой сходимости, соотношение (2.1) эквивалентно тому, что

$$\mathsf{E}\phi(U_k, V_k) \to \mathsf{E}\phi(U, V) \quad (k \to \infty) \tag{2.4}$$

для любой непрерывной ограниченной функции ϕ, откуда следует, что (2.4) имеет место и для $\phi = \phi_t$. Таким образом, (2.1) влечет (2.3) и, следовательно, (2.2). Теорема доказана.

Заметим, что в некоторых случаях, когда центрирующие и нормирующие постоянные имеют специальный вид, достаточно потребовать слабой сходимости лишь одной из компонент пары (U_k, V_k). Например, в работах (Finkelstein and Tucker, 1989) и (Finkelstein, Kruglov and Tucker, 1994) условия сходимости случайных величин Z_k сформулированы в терминах V_k (мы рассмотрим этот случай подробно при изучении асимптотического поведение процессов риска). В качестве примера рассмотрим ситуацию, где $a_k = c_k = \alpha k$, $b_k = d_k = \sigma k^\gamma$, $\alpha \in \mathbb{R}$, $\sigma > 0$, $0 < \gamma < 1$. Тогда слабая сходимость случайных величин $V_k = \frac{\alpha}{\sigma}\left(\frac{N_k - k}{k^\gamma}\right)$ к собственной случайной величине V при $k \to \infty$ влечет

$$\mathsf{P}\left(\left|\left(\frac{b_{N_k}}{d_k}\right)^{1/\gamma} - 1\right| > x\right) = \mathsf{P}\left(\left|\frac{N_k}{k} - 1\right| > x\right) =$$

$$\mathsf{P}\left(\left|\frac{N_k - k}{k^\gamma}\right| > k^{1-\gamma}x\right) \to 0 \quad (k \to \infty)$$

при каждом $x > 0$, что означает, что $(b_{N_k}/d_k)^{1/\gamma} \to 1$, и потому $U_k \to 1$ при $k \to \infty$. С учетом (1.1), в соответствии с Теоремой 2.1 мы приходим к заключению, что в этом случае распределение, предельное для случайных величин Z_k, имеет вид $(H * W)(x)$, где $W(x) = \mathsf{P}(V < x)$.

Из Теоремы 2.1 видно, что класс распределений, предельных для "нарастающих" случайных сумм при неслучайном центрировании и нормировании состоит из сдвиг-масштабных смесей распределений, предельных для сумм с неслучайным числом слагаемых. Поскольку по предположению нормирующие постоянные b_k и d_k положительны при любом $k \ge 1$, случайная величина U неотрицательна. Поэтому мы можем считать, что

$$H\left(\frac{x - v}{u}\right)\bigg|_{u=0} = \mathbf{I}(v < x),$$

какой бы ни была функция распределения H, где $\mathbf{I}(A)$ – индикаторная функция множества A.

Если Z – случайная величина с функцией распределения $\mathsf{E}H\left(\frac{x-V}{U}\right)$, то

$$Z \stackrel{d}{=} YU + V, \tag{2.5}$$

где случайная величина Y и пара (U, V) независимы. Таким образом, мы приходим к неутешительному выводу: если не делать никаких дополнительных предположений о распределениях индексов или нормирущих и центрирующих постоянных, то множество законов, предельных для "нарастающих" случайных сумм, совпадает с множеством всех функций распределения, так как в случае $\mathsf{P}(U = 0) = 1$ распределение случайной величины Z в (2.5) совпадает с распределением случайной величины V, которая, вообще говоря, может быть совершенно произвольной.

При дополнительных предположениях класс предельных законов сужается. Так, в случае, когда $\mathsf{P}(Y = a) = 1$ для некоторого $a \in \mathbb{R}$ (например, в ситуации, описываемой законом больших чисел), при $\mathsf{P}(V = 0) = 1$ распределение предельной случайной величины Z совпадает с распределением величины aU. Это распределение сосредоточено на полуоси $(-\infty, 0]$ или на полуоси $[0, \infty)$ в зависимости от знака числа a, но среди всех таких распределений оно может быть произвольным.

Таким образом, если мы предположим, что имеет место (1.1), то есть зафиксируем случайную величину Y в представлении (2.5), то предельные закноны можно описать подробнее. Если в дополнение к (1.1) $\mathsf{P}(V = const) = 1$, то с точностью до неслучайного сдвига класс предельных распределений совпадает с множеством масштабных смесей функции распределения $H(x) = \mathsf{P}(Y < x)$. Если $\mathsf{P}(U = const) = 1$, то с точностью до выбора параметра масштаба класс предельных распределений совпадает с множеством сдвиговых смесей функции распределения $H(x)$, что мы видели на примере Теоремы 1.1.

Увы, более детальное описание класса предельных распределений для "нарастающих" случайных сумм представляется невозможным.

Следует отметить, что в доказательствах утверждений этого и предыдущего разделов мы практически не использовали базовое структурное предположение о том, что случайные величины S_k представляют собой накапливаемые суммы независимых случайных величин. Другими словами, приведенные результаты на самом деле справедливы для произвольных случайных последовательностей с независимыми случайными индексами (см. (Королев, 1994)).

3 Необходимые и достаточные условия сходимости случайных сумм

Попытки сформулировать необходимые и достаточные условия слабой сходимости неслучайно центрированных и нормированных "нарастающих" случайных сумм наталкиваются на одно очень серьезное препятствие, которое не позволяет нам доказать не только достаточность, но и необходимость условий (1.1) или (2.1) для (2.2) в общем случае. А именно, в представлении (2.5) случайная величина Y и пара (U, V), независимая от Y, не определяются однозначно по распределению случайной величины Z. (Поскольку мы имеем дело со слабой сходимостью, здесь и далее мы отождествляем случайные величины и их распределения с целью упрощения формулировок и обозначений. Фактически мы рассматриваем классы эквивалентности случайных величин, каждый из которых включает в себя все случайные величины с одним и тем же распределением.)

Чтобы преодолеть это препятствие, для произвольной случайной величины Z введем множество

$$\mathcal{V}(Z) = \{(Y, U, V) : Z \stackrel{d}{=} YU + V, \quad Y \text{и пара } (U, V) \text{ независимы}\}.$$

Какой бы ни была случайная величина Z, множество $\mathcal{V}(Z)$ непусто, так как всегда $(Y, 0, Z) \in \mathcal{V}(Z)$, где Y – произвольная случайная величина, независимая от Z. Из этого примера также следует, что множество $\mathcal{V}(Z)$ всегда содержит более одного элемента. Множество $\mathcal{V}(Z)$ может содержать различные тройки и с ненулевой случайной величиной U. Например, пусть $Z \stackrel{d}{=} W - W_1$, где W и W_1 – независимые случайные величины с одним и тем же гамма-распределением с некоторыми параметрами масштаба $\lambda > 0$ и формы $\alpha > 0$. Тогда множество $\mathcal{V}(Z)$, очевидно, будет содержать тройки $(1, W, -W_1)$ и $(W, 1, -W_1)$ наряду с тривиальными тройками $(Y, 0, W - W_1)$. Более того, в данном случае множество $\mathcal{V}(Z)$ также содержит и тройку $(Y, W_2^{1/2}, 0)$, где случайная величина Y имеет стандартное нормальное распределение и независима от случайной величины W_2, имеющей гамма-распределение с параметром масштаба $\frac{1}{2}\lambda^2$ и параметром формы α. Чтобы убедиться в этом, заметим, что характеристическая функция разности $W - W_1$ равна

$$\mathsf{E}e^{it(W-W_1)} = \frac{1}{\left(1 - \dfrac{it}{\lambda}\right)^{\alpha}} \cdot \frac{1}{\left(1 + \dfrac{it}{\lambda}\right)^{\alpha}} = \left(\frac{\lambda^2}{\lambda^2 + t^2}\right)^{\alpha}. \tag{3.1}$$

Пусть W_2 – гамма-распределенная случайная величина с параметром формы α и некоторым параметром масштаба μ и пусть Y – независимая от нее случайная величина со стандартным нормальным распределением. Тогда

$$\mathsf{E}e^{itYW_2^{1/2}} = \frac{\mu^\alpha}{\Gamma(\alpha)}\int_0^\infty e^{-x\left(\frac{1}{2}t^2+\mu\right)}x^{\alpha-1}\,dx =$$
$$\frac{\mu^\alpha}{\Gamma(\alpha)\left(\frac{1}{2}t^2+\mu\right)^\alpha}\int_0^\infty e^{-y}y^{\alpha-1}\,dy = \left(\frac{2\mu}{2\mu+t^2}\right)^\alpha. \tag{3.2}$$

Правые части (3.1) и (3.2) совпадают, если $\mu = \frac{1}{2}\lambda^2$, что и требовалось доказать.

Пусть $L_1(\cdot,\cdot)$ и $L_2(\cdot,\cdot)$ – метрики соответственно в пространствах одно- и двумерных распределений (или, что то же самое, случайных величин), метризующие слабую сходимость. Например, $L_1(\cdot,\cdot)$ – это метрика Леви, $L_2(\cdot,\cdot)$ – метрика Леви-Прохорова (см. Введение). В соответствии со сказанным выше, если X и Y – случайные величины с функциями распределения G и H соответственно, то мы не будем различать $L_1(X,Y)$ и $L_1(G,H)$. Аналогично, если (X_1,X_2) и (Y_1,Y_2) – двумерные случайные векторы соответственно с функциями распределения G и H, то мы не будем делать различий между $L_2(G,H)$ и $L_2((X_1,X_2),(Y_1,Y_2))$.

Следующая теорема играет основную роль. Все остальные утверждения этого раздела (и большинство утверждений других разделов) являются ее следствиями.

Теорема 3.1. *Предположим, что*

$$N_k \xrightarrow{P} \infty \qquad (k\to\infty), \tag{3.3}$$

и последовательности $\{a_k\}$ и $\{b_k\}$, $b_k>0$, $b_k\to\infty$ $(k\to\infty)$, обеспечивают слабую компактность семейства случайных величин $\{(S_k-a_k)/b_k\}_{k\ge1}$. Сходимость

$$\frac{1}{d_k}\left(\sum_{j=1}^{N_k}X_j - c_k\right)\Rightarrow Z \qquad (k\to\infty) \tag{3.4}$$

к некоторой случайной величине Z имеет место при некоторых последовательностях положительных чисел $\{d_k\}$, $d_k\to\infty$ $(k\to\infty)$, и вещественных чисел $\{c_k\}$ тогда и только тогда, когда существует слабо компактная последовательность троек случайных величин $\{(Y_k',U_k',V_k')\}_{k\ge1}$ таких, что $(Y_k',U_k',V_k')\in\mathcal{V}(Z)$ при каждом $k\ge1$ и

$$L_1\left(\frac{1}{b_k}\left(\sum_{j=1}^{k}X_j - a_k\right), Y_k'\right)\to 0 \qquad (k\to\infty), \tag{3.5}$$

$$L_2\left(\left(\frac{b_{N_k}}{d_k}, \frac{a_{N_k}-c_k}{d_k}\right), (U_k',V_k')\right)\to 0 \qquad (k\to\infty). \tag{3.6}$$

Доказательство. Необходимость. Пусть X_j' – случайная величина, независимая от X_j и такая, что $X_j' \stackrel{d}{=} X_j$. Обозначим $X_j^{(s)} = X_j' - X_j$ ($X_j^{(s)}$ – это симметризация

случайной величины X_j). Пусть $q\in(0,1)$. Точную нижнюю грань q-квантилей случайной величины N_k обозначим $l_k(q)$. Предполагая, что случайные величины N_k, X_1, $X_2,\ldots$, X_1', $X_2',\ldots$ независимы при каждом $k\ge1$, определим

$$T_k = \frac{1}{d_k}\sum_{j=1}^{N_k}X_j^{(s)}.$$

Используя неравенство симметризации (см. Лоэв, 1962)

$$\mathsf{P}(|X_1^{(s)}|\ge x)\le 2\mathsf{P}\left(|X_1-a|\ge\frac{x}{2}\right),$$

справедливое для любой случайной величины X_1 и любых $a\in\mathbb{R}$, $x>0$, получим

$$\mathsf{P}(|T_k|\ge x) = \sum_{n=1}^\infty \mathsf{P}(N_k=n)\mathsf{P}\left(\left|\frac{1}{d_k}\sum_{j=1}^n X_j^{(s)}\right|\ge x\right)\le$$
$$2\sum_{n=1}^\infty \mathsf{P}(N_k=n)\mathsf{P}\left(\left|\frac{1}{d_k}\left(\sum_{j=1}^n X_j - c_k\right)\right|\ge\frac{x}{2}\right) =$$
$$2\mathsf{P}\left(\left|\frac{1}{d_k}\left(\sum_{j=1}^{N_k} X_j - c_k\right)\right|\ge\frac{x}{2}\right)$$

для любого $x>0$. Отсюда следует, что

$$\lim_{x\to\infty}\sup_k \mathsf{P}(|T_k|\ge x)\le$$
$$2\lim_{x\to\infty}\sup_k \mathsf{P}\left(\left|\frac{1}{d_k}\left(\sum_{j=1}^{N_k} X_j - c_k\right)\right|\ge\frac{x}{2}\right) = 0$$

в силу условия (3.4). Согласно Теореме 3 на с. 43 в (Гнеденко и Колмогоров, 1949), это означает, что последовательность $\{T_k\}_{k\ge1}$ слабо компактна. Наша ближайшая цель – установить слабую компактность последовательности $\{b_{N_k}/d_k\}_{k\ge1}$.

С этой целью убедимся, что при каждом $q\in(0,1)$

$$C(q)\equiv\sup_k\frac{b_{l_k(q)}}{d_k}<\infty. \tag{3.7}$$

Используя неравенство Леви

$$\mathsf{P}\left(\max_{1\le r\le n}\left|\sum_{j=1}^r X_j^{(s)}\right|\ge x\right)\le 2\mathsf{P}\left(\left|\sum_{j=1}^n X_j^{(s)}\right|\ge x\right),$$

справедливое для любых независимых случайных величин $X_1, \ldots, X_n$ и любого $x>0$ (см., например, (Лоэв, 1962)), мы получим

$$\mathsf{P}(|T_k| \geq x) = \mathsf{P}\left(\left|\frac{1}{d_k}\sum_{j=1}^{N_k} X_j^{(s)}\right| \geq x\right) =$$
$$\sum_{n=1}^{\infty}\mathsf{P}(N_k = n)\mathsf{P}\left(\left|\frac{1}{d_k}\sum_{j=1}^{n} X_j^{(s)}\right| \geq x\right) \geq$$
$$\sum_{n>l_k(q)}\mathsf{P}(N_k = n)\mathsf{P}\left(\left|\frac{1}{d_k}\sum_{j=1}^{n} X_j^{(s)}\right| \geq x\right) \geq$$
$$\frac{1}{2}\sum_{n>l_k(q)}\mathsf{P}(N_k = n)\mathsf{P}\left(\left|\frac{1}{d_k}\sum_{j=1}^{l_k(q)} X_j^{(s)}\right| \geq x\right) \geq$$
$$\frac{1}{2}(1-q)\mathsf{P}\left(\left|\frac{1}{d_k}\sum_{j=1}^{l_k(q)} X_j^{(s)}\right| \geq x\right).$$

Таким образом, слабая компактность последовательности $\{T_k\}_{k\geq 1}$, установленная выше, по упомянутой выше Теореме 3 из книги (Гнеденко и Колмогоров, 1949), с. 43, влечет слабую компактность последовательности функций распределения $\{F_k^{(q)}\}_{k\geq 1}$ при каждом $q \in (0,1)$, где

$$F_k^{(q)}(x) = \mathsf{P}\left(\left|\frac{1}{d_k}\sum_{j=1}^{l_k(q)} X_j^{(s)}\right| \leq x\right).$$

Предположим, что (3.7) не имеет места. Тогда существуют $q \in (0,1)$ и последовательность $\mathcal{K}$ натуральных чисел такие, что

$$\frac{b_{l_k(q)}}{d_k} \to \infty \qquad (k \to \infty, k \in \mathcal{K}). \tag{3.8}$$

Так как $N_k \xrightarrow{P} \infty$ $(k \to \infty)$, то $l_k(q) \to \infty$ $(k \to \infty)$. В силу слабой компактности семейства $\{Y_k\}$, $Y_k = (S_k - a_k)/b_k$, из последовательности $\mathcal{K}$ можно выделить подпоследовательность $\mathcal{K}_1 \subseteq \mathcal{K}$ такую, что

$$Y_{l_k(q)} = \frac{1}{b_{l_k(q)}}\left(\sum_{j=1}^{l_k(q)} X_j - a_{l_k(q)}\right) \Rightarrow Y \tag{3.9}$$

при $k \to \infty$, $k \in \mathcal{K}_1$. Пусть Y' – случайная величина, независимая от Y и такая, что $Y' \stackrel{d}{=} Y$. Функцию распределения $F_k^{(q)}$ можно записать в виде

$$F_k^{(q)}(x) = \mathsf{P}\left(\frac{1}{b_{l_k(q)}}\left(\sum_{j=1}^{l_k(q)} X_j - a_{l_k(q)}\right) - \right.$$
$$\left.\frac{1}{b_{l_k(q)}}\left(\sum_{j=1}^{l_k(q)} X_j' - a_{l_k(q)}\right) \leq \frac{d_k x}{b_{l_k(q)}}\right).$$

Поэтому из (3.8) и (3.9) вытекает, что при любом $x \in \mathbb{R}$

$$F_k^{(q)}(x) \to \mathsf{P}(Y - Y' \leq 0) \geq \frac{1}{2} \qquad (k \to \infty,\ k \in \mathcal{K}_1).$$

Но это противоречит слабой компактности семейства функций распределения $\{F_k^{(q)}(x)\}_{k\geq 1}$. Таким образом, мы доказали, что (3.7) имеет место для любого $q \in (0,1)$.

Несложно убедиться, что случайная величина N_k имеет такое же распределение, как $l_k(W)$, где W – случайная величина, равномерно распределенная на интервале $(0,1)$. Поэтому, используя (3.7), для любого $x \geq 0$ мы получим

$$\mathsf{P}\left(\frac{b_{N_k}}{d_k} \geq x\right) = \mathsf{P}\left(\frac{b_{l_k(W)}}{d_k} \geq x\right) = \int_0^1 \mathsf{P}\left(\frac{b_{l_k(q)}}{d_k} \geq x\right)dq \leq$$
$$\int_0^1 \mathsf{P}(C(q) \geq x)dq = \mathsf{P}(C(W) \geq x).$$

Поэтому

$$\lim_{x\to\infty}\sup_k \mathsf{P}\left(\frac{b_{N_k}}{d_k} \geq x\right) \leq \lim_{x\to\infty}\mathsf{P}(C(W) \geq x) = 0,$$

то есть последовательность случайных величин $\{b_{N_k}/d_k\}_{k\geq 1}$ слабо компактна.

Теперь докажем, что последовательность $\{(a_{N_k} - c_k)/d_k\}_{k\geq 1}$ слабо компактна. Для произвольного $R > 0$ по формуле полной вероятности мы будем иметь

$$\mathsf{P}\left(\left|\frac{a_{N_k} - c_k}{d_k}\right| > R\right) =$$
$$\sum_{n=1}^{\infty}\mathsf{P}(N_k = n)\mathsf{P}\left(\left|\frac{S_n - c_k}{d_k} - \frac{b_n}{d_k}\left(\frac{S_n - a_n}{b_n}\right)\right| > R\right) \leq$$
$$\sum_{n=1}^{\infty}\mathsf{P}(N_k = n)\mathsf{P}\left(\left|\frac{S_n - c_k}{d_k}\right| > \frac{R}{2}\right) +$$
$$\sum_{n=1}^{\infty}\mathsf{P}(N_k = n)\mathsf{P}\left(\left|\frac{S_n - a_n}{b_n}\right| > \frac{Rd_k}{2b_n}\right) \equiv I_1(k,R) + I_2(k,R). \tag{3.10}$$

Сначала рассмотрим $I_2(k,R)$. Для $M > 0$ обозначим

$$\mathcal{N}_0 = \mathcal{N}_0(k,M) = \{n : b_n \leq Md_k\}.$$

Имеем

$$I_2(k,R) = \sum_{n\in\mathcal{N}_0(k,M)}\mathsf{P}(N_k = n)\mathsf{P}\left(|Y_k| > \frac{Rd_k}{2b_n}\right) +$$

$$\sum_{n \notin \mathcal{N}_0(k,M)} \mathsf{P}(N_k = n)\mathsf{P}\left(|Y_k| > \frac{Rd_k}{2b_n}\right) \leq$$

$$\sup_k \mathsf{P}\left(|Y_k| > \frac{R}{2M}\right) + \mathsf{P}\left(\frac{b_{N_k}}{d_k} \geq M\right). \tag{3.11}$$

Фиксируем произвольное $\varepsilon > 0$. Выберем $M = M(\varepsilon) > 0$ так, чтобы

$$\sup_k \mathsf{P}\left(\frac{b_{N_k}}{d_k} \geq M(\varepsilon)\right) < \varepsilon. \tag{3.12}$$

Такой выбор возможен вследствие слабой компактности семейства $\{b_{N_k}/d_k\}_{k\geq 1}$, установленной выше. Теперь выберем $R_2 = R_2(\varepsilon)$ так, чтобы для любого $R > R_2$

$$\sup_k \mathsf{P}\left(|Y_k| > \frac{R}{2M(\varepsilon)}\right) < \varepsilon. \tag{3.13}$$

Такой выбор возможен вследствие слабой компактности семейства $\{Y_k\}_{k\geq 1}$, $Y_k = (S_k - a_k)/b_k$. Таким образом, из (3.11), (3.12) и (3.13) мы получаем

$$I_2(k, R) < 2\varepsilon \tag{3.14}$$

для любого $R > R_2$ и любого $k \geq 1$.

Теперь рассмотрим $I_1(k, R)$. Из (3.4) следует, что для любого $\varepsilon > 0$ существует $R_1 = R_1(\varepsilon)$ такое, что

$$I_1(k, R) < \varepsilon \tag{3.15}$$

Для любого $R > R_1$ и любого $k \geq 1$.

Из (3.10), (3.14) и (3.15) следует, что при $R > \max(R_1, R_2)$ мы имеем

$$\sup_k \mathsf{P}\left(\left|\frac{a_{N_k} - c_k}{d_k}\right| > R\right) < 3\varepsilon,$$

что означает, что последовательность случайных величин $\{(a_{N_k} - c_k)/d_k\}_{k\geq 1}$ слабо компактна в силу произвольности ε. Вместе со слабой компактностью семейства $\{b_{N_k}/d_k\}_{k\geq 1}$ это означает, что слабо компактно семейство пар $\{(b_{N_k}/d_k, (a_{N_k} - c_k)/d_k)\}_{k\geq 1}$.

Обозначим

$$U_k = \frac{b_{N_k}}{d_k}, \quad V_k = \frac{a_{N_k} - c_k}{d_k},$$

$$\gamma_k = \inf\{L_1(Y_k, Y) + L_2((U_k, V_k,), (U, V)) : (Y, U, V) \in \mathcal{V}(Z)\}.$$

Докажем, что $\gamma_k \to 0$ при $k \to \infty$. Предположим противное. В этом случае при некотором $\delta > 0$ и всех k из некоторой последовательности $\mathcal{K}_2$ натуральных чисел должно выполняться неравенство $\gamma_k \geq \delta$. Выберем подпоследовательность $\mathcal{K}_3 \subseteq \mathcal{K}_2$

так, чтобы последовательности случайных величин $\{Y_k\}_{k\in\mathcal{K}_3}$ и пар $\{(U_k, V_k)\}_{k\in\mathcal{K}_3}$ слабо сходились соответственно к некоторой случайной величине Y и паре (U, V) при $k \to \infty$, $k \in \mathcal{K}_3$. При этом для всех достаточно больших $k \in \mathcal{K}_3$ должно выполняться неравенство

$$L_1(Y_k, Y) + L_2((U_k, V_k), (U, V)) < \delta. \tag{3.16}$$

Применяя Теорему 2.1 к последовательностям $\{Y_k\}_{k\in\mathcal{K}_3}$ и $\{(U_k, V_k)\}_{k\in\mathcal{K}_3}$, мы убеждаемся, что $(Y, U, V) \in \mathcal{V}(Z)$, так как выполнено (3.4) и потому предел должен быть одним и тем же для всех сходящихся последовательностей. Но тогда (3.16) противоречит предположению о том, что $\gamma_k \geq \delta$ при всех $k \in \mathcal{K}_3$. Тем самым мы доказали, что $\gamma_k \to 0$ при $k \to \infty$.

Для каждого $k = 1, 2, \ldots$ выберем тройку (Y_k', U_k', V_k') из $\mathcal{V}(Z)$, для которой

$$L_1(Y_k, Y_k') + L_2((U_k, V_k), (U_k', V_k')) \leq \gamma_k + \frac{1}{k}.$$

Последовательность троек (Y_k', U_k', V_k'), очевидно, удовлетворяет условиям (3.5) и (3.6). Ее слабая компактность вытекает из (3.5) и (3.6) и того факта, что слабо компактными являются последовательности $\{Y_k\}_{k\geq 1}$ и $\{(U_k, V_k)\}_{k\geq 1}$.

Достаточность. Предположим, что последовательность $\{Z_k\}$, $Z_k = (S_{N_k} - c_k)/d_k$, не сходится слабо к Z. Это означает, что для некоторого $\delta > 0$ и всех k из некоторой последовательности $\mathcal{K}$ натуральных чисел выполняется неравенство $L_1(Z_k, Z) \geq \delta$. Слабая компактность последовательности $\{(Y_k', U_k', V_k')\}_{k\geq 1}$ позволяет нам выбрать подпоследовательность $\mathcal{K}_1 \subseteq \mathcal{K}$ так, что $Y_k' \Rightarrow Y$ и $(U_k', V_k') \Rightarrow (U, V)$ при $k \to \infty$, $k \in \mathcal{K}_1$, где Y, U и V – некоторые случайные величины. Повторяя рассуждения, использованные при доказательстве Теоремы 2.1, мы убеждаемся, что для каждого $t \in \mathbb{R}$

$$\mathsf{E}\exp\{itZ\} = \mathsf{E}h_k'(tU_k')\exp\{itV_k'\} \to$$
$$\mathsf{E}h(Ut)\exp\{itV\} = \mathsf{E}\exp\{it(YU + U)\}$$

при $k \to \infty$, $k \in \mathcal{K}_1$, где случайная величина Y и пара (U, V) независимы, $h_k'(t) = \mathsf{E}\exp\{itY_k'\}$, $h(t) = \mathsf{E}\exp\{itY\}$. Это означает, что $(Y, U, V) \in \mathcal{V}(Z)$. Из неравенства

$$L_1(Y_k, Y) \leq L_1(Y_k, Y_k') + L_1(Y_k', Y)$$

и условия (3.5) вытекает, что $L_1(Y_k, Y) \to 0$ при $k \to \infty$, $k \in \mathcal{K}_1$. Аналогично, неравенство

$$L_2((U_k, V_k), (U, V)) \leq$$
$$L_2((U_k, V_k), (U_k', V_k')) + L_2((U_k', V_k'), (U, V))$$

и условие (3.6) влекут $L_2((U_k, V_k), (U, V)) \to 0$ при $k \to \infty$, $k \in \mathcal{K}_1$. Применим Теорему 2.1 к последовательностям $\{Y_k\}_{k\in\mathcal{K}_1}$ и $\{(U_k, V_k)\}_{k\in\mathcal{K}_1}$. В результате получим $L_1(Z_k, Z) \to 0$ при $k \to \infty$, $k \in \mathcal{K}_1$, что противоречит предположению о том, что $L_1(Z_k, Z) \geq \delta$ при всех $k \in \mathcal{K}_1$. Теорема доказана.

Формулировка Теоремы 3.1 содержит условие слабой компактности семейства случайных величин $\{(S_k - a_k)/b_k\}_{k\ge1}$. Рассмотрим вопрос о том, насколько это условие ограничительно. Оказывается, что в случае, когда слагаемые в суммах распределены одинаково, от этого условия удается избавиться совсем, но, к сожалению, подробное обсуждение этой ситуации выходит далеко за пределы данного курса. Однако, совсем просто убедиться, что это условие автоматически выполняется, если мы потребуем существования моментов слагаемых хоть какого-нибудь положительного порядка. Сказанное будет проиллюстрировано Теоремой 3.2. Наконец, во многих ситуациях вполне уместно заменить условие слабой компактности последовательности $\{Y_k\}$ условием ее слабой сходимости. Как мы увидим в Теореме 3.4, в этом случае формулировки существенно упрощаются.

Введем следующие обозначения: $m_j = \mathsf{E}X_j$, $A_j = m_1 + \ldots + m_j$, $\sigma_j^2 = \mathsf{D}X_j$, $B_j^2 = \sigma_1^2 + \ldots + \sigma_j^2$.

Теорема 3.2. *Предположим, что выполнено (3.3) и $\sigma_j^2 < \infty, j \ge 1$, так, что $B_k^2 \to \infty (k \to \infty)$. Сходимость (3.4) имеет место с некоторыми последовательностями положительных чисел $\{d_k\}_{k\ge1}, d_k \to \infty (k \to \infty)$, и вещественных чисел $\{c_k\}_{k\ge1}$, если и только если существует слабо компактная последовательность троек случайных величин $\{(Y_k', U_k', V_k')\}_{k\ge1}$ таких, что $(Y_k', U_k', V_k') \in \mathcal{V}(Z)$ при каждом $k \ge 1$ и выполнены условия (3.5) и (3.6) с $b_k = B_k$ и $a_k = A_k$.*

Доказательство сводится к Теореме 3.1, поскольку слабая компактность семейства $\{(S_k - a_k)/b_k\}_{k\ge1}$ вытекает из неравенства Чебышева:

$$\lim_{x\to\infty} \sup_k \mathsf{P}\left(\left|\frac{S_k - A_k}{B_k}\right| \ge x\right) \le \lim_{x\to\infty} \frac{1}{x^2} = 0.$$

Теорема доказана.

Теорема 3.1 представляет собой полное обращение Теоремы 2.1. Однако возможны ситуации, когда либо свойства слагаемых, либо свойства индексов обеспечивают выполнение одного из условий (1.1) или (2.1). Сейчас мы приведем две соответствующие модификации Теоремы 3.1.

Начнем с ситуации, когда выполняется (2.1). Для произвольной случайной величины Z и произвольной пары (U, V) с $\mathsf{P}(U \ge 0) = 1$ введем множество

$$\mathcal{U}(Z \mid U, V) = \{Y : Z \overset{d}{=} YU + V,\ Y \text{ и } (U, V) \text{ независимы}\}.$$

Основное отличие этой ситуации от той, которая рассматривалась в Теореме 3.1, заключается в том, что при некоторых Z, U и V множество $\mathcal{U}(Z \mid U, V)$ может оказаться пустым. Действительно, пусть, например, $\mathsf{P}(V = 0) = 1$ и пусть случайная величина U имеет стандартное показательное распределение. Пусть Z – положительная случайная величина. Тогда условие $\mathsf{P}(Y > 0) = 1$ необходимо для того, чтобы было справедливо представление $Z \overset{d}{=} YU$. В этом случае для любого $x \ge 0$

$$\mathsf{P}(YU > x) = \int_0^\infty e^{-sx} dG(s), \qquad (3.17)$$

где $G(s) = \mathsf{P}(Y^{-1} < s)$. Но в правой части (3.17) стоит преобразование Лапласа-Стилтьеса функции распределения G. Поэтому, если в качестве Z мы возьмем такую случайную величину, что фуекция $Q(s) = \mathsf{P}(Z > s)$ не является вполне монотонной, то представление $Z \overset{d}{=} YU$ станет невозможным ни при какой случайной величине Y, независимой от U, поскольку в соответствии с теоремой С. Н. Бернштейна (см., например, (Феллер, 1984), глава XIII), любое преобразование Лапласа-Стилтьеса вполне монотонно.

Теорема 3.3. *Предположим, что последовательности $\{a_k\}$, $\{b_k\}$, $\{c_k\}$, $\{d_k\}$, $b_k > 0$, $b_k \to \infty$, $d_k > 0$, $d_k \to \infty$ $(k \to \infty)$ обеспечивают сходимость (2.1) и слабую компактность семейства случайных величин $\{(S_k - a_k)/b_k\}_{k\ge1}$. Сходимость (3.4) имеет место тогда и только тогда, когда существует слабо компактная последовательность случайных величин $\{Y_k'\}_{k\ge1}$ таких, что $Y_k' \in \mathcal{U}(Z \mid U, V)$ при каждом $k \ge 1$ и выполнено (3.5).*

Доказательство этой теоремы сводится к Теореме 3.1. Заметим, что условие (3.3) в рассматриваемом случае излишне, так как в доказательстве Теоремы 3.1 оно использовалось лишь для того, чтобы установить слабую компактность последовательности $\{b_{N_k}/d_k\}_{k\ge1}$, которая в данном случае имеет место автоматически вследствие условия (2.1). Теорема доказана.

Чтобы сформулировать еще один вариант частичного обращения Теоремы 2.1, для двух произвольных случайных величин Z и Y введем множество

$$\mathcal{W}(Z|Y) = \{(U, V) : Z \overset{d}{=} YU + V,\ Y \text{ и } (U, V) \text{ независимы}\},$$

содержащее все пары случайных величин (U, V), независимые от Y, которые обеспечивают представление $Z \overset{d}{=} YU + V$. Используя те же рассуждения, что и при описании свойств множества $\mathcal{V}(Z)$, мы можем убедиться, что, во-первых, множество $\mathcal{W}(Z|Y)$ всегда не пусто, и во-вторых, при некоторых Z и Y множество $\mathcal{W}(Z|Y)$ может содержать более одного элемента. Следующая теорема дает возможность привлечь понятие идентифицируемости семейства смесей к изучению асимптотических свойств "нарастающих" случайных сумм. Позднее мы убедимся, что это приводит к значительному упрощению формулировок.

Теорема 3.4. *Предположим, что имеет место (3.3) и выполнено (1.1) с некоторыми последовательностями положительных чисел $\{b_k\}$, $b_k \to \infty$, $(k \to \infty)$, вещественных чисел $\{a_k\}$ и случайной величиной Y. Сходимость (3.4) к некоторой случайной величине Z имеет место с некоторыми последовательностями положительных чисел $\{d_k\}$, $d_k \to \infty$, $(k \to \infty)$ и вещественных чисел $\{c_k\}$ тогда и только тогда, когда существует слабо компактная последовательность пар $\{(U_k', V_k')\}_{k\ge1}$ таких, что $(U_k', V_k') \in \mathcal{W}(Z|Y)$ при каждом $k \ge 1$ и выполнено (3.6).*

Эта теорема является прямым следствием Теоремы 3.1.

В свою очередь, Теорема 3.4 позволяет получить совсем простые условия сходимости "нарастающих" случайных сумм центрированных слагаемых.

Определим подмножество

$$\mathcal{S}(Z|Y) = \{(U,V) \in \mathcal{W}(Z|Y) : \mathsf{P}(V=0)=1\}$$

множества $\mathcal{W}(Z|Y)$, которое состоит из тех случайных величин U, которые независимы от Y и допускают представление $Z \stackrel{d}{=} YU$. Поменяв ролями случайные величины Y и U в рассуждениях, которые мы использовали, чтобы описать свойства множества $\mathcal{U}(Z|U,V)$, мы можем убедиться, что при некоторых Z и Y множество $\mathcal{S}(Z|Y)$ может быть пустым. При некоторых Z и Y множество $\mathcal{S}(Z|Y)$ может содержать более одного элемента. Множество $\mathcal{S}(Z|Y)$ содержит не более одного элемента, если семейство масштабных смесей распределения $H(x) = \mathsf{P}(Y < x)$ идентифицируемо (разделимо) (более подробно см. ниже).

Теорема 3.5. *В условиях Теоремы 3.4 сходимость*

$$\frac{1}{d_k}\left(\sum_{j=1}^{N_k} X_j - a_{N_k}\right) \Rightarrow Z \qquad (k\to\infty) \tag{3.18}$$

к некоторой случайной величине Z имеет место с некоторыми последовательностями положительных чисел $\{d_k\}$, $d_k \to \infty$ $(k\to\infty)$ тогда и только тогда, когда существует слабо компактная последовательность случайных величин $\{U_k'\}_{k\geq 1}$ такая, что $U_k' \in \mathcal{S}(Z|Y)$ при каждом $k \geq 1$ и

$$L_1\left(\frac{b_{N_k}}{d_k}, U_k'\right) \to 0 \qquad (k\to\infty).$$

Доказательство. Обозначим $a_0 = 0$, $\alpha_j = a_j - a_{j-1}$, $j \geq 1$. Тогда, полагая $X_j^* = X_j - \alpha_j$, мы сведем доказательство к Теореме 3.4 с $a_k = c_k = 0$. Теорема доказана.

В следующих разделах мы рассмотрим условия сходимости распределений "нарастающих" случайных сумм к законам из идентифицируемых семейств.

4 Идентифицируемость смесей

Понятие идентифицируемости смесей интенсивно используется в прикладных задачах, связанных с разделением смесей, например, таких как статистическая классификация, распознавание образов или идентификация распределений. Библиография по данному кругу вопросов обширна, см., например, обзоры (Исаенко и Урбах, 1976), (Круглов, 1992), книгу (Айвазян, Бухштабер, Енюков и Мешалкин, 1989) и списки литературы в этих источниках. Мы будем интенсивно использовать понятие идентифицируемости семейств смесей в предельных теоремах для случайных сумм в следующих разделах.

Напомним определение идентифицируемых семейств смесей. Оно было предложено в работе (Teicher, 1961). Пусть функция $G(x,y)$ измерима по y при каждом x и является функцией распределения а как функция переменной x при каждом фиксированном y. Пусть $\mathcal{Q}$ – некоторое семейство случайных величин. Обозначим

$$\mathcal{F} = \{F_Q(x) = \mathsf{E}G(x,Q), x \in \mathbb{R} : Q \in \mathcal{Q}\}. \tag{4.1}$$

Определение 4.1. Семейство $\mathcal{F}$, определяемое ядром G и множеством $\mathcal{Q}$, называется идентифицируемым, если из равенства

$$\mathsf{E}G(x,Q_1) = \mathsf{E}G(x,Q_2), \qquad x \in \mathbb{R},$$

где $Q_1 \in \mathcal{Q}$, $Q_2 \in \mathcal{Q}$, вытекает, что $Q_1 \stackrel{d}{=} Q_2$.

В рассматриваемом нами случае $y = (u,v)$, $G(x,y) = H((x-v)/u)$. Поэтому мы можем воспользоваться хорошо известными результатами об идентифицируемости семейств одномерных распределений.

Для начала рассмотрим ситуацию, когда смешивание производится либо по параметру сдвига, либо по параметру масштаба, то есть рассмотрим однопараметрические семейства. Условия идентифицируемости таких семейств хорошо известны. Мы приведем некоторые из них.

Определение 4.2. Семейство функций распределения $\{G(x,y) : y > 0\}$ называется аддитивно замкнутым, если для любых $y_1 > 0, y_2 > 0$

$$G(x,y_1) * G(x,y_2) \equiv G(x, y_1+y_2). \tag{4.2}$$

Иногда свойство (4.2) семейств распределений называют воспроизводимостью по параметру y.

Семейство нормальных законов с нулевым средним является очевидным примером аддитивно замкнутого семейства. Несложно проверить, что аддитивно замкнутыми являются также семейства пуассоновских и биномиальных распределений (в последнем случае вероятность успеха фиксирована), гамма-распределений с фиксированным параметром масштаба.

Характеристическим свойством законов из аддитивно замкнутых семейств является то, что их характеристические функции можно представить в виде

$$\int\limits_{-\infty}^{\infty} e^{itx} dG(x,y) = [\phi(t)]^y, \tag{4.3}$$

где $\phi(t)$ – некоторая (комплекснозначная) функция, не зависящая от параметра y.

Пусть $\mathcal{Q} = \{Q : \mathsf{P}(Q > 0) = 1\}$, $P_Q(x) = \mathsf{P}(Q < x)$, $F_Q(x) = \mathsf{E}G(x, Q)$. Положим

$$f_Q(t) = \int\limits_{-\infty}^{\infty} e^{itx} dF_Q(x).$$

Если семейство, порожденное ядром $G(x,y)$, аддитивно замкнуто, то с учетом (4.3) мы имеем

$$f_Q(t) = \int\limits_{-\infty}^{\infty} e^{itx} d\,\mathsf{E}G(x,Q) = \int\limits_{-\infty}^{\infty} e^{itx} d\left(\int\limits_0^{\infty} G(x,y) dP_Q(y)\right) =$$

$$\int\limits_0^{\infty}\left(\int\limits_{-\infty}^{\infty} e^{itx} dG(x,y)\right) dP_Q(y) = \int\limits_0^{\infty} [\phi(t)]^y dP_Q(y).$$

Положим $z = \phi(t)$. Функция $\psi_Q(z) = \int_0^\infty z^y dP_Q(y)$ аналитична по крайней мере в области $D = \{z : 0 < |z| < 1\}$.

Предположим, что случайные величины Q_1 и Q_2 ($Q_1 \in \mathcal{Q}, Q_2 \in \mathcal{Q}$) с различными функциями распределения P_{Q_1} и P_{Q_2} образуют одну и ту же смесь, то есть $F_{Q_1}(x) \equiv F_{Q_2}(x)$. Но тогда в силу взаимно однозначного соответствия между распределениями и их характеристическими функциями функции $\psi_{Q_1}(z)$ и $\psi_{Q_2}(z)$ должны совпадать при всех z из множества $D_\phi = \{z = \phi(t) : -\infty < t < \infty\}$. Поскольку $D_\phi \subseteq D$ и функции $\psi_{Q_1}(\phi(t))$ и $\psi_{Q_2}(\phi(t))$ допускают аналитическое продолжение на всю область D, они также должны совпадать во всех точках $z \in D$. Следовательно, $\psi_{Q_1}(\rho e^{it}) = \psi_{Q_2}(\rho e^{it})$ для всех $\rho \in (0,1)$. Положим $\rho_n = n/(n+1)$, $n = 1, 2, \ldots$. Тогда по теореме о мажорируемой сходимости мы имеем

$$\lim_{n\to\infty} \int\limits_0^{\infty} (\rho_n e^{it})^y dP_Q(y) = \int\limits_0^{\infty} e^{ity} dP_Q(y),$$

и равенство $\psi_{Q_1}(\rho e^{it}) = \psi_{Q_2}(\rho e^{it})$ оказывается справедливым и при $\rho = 1$. Но $\psi_Q(e^{it})$ – не что иное как характеристическая функция случайной величины Q. Поэтому равенство $\psi_{Q_1}(e^{it}) = \psi_{Q_2}(e^{it})$, $t \in \mathbb{R}$, означает совпадение характеристических функций случайных величин Q_1 и Q_2. Следовательно, $Q_1 \stackrel{d}{=} Q_2$. Мы пришли к противоречию с изначальным предположением о том, что эти распределения различны, и тем самым доказали следующее утверждение.

Теорема 4.1. *Семейство смесей (4.1) функций распределения $G(x,\cdot)$ из аддитивно замкнутого множества идентифицируемо.*

Смеси, определяемые ядрами из аддитивно замкнутых семейств, конечно же, не исчерпывают все случаи идентифицируемых смесей. Рассмотрим масштабные смеси распределений, сосредоточенных на положительной полуоси.

Теорема 4.2. *Пусть $G(x,y) = G(xy)$, $y \geq 0$, $G(0) = 0$. Предположим, что преобразование Фурье функции $G^*(y) = G(e^y)$, $y \geq 0$, не обращается в нуль. Тогда семейство смесей*

$$\mathcal{F} = \{F_Q(x) = \mathsf{E}G(xQ), x \geq 0 : \mathsf{P}(Q > 0) = 1\}$$

идентифицируемо.

Доказательство. Вводя переменные u и v вместо x и y, полагая $x = e^u$, $y = e^{-v}$ и обозначая $F_Q^*(u) = F_Q(e^u)$, $P_Q^*(v) = \mathsf{P}(Q \geq e^{-v})$, запишем смесь $F_Q(x)$ в виде свертки функций G^* и P_Q^*:

$$F_Q(x) = F_Q(e^u) = F_Q^*(u) = \int\limits_{-\infty}^{\infty} G^*(u-v) dP_Q^*(v).$$

Но преобразование Фурье-Стилтьеса $\psi_{F_Q^*}$ свертки F_Q^* равно произведению преобразований Фурье-Стилтьеса $\psi_{G^*}(t)$ и $\psi_{P_Q^*}(t)$ функций G^* и P_Q^* соответственно:

$$\psi_{F_Q^*}(t) = \psi_{G^*}\psi_{P_Q^*}(t).$$

Поэтому, если $\psi_{G^*}(t)$ ни в одной точке не обращается в нуль, то равенство $G^* * P_{Q_1}^* = G^* * P_{Q_2}^*$ влечет $Q_1 \stackrel{d}{=} Q_2$. Теорема доказана.

Некоторые семейства сдвиговых смесей обладают аналогичным свойством.

Теорема 4.3. *Семейство сдвиговых смесей*

$$\mathcal{F} = \{F_Q(x) = \mathsf{E}G(x - Q), x \in \mathbb{R} : Q \in \mathcal{Q}\},$$

где $\mathcal{Q}$ – множество всех вещественных случайных величин, идентифицируемо, если характеристическая функция, соответствующая функции распределения $G(x)$, нигде не обращается в нуль.

Доказательство. Достаточно записать условие $F_Q(x) = \mathsf{E}G(x - Q)$ в терминах характеристических функций.

Теорема 4.1 обеспечивает идентифицируемость масштабных смесей нормальных законов с нулевым средним, равно как и других симметричных строго устойчивых распределений.

Теорема 4.3 обеспечивает идентифицируемость сдвиговых смесей любого безгранично делимого распределения.

Условия идентифицируемости однопараметрических смесей, представленные в Теоремах 4.1-4.3, были доказаны Тейчером (Teicher, 1961). Мы ограничились упоминанием тех идентифицируемых семейств, котрые будут рассмотрены ниже в связи с изучением условий сходимости некоторых специальных случайных сумм. Многочисленные примеры других идентифицируемых семейств можно найти в работах (Teicher, 1961), (Teicher, 1963), (Yakowitz and Spragins, 1968).

Некоторые необходимые и достаточные условия идентифицируемости семейств смесей доказаны в работе (Tallis, 1969), см. также обзор (Круглов, 1991).

К сожалению, в общем случае примеры идентифицируемых семейств двухпараметрических сдвиг-масштабных смесей неизвестны. Такие примеры известны только в том случае, когда Q – двумерный дискретный случайный вектор, принимающий лишь конечное число значений. В этом случае идентифицируемым оказывается конечная сдвиг-масштабная смесь нормальных законов (Teicher, 1963).

Покажем, что семейство произвольных сдвиг-масштабных смесей нормальных законов не является идентифицируемым. Пусть Y_1, Y_2, U_1 и U_2 – независимые случайные величины такие, что Y_1 и Y_2 имеют одно и то же стандартное нормальное распределение, $\mathsf{P}(U_1 = 1) = 1$, а случайная величина U_2 невырождена, $\mathsf{P}(U_2 > 0) = 1$. Тогда распределение случайной величины $Z = Y_1U_2 + Y_2$ можно записать как в виде

$$\mathsf{P}(Z < x) = \int\limits_0^\infty \int\limits_{-\infty}^\infty \Phi\left(\frac{x-v}{u}\right) d\mathsf{P}(Y_2 < v) d\mathsf{P}(U_2 < u), \tag{4.4}$$

так и в виде

$$\mathsf{P}(Z < x) = \int\limits_0^\infty \int\limits_{-\infty}^\infty \Phi\left(\frac{x-v}{u}\right) d\mathsf{P}(Y_1U_2 < v) d\mathsf{P}(U_1 < u). \tag{4.5}$$

Как видно, смешивающие распределения в (4.4) и (4.5) различны.

5 Сходимос ь распределений случайных сумм к идентиф ицируемым смесям. Центральная предельна я теорема и закон больших чисел для случа йных сумм

В этом разделе мы пр одолжим исследование условий слабой сходимости "нарастающих" случайных сумм в предположениии, что суммы неслучайного числа слагаемых, будучи надлежащим образом центрированы и нормированы, слабо сходятся к некоторой случайной вели чине Y с функцией распределения $H(x) = \mathsf{P}(Y < x)$. Если при этом семейства сдвиго вых или масштабных смесей функции распределения H оказываются идентифициру емыми, то условия сходимости приобретают особенно простой вид.

Сначала рассмотр им условия сходимости распределений случайных сумм к идентифицируемым сдвиго вым смесям. Как мы видели в разделах 1 и 2, подобные смеси появляются в качест ве предельных законов, если $b_{N_k}/d_k \Rightarrow b$ для некоторого $b > 0$ при $k \to \infty$. Без огра ничения общности, будем считать, что $b = 1$.

Теорема 5.1. *Пр едположим, что $b_k \to \infty$, $d_k \to \infty$, $b_{N_k}/d_k \Rightarrow 1$ и при некоторой последовательности в ещественных чисел $\{\alpha_j\}_{j\ge1}$ имеет место сходимость*

$$\frac{1}{b_k}\sum_{j=1}^k (X_j - \alpha_j) \Rightarrow Y \qquad (k \to \infty), \tag{5.1}$$

к некоторой случайно й величине Y. Пусть, более того, слагаемые $\{X_j\}$ удовлетворяют условию равноме рного асимптотического постоянства: для любого $\varepsilon > 0$

$$\lim_{k\to\infty} \max_{1\le j\le k} \mathsf{P}(|X_j - \alpha_j| > \varepsilon b_k) = 0. \tag{5.2}$$

Сходимость случайны х сумм

$$\frac{1}{d_k}\left(\sum_{j=1}^{N_k}(X_j - \alpha_j) - c_k\right) \Rightarrow Z \qquad (k \to \infty)$$

к некоторой случайн ой величине Z имеет место с некоторой последовательностью вещественных чисел $\{c_k\}_{k\ge1}$ тогда и только тогда, когда существует случайная величина V, удовлетворя ющая условиям

1) $Z \stackrel{d}{=} Y + V$, где Y и V независимы;

2) $\displaystyle\frac{1}{d_k}\left(\sum_{j=1}^{N_k}\alpha_j - c_k\right) \Rightarrow V \qquad (k \to \infty).$

Доказательство. Сведем доказательство к Теореме 3.4. Заметим, что условие (3.3), фигурирующее в формулировке Теоремы 3.4, использовалось только для того, чтобы доказать слабую компактность последовательности $\{b_{N_k}/d_k\}$. Поэтому в рассматриваемой ситуации оно излишне, так как слабая компактность упомянутой последоватльности является следствием ее сходимости к единице. Далее, условие (5.2) означает, что распределение случайной величины Y, будучи предельным для сумм независимых равномерно предельно малых слагаемых, безгранично делимо. Но характеристические функции безгранично делимых законов не обращаются в нуль и потому по Теореме 4.3 семейство сдвиговых смесей функции распределения $H(x) = \mathsf{P}(Y < x)$ идентифицируемо. Поэтому в рассматриваемой ситуации множество $\mathcal{W}_1(Z|Y) = \{V : Z \stackrel{d}{=} Y + V, Y$ и V независимы$\}$ сожержит не более одного элемента. Теперь требуемое утверждение следует из Теоремы 3.4, в которой множество $\mathcal{W}(Z|Y)$ следует заменить на множество $\mathcal{W}_1(Z|Y)$ в силу условия $b_{N_k}/d_k \Rightarrow 1$. Теорема доказана.

Теперь обратимся к изучению условий слабой сходимости распределений "нарастающих" случайных сумм к идентифицируемым масштабным смесям. Мы ограничимся лишь рассмотрением случая, когда предельные распределения являются масштабными смесями нормальных законов с нулевым средним. На то есть несколько причин. Во-первых, условия сходимости к масштабным смесям других распределений можно получить из приводимых ниже результатов с помощью простого переобозначения. Во-вторых, класс масштабных смесей нормальных законов с нулевым средним весьма богат и содержит, в частности, симметризованное гамма-распределение (как мы видели в разделе 3), в том числе распределение Лапласа (двойное экспоненциальное), распределения Коши, Стьюдента, симметричные строго устойчивые законы (см., например, (Золотарев, 1985), Теорема 3.3.1). В третьих, такие распределения имеют широкое применение в метрологии и финансовой математике (см. раздел 7).

Следуя традиции, стандартную нормальную функцию распределения мы будем обозначать $\Phi(x)$, $x \in \mathbb{R}$.

Теорема 5.2. *Предположим, что имеют место (3.3) и (5.1) при некоторых b_k и α_k, $b_k \to \infty$ $(k \to \infty)$, причем $\mathsf{P}(Y < x) = \Phi(x)$, $x \in \mathbb{R}$. Сходимость случайных сумм*

$$\frac{1}{d_k}\sum_{j=1}^{N_k}(X_j - \alpha_j) \Rightarrow Z \qquad (k \to \infty) \tag{5.3}$$

к некоторой случайной величине Z имеет место при некоторой последовательности положительныъх чисел $\{d_k\}_{k\geq 1}$, $d_k \to \infty$ $(k \to \infty)$, тогда и только тогда, когда существует неотрицательная случайная величина U, удовлетворяющая следующим условиям:

1) $\mathsf{P}(Z < x) = \mathsf{E}\Phi\left(\dfrac{x}{U}\right), \quad x \in \mathbb{R}$;

2) $b_{N_k}/d_k \Rightarrow U \quad (k \to \infty)$.

Доказательство. В силу идентифицируемости семейства масштабных смесей нормальных законов множество $\mathcal{S}(Z|Y)$, введенное перед Теоремой 3.5, содержит не более одного элемента. Поэтому требуемый результат немедленно следует из Теоремы 3.5 с $a_k = \alpha_1 + \ldots + \alpha_k$, $k \geq 1$. Теорема доказана.

Важность этой теоремы для приложений обусловлена тем, что она является естественным обобщением центральной предельной теоремы на суммы со случайным числом слагаемых. Она во многих случаях может объяснить отклонения наблюдаемых распределений от нормального. Во многих прикладных задачах, устанавливаемая классическими рассуждениями нормальность наблюдений является скорее исключением, нежели правилом. Это случается, например, в теории измерений, где как правило, распределения ошибок предполагаются нормальными, но в действительности таковыми не являются (см., например, (Новицкий и Зограф, 1991)). Теорема 5.2 может объяснить ненормальность распределения погрешностей. Действительно, когда для обоснования нормальности погрешностей применяется центральная предельная теорема, то рассуждают так. На результат каждого измерения оказывает влияние большое число случайных факторов, ни один из которых не является доминирующим. Погрешность определяется суммарным воздействием этих факторов и потому, согласно центральной предельной теореме, должна иметь нормальное распределение. Однако в подобных рассуждениях есть один существенный изъян: дело в том, что на *разные* измерения воздействует, вообще говоря *разное* число факторов, то есть, число случайных факторов, определяющих погрешность, само является случайным фактором. Поэтому вместо классической предельной теоремы здесь более уместно пользоваться, скажем, Теоремой 5.2, согласно которой распределение погрешностей номально тогда и только тогда, когда число факторов, влияющих на измерения, асимптотически неслучайно.

Из Теоремы 5.2 можно получить еще одно следствие, очень важное с методической точки зрения.

Рассмотрим последовательность независимых необязательно одинаково распределенных случайных величин (с.в.) $\{X_i\}_{i\geq 1}$. Будем считать, что $\mathsf{E}X_i = 0$, $\mathsf{E}X_i^2 = \sigma_i^2 < \infty$, $i \geq 1$. Более того, предположим, что с.в. $\{X_i\}_{i\geq 1}$ удовлетворяют условию Линдеберга: для любого $\tau > 0$

$$\lim_{n\to\infty}\frac{1}{B_n^2}\sum_{i=1}^{n}\int_{|x|>\tau B_n} x^2 dF_i(x) = 0,$$

где $B_n^2 = \sigma_1^2 + \ldots + \sigma_n^2$, $F_i(x) = \mathsf{P}(X_i < x)$. Пусть $\{N_n\}_{n\geq 1}$ – последовательность целочисленных положительных с.в. таких, что при каждом n с.в. N_n и $\{X_i\}_{i\geq 1}$ независимы. Для натурального k обозначим $S_k = X_1 + \ldots + X_k$. Хорошо известно, что в сделанных выше предположениях

$$\lim_{n\to\infty}\mathsf{P}\left(\frac{S_n}{B_n} < x\right) = \Phi(x)$$

равномерно по x. Покажем, что несмотря на такое "хорошее" поведение слагаемых, гарантирующее очень легкие (нормальные) хвосты предельного закона для сумм S_n, случайные суммы S_{N_n} за счет "плохого" поведения индекса N_n могут иметь предельные распределения с тяжелыми хвостами, например, устойчивые. Тем самым мы попытаемся поколебать устоявшееся представление, бытующее среди специалистов-прикладников, согласно которому тяжелые хвосты закона, предельного для сумм независимых случайных величин (присущие, например, устойчивым распределениям), непременно обусловлены наличием тяжелых хвостов у слагаемых (например, отсутствием у них моментов порядка меньше второго). Это представление основано на хорошо известных результатах (см., например, (Гнеденко и Колмогоров, 1949), гл. 7. раздел 35, Теорема 2), согласно которым сходимость распределений сумм неслучайного числа независимых одинаково распределенных слагаемых к устойчивым законам возможна тогда и только тогда, когда слагаемые имеют в определенном смысле тяжелые хвосты. Конечно же, мы не будем подвергать сомнению справедливость этих классических результатов. Мы просто обращаем внимание на то, что при решении прикладных задач (особенно в финансовой и актуарной математике, где в последнее время устойчивые законы находят весьма широкое применение) очень важен правильный выбор структурной модели. Из приводимых ниже результатов вытекает, что тяжелые хвосты распределения, предельного для сумм, могут возникать за счет случайности числа слагаемых в сумме. Мы покажем, что случайные суммы независимых слагаемых с описанными выше свойствами асимптотически строго устойчивы тогда и только тогда, когда асимптотически строго устойчивы их индексы.

Пусть $G_{\alpha,\theta}(x)$ – строго устойчивая функция распределения с характеристическим показателем α и параметром θ, которой соответствует характеристическая функция

$$g_{\alpha,\theta}(t)=\exp\left\{-|t|^{\alpha}\exp\left\{-i\frac{\pi\theta\alpha}{2}\mathrm{sign}t\right\}\right\},$$

где $t\in\mathbb{R}$, $0<\alpha\le 2$, $|\theta|\le\theta_\alpha=\min(1,2/\alpha-1)$. Пусть $\{D_n\}_{n\ge1}$ – последовательность положительных чисел такая, что $D_n\to\infty$ при $n\to\infty$.

Теорема 5.3. *Предположим, что $N_n\to\infty$ по вероятности при $n\to\infty$. Для того чтобы*

$$\lim_{n\to\infty}\mathsf{P}\left(\frac{S_{N_n}}{D_n}<x\right)=G_{\alpha,0}(x),\quad x\in\mathbb{R},\tag{5.4}$$

необходимо и достаточно, чтобы

$$\lim_{n\to\infty}\mathsf{P}\left(\frac{B^2_{N_n}}{D^2_n}<x\right)=G_{\alpha/2,1}(x),\quad x\in\mathbb{R}.$$

Доказательство состоит в последовательном применении Теоремы 5.2, согласно которой для того чтобы в сделанных предположениях

$$\mathsf{P}\left(\frac{S_{N_n}}{D_n}<x\right)\Longrightarrow\mathsf{P}(Z<x),\quad n\to\infty,$$

(где Z – некоторая с.в.) необходимо и достаточно, чтобы существовала с.в. $U\ge0$ такая, что, во-первых,

$$\mathsf{P}(Z<x)=\int_0^\infty\Phi(x/u)d\mathsf{P}(U<u),\quad x\in\mathbb{R},$$

и во-вторых,

$$\frac{B_{N_n}}{D_n}\Longrightarrow U,\quad n\to\infty,$$

и Теоремы 3.3.1 из (Золотарев, 1985), согласно которой

$$G_{\alpha,0}(x)=\int_0^\infty\Phi(x/\sqrt{u})dG_{\alpha/2,1}(u),\quad x\in\mathbb{R},$$

с использованием свойства идентифицируемости масштабных смесей нормальных законов и абсолютной непрерывности устойчивых распределений.

Следствие 5.1. *Если в дополнение к условиям Теоремы слагаемые $\{X_j\}_{j\ge1}$ одинаково распределены, то для того чтобы имела место сходимость (5.4), необходимо и достаточно, чтобы*

$$\lim_{n\to\infty}\mathsf{P}\left(\frac{N_n}{D_n^2}<x\right)=G_{\alpha/2,1}(\sigma_1^2x),\quad x\in\mathbb{R}.$$

В случае, когда случайная величина Z в (5.3) сама имеет нормальное распределение, утверждение Теоремы 5.2 можно существенно усилить.

Теорема 5.4. *Предположим, что имеют место (3.3) и (5.1) с некоторыми последовательностями положительных чисел $\{b_k\}$, $b_k\to\infty$ $(k\to\infty)$, и вещественных чисел $\{\alpha_j\}$. Пусть, помимо этого, слагаемые $\{X_j\}$ равномерно асимптотически постоянны. Сходимость*

$$\mathsf{P}\left(\frac{1}{d_k}\sum_{j=1}^{N_k}(X_j-\alpha_j)<x\right)\Rightarrow\Phi(x)\qquad(k\to\infty)\tag{5.5}$$

имеет место при некоторой последовательности положительных чисел $\{d_k\}$, $d_k\to\infty$ $(k\to\infty)$, тогда и только тогда, когда для некоторого $b>0$ выполнены следующие условия:

1) $\mathsf{P}(Y<x)=\Phi(bx)$;

2) $b_{N_k}/d_k\Rightarrow b\qquad(k\to\infty)$.

Доказательство. Необходимость. Условие равномерного асимптотического постоянства слагаемых (5.2) и условие (5.1) влекут безграничную делимость функции распределения случайной величины Y, фигурирующей в (5.1). Так как слагаемые центрированы, вследствие (5.5) любая случайная величина $U\in\mathcal{S}(Z|Y)$ в данном случае должна удовлетворять условию

$$\mathsf{E}h(tU)=e^{-t^2/2},\qquad t\in\mathbb{R},\tag{5.6}$$

где h – безгранично делимая характеристическая функция случайной величины Y. Согласно Лемме 1.2.1 из (Круглов и Королев, 1990) из (5.6) вытекает, что должны существовать функции $a(u)$ и $\sigma^2(u)$ такие, что

$$h(tU) = \exp\{ita(u) - \frac{1}{2}\sigma^2(u)t^2\}, \qquad t \in \mathbb{R},\ u > 0, \tag{5.7}$$

и более того, $\mathsf{E}a(U) = 0$. Специфический вид зависимости аргумента функции в левой части (5.7) от t и u означает, что $a(u) = \alpha u$, где $\alpha \in \mathbb{R}$, и $\sigma(u) = \beta u$, где $\beta \geq 0$. Поэтому мы имеем $\alpha \mathsf{E}U = 0$, что в силу неотрицательности случайной величины U возможно или тогда, когда $\mathsf{P}(U = 0) = 1$ (но тогда нарушается (5.7)), или когда $\alpha = 0$. Таким образом, соотношение (5.7) принимает вид

$$\mathsf{E}h(tU) = \mathsf{E}(\exp\{-\frac{1}{2}t^2\beta^2\})^{U^2} = e^{-t^2/2}, \qquad t \in \mathbb{R}.$$

По Теореме 3 в (Kruglov and Titov, 1986) это возможно лишь тогда, когда случайная U величина вырождена в ненулевой точке. Поэтому любая случайная величина $U \in \mathcal{S}(Z|Y)$ при некотором $b > 0$ удовлетворяет соотношению $\mathsf{P}(U = b) = 1$, и в дополнение к этому функция распределения случайной величины Y равна $\Phi(bx)$. Теперь осталось только сослаться на Теорему 3.5.

Достаточность условий 1) и 2) вытекает из Теоремы 1.2. Теорема доказана.

Наряду с обобщениями центральной предельной теоремы, представленных Теоремами 5.2-5.4, Теорема 3.5 также позволяет получить следующие обобщения закона больших чисел на случайные суммы.

Теорема 5.5. *Предположим, что имеет место (3.3). Пусть сходимость*

$$\frac{1}{b_k}\sum_{j=1}^{k}(X_j - \alpha_j) \Rightarrow 1 \qquad (k \to \infty).$$

имеет место с некоторыми последовательностями вещественных чисел $\{\alpha_k\}_{k\geq1}$ и положительных чисел $\{b_k\}_{k\geq1}$, $b_k \to \infty$ $(k \to \infty)$. Сходимость случайных сумм

$$\frac{1}{d_k}\sum_{j=1}^{N_k}(X_j - \alpha_j) \Rightarrow Z \qquad (k \to \infty)$$

к некоторой случайной величине Z имеет место с некоторой последовательностью положительных чисел $\{d_k\}_{k\geq1}$, $d_k \to \infty$ $(k \to \infty)$, тогда и только тогда, когда

$$\frac{b_{N_k}}{d_k} \Rightarrow Z \qquad (k \to \infty).$$

Доказательство. В рассматриваемом случае мы имеем $h(t) = \mathsf{E}\exp\{itY\} = e^{it}$. Поэтому

$$\mathsf{E}e^{itZ} = \mathsf{E}h(tU) = \mathsf{E}e^{itU}, \qquad t \in \mathbb{R},$$

что возможно только лишь, если $Z \stackrel{d}{=} U$, то есть семейство масштабных смесей вырожденного не в нуле закона тривиальным образом идентифицируемо. Теперь требуемое вытекает из Теоремы 3.5. Теорема доказана.

В свою очередь, из Теоремы 5.5 вытекает следующее уточнение знаменитой теоремы Реньи-Модьороди (Rényi, 1956), (Mogyoródi, 1971). Эта теорема играет важную роль в теории надежности.

Пусть $X_1, X_2, \ldots$ – независимые одинаково распределенные случайные величины с $\mathsf{E}X_1 = m \neq 0$. Для определенности, предположим, что $m > 0$. Как и ранее, пусть положительные случайные величины N_k независимы от последовательности $\{X_j\}_{j\geq1}$ при каждом $k \geq 1$.

Теорема 5.6. *Пусть имеет место (3.3) и $F(x)$ – произвольная функция распределения. Тогда*

$$\mathsf{P}\left(\frac{1}{k}\sum_{j=1}^{N_k}X_j < mx\right) \Rightarrow F(x) \qquad (k \to \infty),$$

если и только если

$$\mathsf{P}(N_k < kx) \Rightarrow F(x) \qquad (k \to \infty).$$

Доказательство этого результата сводится к Теореме 5.5, так как случайные величины $\{X_j\}_{j\geq1}$, обладающие указанными свойствами, очевидным образом удовлетворяют закону больших чисел.

Теорема 5.6 представляет собой закон больших чисел для случайных сумм. обратим внимание на одну удивительную особенность этого результата: предел для "средних арифметических" случайного числа независимых одинаково распределенных случайных величин, в отличие от классической ситуации, может быть случайным и полностью определяется предельным поведением случайного числа слагаемых.

6 Предельные теоремы для процессов риска

Рассмотрим случайный процесс

$$S(t) = S_0 + ct - \sum_{i=1}^{N(t)} X_i, \qquad t \geq 0, \tag{6.1}$$

где $N(t)$, $t \geq 0$, – целочисленный неотрицательный случайный процесс с неубывающими траекториями, $\{X_j\}_{j\geq 1}$ – независимые случайные величины, причем случайная величина $N(t)$ и последовательность $\{X_j\}_{j\geq 1}$ независимы при каждом $t \geq 0$, $c > 0$, $S_0 \geq 0$ (по определению, $\sum_{i=1}^{0} X_j = 0$).

Процесс (6.1) мы будем называть процессом риска. Если $N(t)$ – однородный пуассоновский процесс, а случайные величины $\{X_j\}_{j\geq 1}$ одинаково распределены и имеют конечные дисперсии, то процесс (6.1) называется классическим процессом риска.

Процессы вида (6.1) играют важную роль в актуарной математике. Они описывают поведение капитала страховой компании в зависимости от времени t. Такая интерпретация становится ясной, если мы предположим, что в (6.1) S_0 – это начальный капитал страховой компании, c – положительный коэффициент, характеризующий линейный рост капитала страховой компании благодаря страховым взносам клиентов, а случайные величины $\{X_j\}$ положительны и описывают выплаты страховой компании по страховым случаям, число которых до момента времени t равно $N(t)$. Конечно, эта модель не является всеобъемлющей и не учитывает, например, нелинейный рост капитала, вызванный его вложением в прибыльные проекты, выплату дивидентов держателям акций или полисов компании и инфляцию. Тем не менее, процессы вида (6.1) представляют собой довольно неплохое приближение к описанию реальной ситуации и могут рассматриваться как основа для построения более детализированных моделей различных аспектов деятельности страховых компаний.

Литература по актуарной математике, в которой описываются свойства процессов вида (6.1), обширна. Мы упомянем лишь канонические руководства (Bowers et al., 1986) и (Grandell, 1990), где можно найти дальнейшие ссылки.

Хорошо известно, что классический процесс риска асимптотически нормален при $t \to \infty$. Мы приведем доказательство этого факта, основанное на применении Теоремы 2.1. Для простоты, без потери общности, вместо процесса $S(t)$ с непрерывным временем рассмотрим процесс с дискретным временем S_n, $n = 0, 1, \ldots$, полагая $S_n = S(n)$. Это предположение хорошо согласуется с практикой, так как время обычно измеряется дискретными единицами: сутками, часами, минутами, и совсем уж трудно представить себе реальную ситуацию, когда страховая компания фиксирует моменты выплат по страховым случаям с точностью до секунд. Аналогично, $N_n = N(n)$.

Итак, вначале предположим, что случайные величины $\{X_j\}_{j\geq 1}$ одинаково распределены с $\mathsf{E}X_1 = a, \mathsf{D}X_1 = \sigma^2 < \infty$. Пусть $N(t)$ – однородный пуассоновский процесс с

интенсивностью $\lambda > 0$.

Теорема 6.1. *Классический процесс риска асимптотически нормален: для любого $x \in \mathbb{R}$*

$$\lim_{n\to\infty} \mathsf{P}\left(\frac{S_n - n(c - a\lambda) - S_0}{\sqrt{n\lambda(a^2+\sigma^2)}} < x\right) = \Phi(x).$$

Доказательство. Поскольку

$$S_n - n(c - a\lambda) - S_0 = -\left(\sum_{j=1}^{N_n} X_j - na\lambda\right), \tag{6.2}$$

мы можем свести доказательство к Теореме 2.1. В нашем случае

$$\mathsf{E}\sum_{j=1}^{N_n} X_j = an\lambda, \quad \mathsf{D}\sum_{j=1}^{N_n} X_j = n\lambda(a^2+\sigma^2).$$

Положим $a_n = na$, $c_n = na\lambda$, $b_n = \sigma\sqrt{n}$, $d_n = \sqrt{n\lambda(\sigma^2+a^2)}$. Тогда

$$\frac{a_{N_n} - c_n}{d_n} = \frac{a(N_n - n\lambda)}{\sqrt{n\lambda(a^2+\sigma^2)}} = \frac{a}{\sqrt{\sigma^2+a^2}} \cdot \frac{N_n - n\lambda}{\sqrt{n\lambda}} \Rightarrow V \ (n\to\infty), \tag{6.3}$$

где $\mathsf{P}(V < x) = \Phi\left(x\sqrt{\frac{a^2+\sigma^2}{a^2}}\right)$, $x \in \mathbb{R}$, согласно хорошо известному свойству распределения Пуассона:

$$\lim_{n\to\infty} \mathsf{P}\left(\frac{N_n - n\lambda}{\sqrt{n\lambda}} < x\right) = \Phi(x), \quad x \in \mathbb{R}.$$

Далее,

$$\frac{b_{N_n}}{d_n} = \sqrt{\frac{N_n}{n\lambda} \cdot \frac{\sigma^2}{a^2+\sigma^2}}.$$

Для произвольного $\varepsilon > 0$ мы имеем

$$\mathsf{P}\left(\left|\frac{N_n}{n\lambda} - 1\right| > \varepsilon\right) \leq \frac{\mathsf{D}N_n}{\varepsilon^2 n^2\lambda^2} = \frac{1}{\varepsilon^2 n\lambda} \to 0$$

при $n \to \infty$, так что

$$\frac{b_{N_n}}{d_n} \Rightarrow \sqrt{\frac{\sigma^2}{a^2+\sigma^2}} = U, \qquad (n\to\infty). \tag{6.4}$$

Наконец, в силу центральной предельной теоремы

$$\frac{1}{b_n}\left(\sum_{j=1}^{n} X_j - a_n\right) = \frac{1}{\sigma\sqrt{n}}\left(\sum_{j=1}^{n} X_j - na\right) \Rightarrow Y \quad (n\to\infty), \tag{6.5}$$

где $\mathsf{P}(Y < x) = \Phi(x)$, $x \in \mathbb{R}$. Поэтому, применяя Теорему 2.1 к случайным величинам

$$Z_n = \frac{\sum_{j=1}^{N_n} X_j - na\lambda}{\sqrt{n\lambda(a^2+\sigma^2)}}$$

с учетом (6.3), (6.4) и (6.5), имея в виду вид полученных нами распределений предельных случайных величин Y, U и V, мы получим

$$\lim_{n\to\infty} \mathsf{P}\left(\frac{\sum_{j=1}^{N_n} X_i - na\lambda}{\sqrt{n\lambda(a^2+\sigma^2)}} < x\right) = \mathsf{P}(YU + V < x) =$$
$$\Phi\left(x\sqrt{\frac{a^2+\sigma^2}{\sigma^2}}\right) * \Phi\left(x\sqrt{\frac{a^2+\sigma^2}{a^2}}\right) = \Phi(x), \quad x \in \mathbb{R}.$$

Заметим, что предельная случайная величина U вырождена, а вырожденная случайная величина независима от любой другой. Поэтому вместо условия слабой сходимости совместных распределений пар, фигурирующего в Теореме 2.1, мы можем ограничиться условиями сходимости маргинальных распределений. Таким образом, принимая во внимание (6.2), мы будем иметь

$$\lim_{n\to\infty} \mathsf{P}\left(\frac{S_n - n(c - a\lambda) - S_0}{\sqrt{n\lambda(a^2+\sigma^2)}} < x\right) =$$
$$\lim_{n\to\infty} \mathsf{P}\left(\frac{\sum_{j=1}^{N_n} X_j - na\lambda}{\sqrt{n\lambda(a^2+\sigma^2)}} > -x\right) =$$
$$1 - \Phi(-x) = \Phi(x), \qquad x \in \mathbb{R},$$

что и требовалось доказать.

Анализ реальных ситуаций показывает, что довольно сильные предположения, определяющие классический процесс риска, на практике можно считать выполненными далеко не всегда. В связи с этим возникают два вопроса. Во-первых, какие распределения могут выступать в качестве предельных для процессов вида (6.1) при ослаблении условий, определяющих классический процесс риска? Во-вторых, насколько можно ослабить эти условия, сохранив адекватность нормальной аппроксимации? Другими словами, в каких случаях можно пользоваться нормальным приближением при ослаблении условий, определяющих классический процесс риска? Остальная часть данного раздела посвящена ответам на эти вопросы.

Мы будем рассматривать ситуацию, которую можно считать обобщением той, в которой страховые требования $\{X_j\}_{j\geq 1}$ предполагаются одинаково распределенными, хотя формально на случайные величины $\{X_j\}$ мы не будем накладывать никаких (в том числе моментных) условий кроме их независимости (за исключением особо оговоренных случаев). Наши дополнительные условия будут связаны с центрирующими и нормирующими константами.

Предположим, что нормирующие постоянные имеют специальный вид, а именно, пусть

$$b_n = d_n = n^{1/\alpha} B(n),$$

где $1 < \alpha \leq 2$, а $B(x)$, $x \in \mathbb{R}$, – медленно меняющаяся функция, то есть такая, что для любого $p > 0$

$$\lim_{x\to\infty} \frac{B(px)}{B(x)} = 1. \tag{6.6}$$

В отношении центрирующих постоянных мы предположим, что $a_n = c_n$; более того, пусть последовательность $\{a_n\}$ монотонно возрастает и

$$\lim_{n\to\infty} \frac{a_n}{n} = A \tag{6.7}$$

для некоторого $A \in (0, \infty)$. Обозначим $\alpha_j = a_j - a_{j-1}$, где для определенности $a_0 = 0$.

Теорема 6.2. *Предположим, что $N_k \xrightarrow{P} \infty$ при $k \to \infty$, страховые требования $\{X_j\}_{j\geq 1}$ равномерно асимптотически постоянны: для любого $\varepsilon > 0$*

$$\lim_{k\to\infty} \max_{1\leq j\leq k} \mathsf{P}(|X_j - \alpha_j| > \varepsilon b_k) = 0,$$

и их неслучайные суммы имеют некоторое предельное распределение:

$$\frac{1}{b_k}\left(\sum_{j=1}^{k} X_j - a_k\right) \Rightarrow Y \qquad (k \to \infty). \tag{6.8}$$

Процесс риска S_n имеет предельное распределение при $n \to \infty$:

$$\frac{1}{b_n}(S_n - S_0 - nc + a_n) \Rightarrow -Z \qquad (n \to \infty), \tag{6.9}$$

тогда и только тогда, когда существует случайная величина V такая, что

1) $Z \stackrel{d}{=} Y + V$, Y и V независимы;

2) $\dfrac{a_{N_k} - a_k}{b_k} \Rightarrow V \qquad (k \to \infty)$.

Доказательство. Необходимость. Поскольку

$$\frac{1}{b_n}(S_n - S_0 - nc + a_n) = -\frac{1}{b_n}\left(\sum_{j=1}^{N_n} X_j - a_n\right),$$

мы сведем доказательство к Теореме 5.1. Покажем, что форма центрирующих и нормирующих постоянных и условия теоремы гарантируют, что $b_{N_k}/b_k \Rightarrow 1$ $(k \to \infty)$.

Действительно, согласно Теореме 3.1 из условий (6.8) и (6.9) с учетом неограниченного стохастического роста N_k вытекает слабая компактность последовательности пар случайных величин $\{(b_{N_k}/b_k, (a_{N_k} - a_k)/b_k)\}_{k\geq 1}$. Как и ранее, точную нижнюю грань q-квантилей случайной величины N_k будем обозначать $l_k(q)$. Поскольку последовательность $\{a_k\}$ монотонно возрастает, q-квантиль случайной величины $(a_{N_k} - a_k)/b_k$, которую мы обозначим $A_k(q)$, равна $(a_{l_k(q)} - a_k)/b_k$. В силу слабой компактности последовательности случайных величин $\{(a_{N_k} - a_k)/b_k\}_{k\geq 1}$ последовательность $\{A_k(q)\}_{k\geq 1}$ равномерно ограничена при каждом $q \in (0,1)$:

$$\sup_k |A_k(q)| \equiv M(q) < \infty.$$

В этом легко убедиться с помощью рассуждений от противного. Поэтому в силу определения b_n и условия (6.7) при каждом $q \in (0,1)$ мы имеем

$$\left|\frac{a_{l_k(q)}}{a_k} - 1\right| = \left|\frac{a_{l_k(q)} - a_k}{b_k}\right| \cdot \frac{b_k}{a_k} = |A_k(q)| \cdot \frac{b_k}{a_k} \leq$$
$$M(q)\frac{k^{1/\alpha}B(k)}{Ak} \cdot \frac{Ak}{a_k} \to 0 \qquad (k \to \infty), \tag{6.10}$$

так как $Ak/a_k \to 1$ согласно условию (6.7) и $k^{1/\alpha}B(k)/k \to 0$ по свойству медленно меняющихся функций. В свою очередь, из (6.10) и (6.7) мы получим, что

$$\frac{l_k(q)}{k} = \frac{Al_k(q)}{a_{l_k(q)}} \cdot \frac{a_{l_k(q)}}{a_k} \cdot \frac{a_k}{Ak} \to 1 \qquad (k \to \infty) \tag{6.11}$$

при каждом $q \in (0,1)$, поскольку неограниченное стохастическое возрастание N_k при $k \to \infty$ означает, что $l_k(q) \to \infty$ $(k \to \infty)$ для любого $q \in (0,1)$. Из (6.11) следует, что $N_k/k \Rightarrow 1$ $(k \to \infty)$. В силу неравенства

$$\mathsf{P}(|W_1 + W_2| > \delta) \leq \mathsf{P}\Big(|W_1| > \frac{\delta}{2}\Big) + \mathsf{P}\Big(|W_2| > \frac{\delta}{2}\Big),$$

справедливого для любых случайных величин W_1 и W_2 и для любого $\delta > 0$, при произвольном $\varepsilon > 0$ мы имеем

$$\mathsf{P}\left(\left|\frac{b_{N_k}}{b_k} - 1\right| > \varepsilon\right) = \mathsf{P}\left(\left|\left(\frac{N_k}{k}\right)^{1/\alpha}\frac{B(N_k)}{B(k)} - 1\right| > \varepsilon\right) =$$
$$\mathsf{P}\left(\left|\left(\frac{N_k}{k}\right)^{1/\alpha}\left(\frac{B(N_k)}{B(k)} - 1\right) + \left(\frac{N_k}{k}\right)^{1/\alpha} - 1\right| > \varepsilon\right) \leq$$
$$\mathsf{P}\left(\left|\left(\frac{N_k}{k}\right)^{1/\alpha}\left(\frac{B(N_k)}{B(k)} - 1\right)\right| > \frac{\varepsilon}{2}\right) + \mathsf{P}\left(\left|\left(\frac{N_k}{k}\right)^{1/\alpha} - 1\right| > \frac{\varepsilon}{2}\right). \tag{6.12}$$

Рассмотрим первое слагаемое в правой части (6.12). Имеет место следующая цепочка неравенств:

$$\mathsf{P}\left(\left|\left(\frac{N_k}{k}\right)^{1/\alpha}\left(\frac{B(N_k)}{B(k)} - 1\right)\right| > \frac{\varepsilon}{2}\right) =$$
$$\sum_{n=1}^{\infty}\mathsf{P}(N_k = n)\mathsf{P}\left(\left|\frac{B(n)}{B(k)} - 1\right|\frac{\varepsilon k^{1/\alpha}}{2n^{1/\alpha}}\right) =$$
$$\sum_{n:\left|\frac{n}{k}-1\right|\leq\frac{1}{2}}\mathsf{P}(N_k = n)\mathsf{P}\left(\left|\frac{B(n)}{B(k)} - 1\right| > \frac{\varepsilon k^{1/\alpha}}{2n^{1/\alpha}}\right) +$$
$$\sum_{n:\left|\frac{n}{k}-1\right|>\frac{1}{2}}\mathsf{P}(N_k = n)\mathsf{P}\left(\left|\frac{B(n)}{B(k)} - 1\right| > \frac{\varepsilon k^{1/\alpha}}{2n^{1/\alpha}}\right) \leq$$
$$\sum_{n:\left|\frac{n}{k}-1\right|\leq\frac{1}{2}}\mathsf{P}(N_k = n)\mathsf{P}\left(\left|\frac{B(n)}{B(k)} - 1\right| > \frac{2^{1/\alpha-1}\varepsilon}{3^{1/\alpha}}\right) + \mathsf{P}\left(\left|\frac{N_k}{k} - 1\right| > \frac{1}{2}\right) \leq$$
$$\sum_{n:\left|\frac{n}{k}-1\right|\leq\frac{1}{2}}\mathsf{P}(N_k = n)\mathsf{P}\left(\sup_{\frac{1}{2}\leq p\leq\frac{3}{2}}\left|\frac{B(kp)}{B(k)} - 1\right| > \frac{2^{1/\alpha-1}\varepsilon}{3^{1/\alpha}}\right) + \mathsf{P}\left(\left|\frac{N_k}{k} - 1\right|\frac{1}{2}\right) \leq$$
$$\mathsf{P}\left(\sup_{\frac{1}{2}\leq p\leq\frac{3}{2}}\left|\frac{B(kp)}{B(k)} - 1\right| > \frac{2^{1/\alpha-1}\varepsilon}{3^{1/\alpha}}\right) + \mathsf{P}\left(\left|\frac{N_k}{k} - 1\right| > \frac{1}{2}\right). \tag{6.13}$$

Согласно Теореме 1.1 в (Сенета, 1985), сходимость (6.6) равномерна на каждом замкнутом интервале значений p. Поэтому существует $k_0 = k_0(\varepsilon)$ такое, что для всех $k \geq k_0$

$$\sup_{\frac{1}{2}\leq p\leq\frac{3}{2}}\left|\frac{B(kp)}{B(k)} - 1\right| < \frac{2^{1/\alpha-1}\varepsilon}{3^{1/\alpha}}. \tag{6.14}$$

Таким образом, в соответствии с (6.13) и (6.14) при всех $k \geq k_0$ мы имеем

$$\mathsf{P}\left(\left|\left(\frac{N_k}{k}\right)\left(\frac{B(N_k)}{B(k)} - 1\right)\right| > \frac{\varepsilon}{2}\right) \leq \mathsf{P}\left(\left|\frac{N_k}{k} - 1\right| > \frac{1}{2}\right),$$

и потому согласно уже доказанному,

$$\lim_{k\to\infty}\mathsf{P}\left(\left|\left(\frac{N_k}{k}\right)^{1/\alpha}\left(\frac{B(N_k)}{B(k)} - 1\right)\right| > \frac{\varepsilon}{2}\right) \leq \lim_{k\to\infty}\mathsf{P}\left(\left|\frac{N_k}{k} - 1\right| > \frac{1}{2}\right) = 0. \tag{6.15}$$

Рассмотрим второе слагаемое в правой части (6.12). Поскольку $N_k/k \Rightarrow 1$ $(k \to \infty)$, $N_k/k \xrightarrow{P} 1$ и потому $(N_k/k)^{1/\alpha} \xrightarrow{P} 1$, то есть

$$\lim_{k\to\infty}\mathsf{P}\left(\left|\left(\frac{N_k}{k}\right)^{1/\alpha} - 1\right| > \frac{\varepsilon}{2}\right) = 0. \tag{6.16}$$

Теперь требуемое соотношение $b_{N_k}/b_k \Rightarrow 1$ $(k \to \infty)$ вытекает из (6.12), (6.15). Ссылка на Теорему 5.1 завершает доказательство необходимости.

Достаточность. Как и при доказательстве необходимости, убеждаемся, что условие 2) влечет $b_{N_k}/b_k \Rightarrow 1$ $(k \to \infty)$, что позволяет свести доказательство к Теореме 2.1. Теорема доказана.

Следствие 6.1. *Предположим, что $N_k \xrightarrow{P} \infty$ и неслучайные суммы страховых требований имеют некоторое предельное распределение, то есть имеет место (6.8). Процесс риска асимптотически нормален*

$$\mathsf{P}\left(\frac{1}{b_k}(S_k - S_0 - kc + a_k) < x\right) \Rightarrow \Phi(x) \qquad (k \to \infty) \tag{6.17}$$

тогда и только тогда, когда существуют числа $\mu \in \mathbb{R}$ и $0 \leq \sigma^2 \leq 1$ такие, что

1) $\mathsf{P}(Y < x) = \Phi\left(\dfrac{x-\mu}{\sigma}\right), \quad x \in \mathbb{R};$

2) $\mathsf{P}\left(\dfrac{a_{N_k} - a_k}{b_k} < x\right) \Rightarrow \Phi\left(\dfrac{x+\mu}{\sqrt{1-\sigma^2}}\right) \quad (k \to \infty).$

Доказательство. По Теореме 6.2 предельная стандартная нормальная случайная величина Z должна удовлетворять соотношению $Z \stackrel{d}{=} Y + V$ с независимыми Y и V. Тогда по теореме Леви-Крамера о разложимости нормального закона только на нормальные компоненты и Y, и V должны иметь нормальное распределение, как известно, являющееся безгранично делимым. Поэтому условие равномерного предельного постоянства страховых требований в рассматриваемом случае излишне. Требуемый результат теперь следует из Теоремы 6.2. Следствие доказано.

В Теореме 6.2 и Следствии 6.1 мы предполагали, что свойства страховых требований обеспечивают слабую сходимость распределений их неслучайных сумм к некоторому закону, а условия сходимости распределений процесса риска формулировались в терминах процесса $N(t)$. Теперь мы будем предполагать, что известны асимптотические свойства процесса $N(t)$ при $t \to \infty$, и сформулируем условия сходимости распределений процесса риска в терминах требований.

Теорема 6.3. *Предположим, что $N_k \xrightarrow{P} \infty$ при $k \to \infty$, имеет место сходимость*

$$\frac{a_{N_k} - a_k}{b_k} \Rightarrow V \qquad (k \to \infty) \tag{6.18}$$

к некоторой случайной величине V, а последовательность случайных величин $\{Y_k\}_{k \geq 1}$,

$$Y_k = \frac{1}{b_k}\left(\sum_{j=1}^{k} X_j - a_k\right), \qquad k \geq 1,$$

слабо компактна. Соотношение (6.9) выполнено для процесса риска S_n тогда и только тогда, когда существует слабо компактная последовательность случайных величин $\{Y_k'\}_{k \geq 1}$ таких, что

1) $Z \stackrel{d}{=} V + Y_k'$ *при каждом k, где V и Y_k' независимы;*

2) $L_1(Y_k, Y_k') \to 0 \qquad (k \to \infty).$

Доказательство. Так же, как и в доказательстве Теоремы 6.2, мы убеждаемся, что условие (6.18) влечет $b_{N_k}/b_k \Rightarrow 1$ при $k \to \infty$, то есть случайные величины U_k' в тройках (Y_k', U_k', V_k'), фигурирующих в Теореме 3.1, должны быть равными единице с вероятностью единица. Теперь требуемое утверждение вытекает из Теоремы 3.1. Теорема доказана.

Следствие 6.2. *Пусть в дополнение к условиям Теоремы 6.3 семейство сдвиговых смесей функции распределения $\mathsf{P}(V < x)$, $x \in \mathbb{R}$, идентифицируемо. Тогда (6.9) имеет место тогда и только тогда, когда существует случайная величина Y такая, что*

1) $Z \stackrel{d}{=} Y + V$, *где Y и V независимы;*

2) $Y_k \Rightarrow Y \quad (k \to \infty).$

Доказательство. В силу условия идентифицируемости семейства сдвиговых смесей функции распределения случайной величины V соотношению $Z \stackrel{d}{=} Y + V$, в правой части которого слагаемые независимы, удовлетворяет не более чем одна случайная величина. Ссылка на Теорему 6.3 завершает доказательство.

Следствие 6.3. *В условиях Теоремы 6.3 процесс риска асимптотически нормален, то есть выполняется соотношение (6.17), тогда и только тогда, когда существуют числа $\mu \in \mathbb{R}$ и $0 \leq \sigma^2 \leq 1$ такие, что*

1) $\mathsf{P}(V < x) = \Phi\left(\dfrac{x-\mu}{\sigma}\right), \quad x \in \mathbb{R};$

2) $\mathsf{P}(Y_k < x) \Rightarrow \Phi\left(\dfrac{x+\mu}{\sqrt{1-\sigma^2}}\right) \quad (k \to \infty).$

Доказательство. В силу Теоремы 6.3 предельная случайная величина Z должна допускать представление в виде $Z \stackrel{d}{=} Y_k' + V$, где Y_k' и V независимы, $k \geq 1$. Но вследствие нормальности Z, каким бы ни было $k \geq 1$, согласно теореме Леви-Крамера представление $Z \stackrel{d}{=} Y_k' + V$ с независимыми слагаемыми в правой части возможно только лишь, если и Y_k', и V имеют нормальные распределения. Более того, поскольку семейство сдвиговых смесей нормальной функции распределения случайной величины V идентифицируемо, случайная величина Y_k' не зависит от k. Следствие доказано.

Наконец, рассмотрим условия сходимости распределений процессов риска без каких бы то ни было предположений о сходимости процесса $N(t)$ или неслучайных сумм страховых требований.

Теорема 6.4. *Предположим, что $N_k \xrightarrow{P} \infty$ при $k \to \infty$ и последовательность $\{Y_k\}_{k\geq 1}$ слабо компактна. Процесс риска S_n слабо сходится (6.9) к некоторой случайной величине Z тогда и только тогда, когда существует слабо компактные последовательности случайных величин $\{Y_k'\}_{k\geq 1}$ и $\{V_k'\}_{k\geq 1}$ такие, что*

1) $Z \stackrel{d}{=} Y_k' + V_k'$, *где Y_k' и V_k' независимы, $k \geq 1$;*

2) $L_1(Y_k, Y_k') \to 0 \quad (k \to \infty)$;

3) $L_1\left(\dfrac{a_{N_k} - a_k}{b_k}, V_k'\right) \to 0 \quad (k \to \infty)$.

Доказательство этой теоремы сводится к Теореме 3.1 с учетом соотношения $b_{N_k}/b_k \Rightarrow 1$ $(k \to \infty)$, установленного в ходе доказательства Теоремы 6.2. Теорема доказана.

Следствие 6.4. *В условиях Теоремы 6.4 процесс риска асимптотически нормален (6.17) тогда и только тогда, когда существуют числовые последовательности $\{\beta_k\}_{k\geq 1}$ и $\{\sigma_k^2\}_{k\geq 1}$ такие, что $\sup\limits_k |\beta_k| < \infty$, $0 \leq \sigma_k^2 \leq 1$, $k \geq 1$, и*

1) $L_1\left(\mathsf{P}(Y_k < x), \Phi\left(\dfrac{x - \beta_k}{\sigma_k}\right)\right) \to 0 \quad (k \to \infty)$;

2) $L_1\left(\mathsf{P}\left(\dfrac{a_{N_k} - a_k}{b_k} < x\right), \Phi\left(\dfrac{x + \beta_k}{\sqrt{1 - \sigma_k^2}}\right)\right) \to 0 \quad (k \to \infty)$.

Доказательство. В силу теоремы Леви-Крамера о разложимости нормального закона только на нормальные компоненты, каждая из случайных величин Y_k' и V_k', фигурирующих в условии 1) Теоремы 6.4 должна иметь нормальное распределение. Более того, условия $\sup\limits_k |\beta_k| < \infty$ и $0 \leq \sigma^2 \leq 1$ необходимы и достаточны для слабой компактности последовательностей $\{Y_k'\}$ и $\{V_k'\}$. Теперь требуемый результат вытекает из Теоремы 6.4. Следствие доказано.

В качестве иллюстрации рассмотрим ситуацию, когда считающий процесс $N(t)$, то есть число страховых случаев, зафиксированных до момента t, является дважды стохастическим пуассоновским процессом (процессом Кокса), получающимся из однородного пуассоновского процесса с помощью случайной замены времени. А именно, пусть $N_1(t)$ – однородный пуассоновский процесс с единичной интенсивностью, а $\Lambda(t), t \geq 0$, – независимый от $N_1(t)$ случайный процесс, обладающий следующими свойствами: $\Lambda(0) = 0$, $\mathsf{P}(\Lambda(t) < \infty) = 1$ при каждом $t > 0$, траектории процесса $\Lambda(t)$ не убывают и непрерывны справа. Тогда дважды стохастический пуассоновский процесс (процесс Кокса) $N(t)$ определяется как

$$N(t) = N_1(\Lambda(t)).$$

В этом случае мы будем говорить, что процесс Кокса $N(t)$ управляется процессом $\Lambda(t)$. Процессы Кокса являются очень удобными и достаточно гибкими моделями, используемыми в страховой математике, поскольку с их помощью можно учитывать реальное непостоянство интенсивности потока страховых случаев, вызываемую, например, сезонными или политическими факторами. Для натуральных n обозначим $\Lambda_n = \Lambda(n)$.

Теорема 6.5. *Предположим, что страховые требования $\{X_j\}_{j\geq 1}$ независимы и одинаково распределены с $\mathsf{E}X_j = a \neq 0$ и $\mathsf{D}X_j = \sigma^2 < \infty$. Пусть $N(t)$ – процесс Кокса, управляемый процессом $\Lambda(t)$, независимый от последовательности $\{X_j\}_{j\geq 1}$. Процесс риска S_n имеет некоторое предельное распределение:*

$$\frac{1}{\sigma\sqrt{n}}(S_n - S_0 - n(c - a)) \Rightarrow -Z \qquad (k \to \infty), \tag{6.19}$$

тогда и только тогда, когда существует случайная величина V такая, что

1) $Z \stackrel{d}{=} \sqrt{\dfrac{a^2 + \sigma^2}{\sigma^2}}W + \dfrac{a}{\sigma}V$, *где W и V независимы, а W имеет стандартное нормальное распределение;*

2) $\dfrac{\Lambda_n - n}{\sqrt{n}} \Rightarrow V \quad (n \to \infty)$.

Доказательство. Применим Теорему 6.2 с $a_n = na_n$, $b_n = \sigma\sqrt{n}$. Согласно этой теореме, (6.19) имеет место тогда и только тогда, когда

$$\frac{a(N_n - n)}{\sigma\sqrt{n}} \Rightarrow V_0 \qquad (n \to \infty), \tag{6.20}$$

где V_0 – случайная величина такая, что $Z \stackrel{d}{=} W_1 + V_0$, W_1 и V_0 независимы и $\mathsf{P}(W_1 < x) = \Phi(x)$, $x \in \mathbb{R}$. Можно показать (см., например, (Gnedenko and Korolev, 1996)), что (6.20) имеет место тогда и только тогда, когда существует случайная величина V такая, что $V_0 \stackrel{d}{=} \frac{a}{\sigma}(W_2 + V)$, где W_2 и V независимы, $\mathsf{P}(W_2 < x) = \Phi(x)$, $x \in \mathbb{R}$, и выполнено 2). Таким образом, с учетом нормальности и независимости случайных величин W_1 и W_2 мы приходим к представлению 1). Теорема доказана.

Следствие 6.5. *В условиях Теоремы 6.5 процесс риска асимптотически нормален*

$$\mathsf{P}\left(\frac{1}{\sigma\sqrt{n}}(S_n - S_0 - n(c - a)) < x\right) \Rightarrow \Phi\left(\frac{x - \mu}{\delta}\right) \qquad (n \to \infty)$$

с некоторыми $\mu \in \mathbb{R}$ и $\delta^2 < \infty$ тогда и только тогда, когда $\delta^2 \geq 1$ и

$$\mathsf{P}\left(\frac{\Lambda_n - n}{\sqrt{n}} < x\right) \Rightarrow \Phi\left(\frac{ax - \sigma\mu}{\sigma\sqrt{\delta^2 - 1}}\right) \qquad (n \to \infty).$$

Это утверждение вытекает из Теоремы 6.5 и теоремы Леви-Крамера о разложимости нормального закона только на нормальные компоненты.

Таким образом, ответы на сформулированные выше вопросы выглядят так.

Предельные распределения для процессов риска при указанном специальном выборе центрирующих и нормирующих постоянных имеют вид свертки двух распределений, одно из которых является предельным для неслучайных сумм страховых требований, а другое является предельным для нормализованного числа страховых случаев.

Нормальная аппроксимация для распределения процесса риска адекватна тогда и только тогда, когда асимптотически нормальны и неслучайные суммы страховых требований, и нормализованные количества страховых случаев.

Мы не использовали свойство положительности случайных величин $\{X_j\}$, которое является следствием их интерпретации как страховых выплат. Поэтому все утверждения данного раздела – это по сути предельные теоремы для "нарастающих" случайных сумм при специальном выборе центрирующих и нормирующих постоянных.

Выбор нормирующих постоянных в виде $b_n = n^{1/\alpha}B(n)$, где $B(x)$ – медленно меняющаяся функция, типичен для ситуации, в которой слагаемые $\{X_j\}$ независимы и одинаково распределены, и обеспечивает сходимость распределения нормированных неслучайных сумм к устойчивому закону с показателем α. Поэтому, формально не используя условие совпадения распределений слагаемых, мы, тем не менее, рассматривали слагаемые, которые "почти одинаково" распределены в указанном смысле.

Мы использовали такие центрирующие и нормирующие постоянные, чтобы вместо сходимости пар случайно индексированных центрирующих и нормирующих постоянных иметь дело только с одной компонентой каждой из этих пар, характеризующей сдвиги, в то время как другая компонента, характеризующая преобразование масштаба, была асимптотически вырожденной. Эти условия позволили нам получить более сильные и одновременно более простые утверждения. Если мы будем рассматривать процессы риска в общей ситуации, без каких бы то ни было условий на центрирующие и нормирующие постоянные, то предельные теоремы для процессов риска с точностью до терминологии будут совпадать с результатами, представленными в разделах 1-3. К сожалению, в общей ситуации не удается получить простых заключений об условиях адекватности нормальной аппроксимации для распределений процессов риска из-за того, что, как мы убедились в разделе 4, двухпараметрические сдвиг-масштабные смеси нормальных законов неидентифицируемы.

7 Некоторые модели финансовой математики

В данном разделе мы рассмотрим проблему построения моделей для распределений приращений процессов биржевых цен. Она привлекала, привлекает и, несомненно, еще долго будет привлекать внимание многих математиков. К настоящему времени предложено много разнообразных моделей. Мы, к счастью, избавлены от необходимости приводить полное описание существующих подходов к построению моделей эволюции финансовых индексов, поскольку имеется превосходный обзор (Ширяев, 1995). Поэтому мы кратко упомянем наиболее существенные результаты и остановим свое внимание на моделях, представимых в виде смесей нормальных законов, поскольку, в отличие от иных моделей, они не только обеспечивают хорошее согласие с реальными данными, но и внутренне непротиворечивы в том смысле, что в рамках этих моделей при помощи предельных теорем для случайных сумм можно сравнительно легко перекинуть мостик между вполне реалистичными структурными микромоделями, описывающими поведение биржевых цен на очень коротких интервалах времени, и широко используемыми в качестве макромоделей динамики биржевых цен винеровскими процессами со случайными сносом и диффузией.

Для простоты мы будем говорить о ценах акций, хотя наши рассуждения могут относиться к любому объекту торговли на бирже. Пусть $P(t)$ — цена акции в момент времени t. Поскольку невозможно абсолютно точно предсказать поведение цены акции, естественно считать, что $P(t)$, $t \geq 0$, — это случайный процесс. Изучение реальных данных показывает, что приращения этого процесса $P(t+\tau_1) - P(t)$ и $P(t) - P(t-\tau_2)$ можно считать некоррелированными для всех t и всех τ_1, τ_2, превосходящих некоторое малое положительное число. Также вполне естественно предположение о том, что $\mathsf{E}(P(t+\tau) - P(t)) = 0$ для всех t и упомянутых τ. Эти предположения могут быть подтверждены следующими нехитрыми рассуждениями: если бы средние приращения процесса биржевой цены были систематически ненулевыми, то внимательные биржевые игроки, сразу заметив это, попытались бы на этом сыграть, одновременно тем самым способствуя возврату ситуации в стационарный режим. (Последнее предположение обуславливает рассмотрений мартингалов в качестве динамических моделей процессов биржевых цен. Поскольку мы будем рассматривать в некотором смысле статические модели, мы обойдемся без привлечения результатов теории мартингалов).

История рассматриваемой задачи тесно связана с развитием теории вероятностей и весьма примечательна. Одну из первых серьезных математических моделей для процесса биржевых цен предложил Л. Башелье в своем основополагающем труде "Теория спекуляции", опубликованном в 1900 году (Bachelier, 1900). Отталкиваясь от предположений некоррелированности и стационарности приращений, он пришел к той же конструкции, на основе которой позже (в 1905 году) А. Эйнштейн предложил модель для процесса броуновского движения (Einstein, 1905), а именно, к одномерному уравнению диффузии. Как известно, подобная конструкция приводит к использованию винеровских процессов в качестве моделей для броуновского движения частицы и, стало быть, для процессов цен акций.

Приращения винеровского процесса имеют нормальное распределение. Винеровские процессы могут считаться в определенном смысле предельными моделями и тесно связаны с центральной предельной теоремой, применение которой к анализу распределений приращений биржевых цен сводится к следующим рассуждениям. Деля интервал времени $[t, t+\tau]$ на некоторое число подынтервалов, мы замечаем, что приращение $P(t+\tau) - P(t)$ может быть представлено в виде суммы приращений процесса $P(t)$ на подынтервалах. В предположении, что приращения на подынтервалах независимы (а отмеченная выше некоррелированность приращений делает это предположение весьма разумным), в соответствии с центральной предельной теоремой распределение приращения $P(t+\tau) - P(t)$ приближается к нормальному при увеличении числа подынтервалов, если выполнено условие Линдеберга. Это условие означает, что ни одно из слагаемых не играет доминирующей роли в сумме. Если подынтервалы имеют одинаковую длину, то естественно считать, что распределения приращений на подынтервалах одинаковы. В этом случае условие Линдеберга сводится к требованию конечности дисперсий элементарных приращений.

Однако, статистический анализ реальных данных показывает, что распределения приращений процессов биржевых цен на интервалах времени сравнительно небольшой длины (до 2–3 недель) отличны от нормальных. Первые работы, в которых отмечено это явление, появились еще в 1915 году. Результаты очень серьезного статистического анализа, подтверждающего отличие упомянутых распределений от нормального, были опубликованы М. Кендаллом в 1953 году (Kendall, 1953). Отмеченный феномен является всеобщим: ненормальность распределений приращений проявляется на всех биржах независимо от объекта торговли. Поэтому винеровские процессы оказались отнюдь не бесспорными моделями динамики биржевых цен.

Отмеченная ненормальность распределений приращений проявляется в том, что в действительности наблюдается заметно больше очень больших и очень маленьких по абсолютной величине значений приращений, нежели их должно быть в соответствии с нормальным распределением. Другими словами, наблюдаемые распределения приращений биржевых цен на интервалах времени умеренной длины являются более островершинными, нежели нормальные, имея заметно более тяжелые хвосты.

Важная попытка объяснить замеченную островершинность этих распределений была предпринята в середине 60-х годов. Обычно эту попытку ассоциируют с исследованиями Б. Мандельброта (Mandelbrot, 1963), (Mandelbrot, 1963) и Э. Фамы (Fama, 1965). Одна из основных идей в их рассуждениях может быть кратко изложена так: отклонение распределений приращений от нормального означает, что классическая центральная предельная теорема здесь неприменима, так как нарушаются ее условия. Если не выходить за рамки общей структурной модели, согласно которой итоговое приращение процесса $P(t)$ на каждом интервале является суммой независимых приращений этого же процесса на подынтервалах одинаковой длины, и предполагать, что элементарные приращения имеют одинаковые распределения,

то единственным условием, которое должно нарушаться, остается требование конечности дисперсии элементарных приращений. Поэтому Б. Мандельброт и Э. Фама предложили вместо классической центральной предельной теоремы использовать предельные теоремы о сходимости сумм независимых слагаемых с бесконечной дисперсией в качестве основы для построения моделей для распределений приращений биржевых цен.

Как известно, в таких теоремах в качестве предельных распределений выступают устойчивые законы. Исследования Б. Мандельброта и Э. Фамы вызвали взрыв интереса как к аналитическим свойствам устойчивых законов (см., например, монографию (Золотарев, 1985) так и к статистическим процедурам, основанным на устойчивых семействах. Однако наряду с большим числом благоприятных аналитических качеств, устойчивые законы обладают одним неприятным свойством: за исключением четырех случаев (нормальное распределение, распределения Коши, Леви и симметричное к распределению Леви), отсутствуют явные выражения устойчивых плотностей (которые всегда существуют) в терминах элементарных функций. Ненамного больше известно устойчивых плотностей, выражающихся через наиболее широко используемые специальные функции. Это обстоятельство затрудняет статистический анализ в рамках устойчивых моделей.

Однако отмеченный недостаток (кстати, практически несущественный при компьютерном анализе данных) не является единственным обстоятельством, заставляющим с некоторым сомнением относиться к целесообразности использования устойчивых законов для моделирования распределений приращений биржевых цен. Напомним, что основным предположением, обусловливающим их применение в подобных задачах, является отсутствие конечных дисперсий у приращений биржевых цен за интервалы времени бесконечно малой длины. Но для отсутствия дисперсии необходимо, чтобы упомянутые элементарные приращения могли с положительной вероятностью принимать как угодно большие по абсолютной величине значения, что на практике невозможно. К тому же по свидетельству практиков, большинство тестов отвергают гипотезу о согласии распределений реальных данных с устойчивыми законами, поскольку, все же, слишком больших по абсолютной величине наблюдений в реальных данных меньше, чем должно быть при устойчивом распределении.

Многие исследователи обращали внимание на то, что реальная интенсивность биржевых торгов крайне изменчива и предпринимали попытки объяснить островершинность распределений приращений биржевых цен, основываясь на этой изменчивости. По-видимому, первым, кто систематически исследовал неоднородность операционного времени на биржах, был П. Кларк (Clark, 1970), (Clark, 1973). Вместо обычных винеровских процессов П. Кларк предложил в качестве моделей динамики биржевых цен использовать подчиненные винеровские процессы со случайным временем вида $W(X(t))$, где $W(t)$, $t \geq 0$ — винеровский процесс, а $X(t)$ — процесс с неубывающими траекториями, начинающимися в нуле.

По Кларку, при использовании центральной предельной теоремы в качестве базы при построении моделей для распределений приращений биржевых цен из-за неоднородности времени нарушается не конечность дисперсий, а другое структурное предположение, именно, предположение о том, что число слагаемых в сумме неслучайно.

Отталкиваясь от догадки Кларка мы сейчас покажем, как построить более реалистичную и довольно простую модель процесса биржевых цен на интервалах времени малой длины (так сказать, элементарную модель на микроуровне), а затем перейти к моделям типа $W(X(t))$ на макроуровне с помощью нескольких версий центральной предельной теоремы для случайных сумм. Эти модели, по сути, являются удобными аппроксимациями для реальных процессов.

Мы начнем с простейшей модели. Рассмотрим процесс цены акции на интервале времени малой длины. Наша основная посылка состоит в том, что формальное выражение цены акции существует только в заключенном на бирже контракте (сделке). Поэтому мы предполагаем, что на интервалах времени между последовательными заключениями контрактов цена акции остается постоянной и равной цене, указанной в последнем заключенном до рассматриваемого момента времени контракте. Тогда изменение цены акции может быть описано маркированным точечным процессом $\{(T_i, P_i)\}_{i\geq 1}$, где T_i — момент заключения i-го контракта и P_i — цена акции, указанная в i-ом контракте (контракты нумеруются в хронологическом порядке; контракты, заключенные одновременно, нумеруются в произвольном порядке). Таким образом, $P(t) = P_i$ при $T_i \leq t < T_{i+1}$. Пусть $N(t)$ — число контрактов, заключенных к моменту t (мы считаем, что отсчет начинается с $t = 0$). Тогда

$$P(t) - P(0) = \sum_{j=1}^{N(t)} (P_j - P_{j-1}), \quad t \geq 0, \quad P_0 = P(0). \tag{7.1}$$

Обозначим $X_j = P_j - P_{j-1}, j \geq 1$. Предположим, что $\{X_j\}_{j\geq 1}$ — независимые (необязательно одинаково распределенные) случайные величины (с.в.). Мы будем изучать поведение распределения с.в. $P(t) - P(0)$ при неограниченном возрастании $N(t)$. Отметим, что мы не предполагаем, что обязательно $t \to \infty$. Интервал $[0, t]$ (пока) остается фиксированным. Для удобства построения аппроксимаций формально рассмотрим последовательность целочисленных неотрицательных случайных величин $\{N_k\}_{k\geq 1}$ и предположим, что при каждом $k \geq 1$ с.в. N_k и $\{X_j\}_{j\geq 1}$ независимы.

Из Леммы 1.1 мы легко получаем следующий результат. Пусть $f_k(s)$ — характеристическая функция (х.ф.) с.в. $P(t) - P(0)$, в которой $N(t)$ заменено на N_k:

$$f_k(s) = \mathsf{E}\exp\left\{is\sum_{j=1}^{N_k} X_j\right\}, \quad s \in \mathbb{R}.$$

Теорема 7.1. *Пусть для некоторых $b_k > 0$, $b_k \to \infty$ $(k \to \infty)$ имеет место сходимость*

$$\mathsf{P}\left(\frac{1}{b_k}\sum_{j=1}^{k} X_j < x\right) \Rightarrow \Phi(x) \qquad (k \to \infty). \tag{7.2}$$

Предположим, что при некоторых $d_k > 0$, $d_k \to \infty$ $(k \to \infty)$ семейство случайных величин $\{b_{N_k}/d_k\}$ слабо компактно. Тогда при каждом $s \in \mathbb{R}$

$$\lim_{k\to\infty}\left| f_k\left(\frac{s}{d_k}\right) - \int_0^\infty \exp\left\{-\frac{1}{2}x^2s^2\right\} dA_k(x)\right| = 0,$$

где $A_k(x) = \mathsf{P}(b_{N_k} < xd_k)$.

Легко видеть, что, если мы обозначим $U_k = b_{N_k}/d_k$, то

$$\int_0^\infty \exp\{-\frac{1}{2}x^2s^2\}dA_k(x) = \mathsf{E}\exp\{is(XU_k)\}, \qquad s \in \mathbb{R}$$

где $\mathsf{P}(X < x) = \Phi(x)$; более того, X и U_k независимы при каждом $k \geq 1$. Таким образом, с помощью Теоремы 1 мы приходим к следующему заключению: если число контрактов, заключенных в течение интервала времени $[0, t]$, довольно велико и последовательные изменения цены акции от контракта к контракту удовлетворяют условиям центральной предельной теоремы, то для некоторой неотрицательной с.в. U справедливо приближенное равенство

$$\mathsf{P}(P(t) - P(0) < x) \approx \mathsf{P}(XU < x) = \mathsf{E}\Phi\left(\frac{x}{U}\right), \qquad x \in \mathbb{R}, \tag{7.3}$$

где $\mathsf{P}(X < x) = \Phi(x)$ и с.в. X и U независимы. Другими словами, распределения, аппроксимирующие распределения приращений процессов цен акций следует искать среди масштабных смесей нормальных законов. Отметим, что если с.в. $X_1, X_2, \ldots$ одинаково распределены, то (7.2) эквивалентно конечности дисперсии X_1; при этом $b_k = \sqrt{k\mathsf{D}X_1}$.

Теорема 7.1 описывает *сближение* распределений приращений биржевых цен. Если воспользоваться утверждениями о *сходимости* распределений случайных сумм, например, Теоремой 5.2, то можно заметить, что тип смешивающего распределения в правой части (7.3) определяется предельным поведением с.в. b_{N_k}/d_k. Теорема 5.2 позволяет ставить вопрос о реконструкции распределения U_k по распределению U, так как согласно этой теореме, в предположении (7.2) условие $U_k \Rightarrow U$ не только достаточно, но и необходимо для возможности замены приближенного неравенства в (7.3) точным в пределе при $N_k \xrightarrow{P} \infty$ при надлежащей нормировке с.в. $P(t) - P(0)$. Справедливость сказанного мы проиллюстрируем одним важным примером, связанным с так называемыми процессами Кокса.

Рассмотрим последовательность $\{N^{(k)}(t)\}_{k\geq 1}$ дважды стохастических пуассоновских процессов (процессов Кокса), то есть $N^{(k)}(t) = N_1(\Lambda_k(t))$, где N_1 — однородный пуассоновский процесс с единичной интенсивностью и Λ_k — независимые от N_1 случайные процессы с неубывающими конечными непрерывными справа траекториями, исходящими из нуля. Пусть с.в. $\{X_j\}_{j\geq 1}$ независимы и одинаково распределены с $\mathsf{E}X_j = 0$ и $\mathsf{D}X_j = \sigma^2 < \infty$. Предположим, что при каждом $k \geq 1$ процесс $N^{(k)}(t)$ независим от последовательности $\{X_j\}_{j\geq 1}$.

Теорема 7.2. *Предположим, что $\Lambda_k(t) \xrightarrow{P} \infty$ и последовательность случайных величин $\{\Lambda_k(t)/d_k^2\}_{k\geq 1}$ слабо компактна при некоторых $d_k > 0$. Тогда сходимость*

$$\frac{1}{d_k}(P(t) - P(0)) \Rightarrow Z \qquad (k \to \infty)$$

к некоторой случайной величине Z имеет место тогда и только тогда, когда существует неотрицательная случайная величина U такая что

1) $$\mathsf{P}(Z < x) = \mathsf{E}\Phi\left(\frac{x}{U}\right), \qquad x \in \mathbb{R}; \tag{7.4}$$

2) $$\frac{\sigma\sqrt{\Lambda_k(t)}}{d_k} \Rightarrow U \qquad (k \to \infty). \tag{7.5}$$

Это утверждение можно доказать так же, как доказывалась Теорема 5.2 с учетом замечания, сделанного в самом конце раздела 2.

Теорема 7.2 служит очень наглядной иллюстрацией того, как реализуется принцип неоднородности биржевого времени, предложенный Кларком. В некотором смысле $\Lambda_k(t)$ — это среднее число контрактов, заключенных к моменту t. Если траектории процесса $\Lambda_k(t)$ дифференцируемы, то $\lambda_k(t) = \Lambda_k'(t)$ определяет интенсивность торгов в момент t. При интерпретации рассматриваемой модели может оказаться разумным отождествление функции $\lambda_k(t)$ с волатильностью акции в момент t.

Обсудим, насколько хорошо выводы, получаемые с помощью Теорем 7.1 и 7.2, согласуются с реальными данными. Изучение реальных данных показывает, что феномен ненормальности распределений приращений $P(t) - P(0)$ затухает с ростом t. При довольно больших t (начиная с 3–4 недель) эти распределения практически неотличимы от нормального. С помощью Теоремы 7.2 мы можем предложить следующее объяснение ненормальности распределений с.в. $P(t) - P(0)$ при умеренных t и "унормаливания" этих распределений с ростом t.

С одной стороны, если t достаточно велико, чтобы считать $\Lambda_k(t)$ большим, но одновременно достаточно мало́, чтобы считать разброс значений с.в. $\Lambda_k(t)/d_k^2$ на непересекающихся интервалах длины t большим, то естественно полагать, что распределение с.в. U в Теореме 7.2 отлично от вырожденного, то есть распределение с.в. Z ненормально.

С другой стороны, обратим внимание, что до сих пор вспомогательный параметр k никак не был формально связан со временем. Поэтому мы можем рассмотреть целочисленное неотрицательное время, положить $k = t$ и рассмотреть в Теореме 7.2 вместо последовательности процессов $\{\Lambda_k(t)\}_{k\geq 1}$ один процесс $\Lambda(t)$ и считать, что $\Lambda_k = \Lambda(k)$. В этом случае константы d_k, очевидно, зависят от времени. На практике флуктуации процесса $\Lambda(t)$ на интервалах большой длины относительно невелики в том смысле, что $\Lambda_k/k \xrightarrow{P} \mathrm{const}$ при $k \to \infty$. Полагая $d_k^2 = \sigma^2 k$, мы приходим к выводу, что с.в. U в (7.5) становится константой, то есть распределение нормализованного приращения $(P(t) - P(0))/(\sigma\sqrt{t})$ становится нормальным при $t \to \infty$.

Как мы уже отмечали в разделе 5, класс масштабных смесей нормальных законов $\mathsf{E}\Phi\left(\frac{x}{U}\right)$ очень богат и содержит, помимо само́го нормального распределения, например, распределения Коши, Стьюдента, Лапласа (двойное или двустороннее экспоненциальное распределение), симметричные устойчивые распределения. Имеется большое число публикаций, посвященных изучению конкретных смесей со специальными смешивающими законами. Помимо упоминавшихся выше работ П. Кларка заслуживают быть отмеченными исследования (Praetz, 1972) и (Blattberg and Gonedes, 1974) (где рассматривается распределение Стьюдента), (Brada and Van Tassel, 1966), (Press, 1967), (Grander and Morgenstern, 1970), (Boness, Chen and Jatusipitak, 1974), (Oldfield, Rogalski and Jarrow, 1977), (Cox and Ross, 1976), (Merton, 1973), (Kon, 1984) (где рассматриваются нормальные смеси с различными смешивающими распределениями). Однако основываясь на Теореме 7.2, можно высказать предположение, что, скорее всего, не существует одной универсальной с.в. U, точнее, одного разумным образом параметризованного семейства таких с.в., которое удовлетворительно описывает поведение всех акций на всех биржах. Действительно, условие (7.5) и комментарий сразу же после Теоремы 7.2 напрямую связывает с.в. U с изменением интенсивностей торгов или волатильностями, которые, разумеется, сильно зависят от акции, времени и места торгов. Тем не менее, для каждой конкретной акции и каждой конкретной биржи можно пробовать подобрать свою с.в. U (то есть ее распределение).

Со смесями $\mathsf{E}\Phi\left(\frac{x}{U}\right)$ связан один забавный, скорее психологический, парадокс. Как мы уже отмечали, истинные распределения приращений цены акции на умеренных интервалах всегда заметно более островершинны нежели нормальное распределение. В то же время, на первый взгляд, не совсем ясно, всегда ли смеси $\mathsf{E}\Phi\left(\frac{x}{U}\right)$, предлагаемые в качестве моделей для распределений рассматриваемых приращений, более островершинны, чем нормальные законы. Действительно,

$$\mathsf{E}\Phi\left(\frac{x}{U}\right) = \mathsf{P}(XU < x), \qquad x \in \mathbb{R},$$

где X и U независимы и $\mathsf{P}(X < x) = \Phi(x)$. Как известно, нормальные распределения с большими дисперсиями обладают плотностями с гладкими вершинами, в то время как вершины нормальных плотностей с малыми дисперсиями остры. Поэтому кажется, что, умножая нормальную с.в. X на с.в. U, значения которой концентрируются недалеко от нуля, мы можем получить более островершинную плотность, а умножая X на с.в. U со значениями, существенно превосходящими единицу, мы получим менее островершинную плотность. К счастью, эти рассуждения кажутся разумными лишь до тех пор, пока они не проверены вычислениями. В качестве численной характеристики островершинности рассмотрим коэффициент эксцесса $\text{æ}(Y)$, который для с.в. Y с $\mathsf{E}Y^4 < \infty$ определяется как

$$\text{æ}(Y) = \mathsf{E}\left(\frac{Y - \mathsf{E}Y}{\sqrt{\mathsf{D}Y}}\right)^4.$$

Если $\mathsf{P}(X < x) = \Phi(x)$, то $\text{æ}(X) = 3$. Плотности с более острыми вершинами (и, соответственно, более тяжелыми хвостами), чем у нормальной плотности, имеют $\text{æ} > 3$, а для плотностей с менее острой вершиной $\text{æ} < 3$. Следующее утверждение разрушает приведенные выше "аргументы", основанные на "здравом" смысле.

Лемма 7.1. *Пусть X и U — независимые случайные величины с конечными четвертыми моментами. Предположим, что $\mathsf{E}X = 0$ и $\mathsf{P}(U \geq 0) = 1$. Тогда*

$$æ(XU) \geq æ(X).$$

Более того, $æ(XU) = æ(X)$ *тогда и только тогда, когда* $\mathsf{P}(U = const) = 1$.

Доказательство. Из независимости X и U следует, что

$$æ(XU) = \mathsf{E}\left(\frac{XU - \mathsf{E}XU}{\sqrt{\mathsf{D}XU}}\right)^4 = \frac{\mathsf{E}(XU - \mathsf{E}XU)^4}{(\mathsf{E}(XU - \mathsf{E}XU)^2)^2} =$$

$$\frac{\mathsf{E}(XU - \mathsf{E}X\mathsf{E}U)^4}{(\mathsf{E}(XU - \mathsf{E}X\mathsf{E}U)^2)^2} = \frac{\mathsf{E}X^4\mathsf{E}U^4}{(\mathsf{E}X^2)^2(\mathsf{E}U^2)^2} = æ(X) \cdot \frac{\mathsf{E}U^4}{(\mathsf{E}U^2)^2}. \tag{7.6}$$

Но согласно неравенству Йенсена, $\mathsf{E}U^4 \geq (\mathsf{E}U^2)^2$. Поэтому правая часть соотношения (7.6) всегда не меньше, чем $æ(X)$. Более того, она равна $æ(X)$ тогда и только тогда, когда $\mathsf{E}U^4 = (\mathsf{E}U^2)^2$, что возможно лишь тогда, когда $\mathsf{P}(U = const) = 1$. Лемма доказана.

Таким образом, масштабные смеси нормальных законов всегда более островершинны, нежели само нормальное распределение, и следовательно, имеют более тяжелые хвосты.

Иногда есть основания считать, что слагаемые X_j в Теореме 7.1 не центрированы. Тем не менее, Теоремы 7.1 и 7.2 дают некоторую возможность анализировать тренды процессов биржевых цен. Трендовые индикаторы весьма популярны среди тех, кто играет на бирже. Пусть $M(t)$ — тренд (основная тенденция) цены акции в момент t. Отклонения цены акции от $M(t)$ в конкретных контрактах обусловлены намерениями конкретных сторон, вовлеченных в контракты, и потому могут считаться "более случайными", нежели $M(t)$. С формальной точки зрения это сводится к тому, что мы будем считать $M(t)$ неслучайной функцией. Пусть $X_j^* = P_j - P_{j-1}$ — приращение цены акции от $(j-1)$-го к j-му контракту и пусть T_j — момент заключения j-го контракта. Положим $\alpha_j^* = M(T_j) - M(T_{j-1})$. Тогда мы можем считать, что с.в. $X_j = X_j^* - \alpha_j^*$ центрированы, и с помощью Теорем 7.1 и 7.2 мы можем построить аппроксимации распределений центрированных приращений $Z(t) = (P(t) - M(t)) - (P(0) - M(0))$. В связи с этим мы получаем простой принцип проверки правильности подгонки тренда: если наблюдаемое распределение с.в. $Z(t)$ симметрично и островершинно, то тренд построен правильно.

Несимметричность распределения с.в. $Z(t)$ может также быть обусловлена неправильным выбором шкалы. Выше мы рассматривали поведение процесса $P(t)$. Однако, как отмечено в книге (Hull, 1989), ключевым моментом формирования цен акций и стратегии игры на бирже является то, что инвесторы главным образом заинтересованы в прибыли, выраженной в процентах от стоимости акций, и не зависящей от самой цены. В этой связи вместо процесса $P(t) - P(0)$ следует рассматривать процесс $\log(P(t)/P(0))$ (переход к логарифмической шкале приводит к рассмотрению так называемого геометрического броуновского движения, см. ниже). При этом

все рассуждения, приведенные выше, остаются справедливыми и для приращений такого процесса, которые, очевидно, по смыслу задачи понимаются не в аддитивном, а в мультипликативном смысле. Таким образом, Теоремы 7.1 и 7.2 также можно использовать не только для подгонки тренда, но и для проверки правильности выбора шкалы по симметричности и островершинности распределений приращений рассматриваемых процессов.

До сих пор мы рассматривали модели распределений приращений биржевых цен, полученные с помощью предельных теорем для сумм со случайным числом слагаемых в схеме "нарастающих" сумм, и предполагали, что элементарные приращения цены акции, то есть ее изменения от контракта к контракту, имеют нулевые ожидания. Это предположение и общие рассуждения, типичные для центральной предельной теоремы, привели нас к рассмотрению масштабных смесей нормальных законов в качестве искомых моделей. Такие смеси островершинны и симметричны. Однако распределения, наблюдаемые на практике без учета отклонений от тренда, о чем говорилось выше, не только островершинны, но, как правило, еще и асимметричны.

Переходя от рассмотрения сумм центрированных слагаемых к центрированным суммам (что можно интерпретировать как желание автоматически учесть, вообще говоря, ненулевой тренд), с помощью Теоремы 2.1 или Теоремы 3.1 в предположении асимптотической нормальности сумм неслучайного числа с.в. мы приходим к моделям распределений приращений биржевых цен, имеющих вид сдвиг-масштабных смесей нормальных законов

$$\mathsf{E}\Phi\left(\frac{x - V}{U}\right),$$

где с.в. U неотрицательна и с.в. U и V необязательно независимы.

Чтобы избежать появления только симметричных моделей, обусловленного приведенными выше результатами, мы обобщим нашу основную модель (7.1) и рассмотрим схему серий. В отличие от простейшей ситуации, рассмотренной выше, на разных интервалах времени приращения процесса цены акции могут иметь разный систематический снос (различные ненулевые средние значения), поскольку можно выделить интервалы, на которых цены довольно стабильно идут вверх или стабильно идут вниз. Поэтому мы можем предположить совпадение распределений элементарных приращений в пределах одного интервала стабильности, но совсем не обязательно предполагать, что эти распределения одинаковы в течение длинных интервалов времени, включающих много подынтервалов относительной стабильности. Подобные рассуждения естественно приводят нас к схеме серий.

Еще один довод в пользу привлечения схемы серий заключается в следующем. Пусть n — вспомогательный натуральный папраметр и пусть нас интересует суммарное приращение цены акции на интервале длины t. Пусть $T = nt$ — суммарная длина интервала времени, в течение которого проводятся наблюдения за приращениями на подынтервалах длины t, так что можно считать n объемом выборки, под которой мы понимаем набор доступных нам наблюдений за приращениями на подынтервалах длины t. Пусть N_n — число контрактов, заключенных в течение n-го подынтервала длины t, и пусть $X_{n,1}, \ldots, X_{n,N_n}$ — элементарные приращения цены акции в течение последнего, n-го, интервала. Пусть теперь T увеличивается, что означает увеличение n, если t остается постоянным. При этом последний, n-ый подынтервал движется направо вдоль числовой оси. Несмотря на то, что мы будем предполагать, что при каждом $n \geq 1$ с.в. $\{X_{n,j}\}_{j\geq 1}$ одинаково распределены, распределения с.в. $X_{n,j}$ при разных n (то есть на разных подынтервалах длины t) могут не совпадать. Таким образом, мы видим, что рассмотрение схемы серий может оказаться полезным при построении моделей, учитывающих возможный дрейф распределений приращений биржевых цен во времени.

Перейдем к формальному построению модели. Наша базовая модель теперь будет иметь вид

$$P_n(t) - P_n(0) = \sum_{j=1}^{N_n} X_{n,j} \tag{7.7}$$

где $P_n(t)$ и $P_n(0)$ — значения цены акции соответственно в конце n-го интервала длины t и в его начале; $\{X_{n,j}\}_{j\geq 1}$ — с.в., предполагаемые независимыми и одинаково распределенными при каждом фиксированном n, которые интерпретируются как элементарные изменения цены акции в течение n-го интервала времени длины t; N_n — натуральнозначные с.в., предполагаемые независимыми от $\{X_{n,j}\}_{j\geq 1}$ при каждом $n \geq 1$. Они играют роль числа контрактов, заключенных в течение n-го интервала времени. Обязательно надо отметить, что возможны различные интерпретации модели (7.7), включая как приведенные выше, так и те, в которых параметр n не имеет никакого "физического" смысла, а будучи чисто вспомогательным, используется только для построения предельных аппроксимаций. Например, мы вполне можем связать n с t и изучать распределения приращений биржевых цен за интервалы увеличивающейся длины, что естественно при исследовании эффекта "унормаливания" распределений приращений биржевых цен на интервалах увеличивающейся длины (см. выше).

Мы имеем дело с абсолютными приращениями. Но если интересоваться относительными приращениями, то нашу аддитивную модель (7.7) можно переписать по отношению к логарифмам рассматриваемых приращений.

Таким образом, мы видим, что схема серий (пусть даже одинаково в каждой серии распределенных с.в.) достаточно гибка и допускает много разных способов интерпретации параметров и случайных величин, участвующих в формальной модели.

Если k — натуральное, то через $S_{n,k}$ мы будем обозначать сумму $X_{n,1} + \ldots + X_{n,k}$. Относительно слагаемых $X_{n,j}$ мы сделаем "хорошие" предположения, что они предельно малы и удовлетворяют условиям центральной предельной теоремы в том смысле, что существуют натуральные k_n такие, что $k_n \to \infty$ при $n \to \infty$ и

$$\mathsf{P}(S_{n,k_n} < x) \Rightarrow \Phi\left(\frac{x-a}{\sigma}\right) \qquad (n \to \infty) \tag{7.8}$$

при некоторых $a \in \mathbb{R}$ и $0 < \sigma < \infty$ (включая случаи $a = 0$ и $\sigma = 1$). Заметим, что (7.8) имеет место в следующей довольно общей ситуации. В соответствии с приведенной выше аргументацией мы будем предполагать, что $X_{n,j}$ имеют конечные дисперсии.

Предположим, что $X_{n,j}$ можно представить в виде $X_{n,j} = X^*_{n,j} + a_n$, где a_n неслучайны и описывают систематический снос элементарного приращения $X_{n,j}$ на n-ом интервале времени, и $\mathsf{E}X^*_{n,j} = 0$, $\mathsf{D}X^*_{n,j} = \sigma_n^2 < \infty$, так что $\mathsf{E}X_{n,1} = a_n$ и $\mathsf{D}X_{n,1} = \sigma_n^2$. Предположим, что $a_n k_n \to a$ и $k_n\sigma_n^2 \to \sigma^2$ при $n \to \infty$. Тогда из классического результата об условиях асимптотической нормальности распределений сумм независимых с.в. с конечными дисперсиями в схеме серий (см., например, (Гнеденко и Колмогоров, 1949)) мы получаем, что (7.8) имеет место тогда и только тогда, когда выполнено условие Линдеберга: для любого $\varepsilon > 0$

$$\lim_{n\to\infty} k_n \mathsf{E}(X^*_{n,1})^2 \mathsf{I}(|X^*_{n,1}| \geq \varepsilon) = 0,$$

то есть, квадратичные хвосты распределений элементарных приращений должны убывать достаточно быстро (здесь $\mathsf{I}(A)$ обозначает индикатор множества A). Кстати, это условие, очевидно, противоречит гипотезе о бесконечности дисперсии элементарных приращений, выдвинутой Мандельбротом и Фамой, но, как мы уже видели и еще увидим, также может привести к асимметричным и островершинным распределениям.

Теперь в нашем распоряжении есть все, чтобы сформулировать следующий основной результат. Напомним, что символом L_1 мы обозначаем метрику, метризующую слабую сходимость в пространстве случайных величин (или, что фактически то же самое, в пространстве функций распределения). Например, в качестве L_1 можно рассматривать метрику Леви, определенную во Введении.

Теорема 7.3. *Пусть $\{X_{n,j}\}$ и N_n удовлетворяют описанным выше условиям. Предположим, что существует последовательность натуральных чисел $\{k_n\}_{n\geq 1}$ такая, что имеет место (7.8) с некоторыми $a \in \mathbb{R}$ и $\sigma < \infty$. Тогда слабая сходимость приращений процесса цены акции к некоторой случайной величине Z*

$$P_n(t) - P_n(0) - c_n \Rightarrow Z \qquad (n \to \infty) \tag{7.9}$$

с некоторыми $c_n \in \mathbb{R}$ имеет место тогда и только тогда, когда существует слабо компактная последовательность случайных величин $\{U_n\}_{n\geq 1}$ и ограниченная числовая последовательность $\{v_n\}_{n\geq 1}$, такие что

$$\mathsf{E}\exp\{isZ\} = \mathsf{E}\exp\left\{is(U_n + v_n) - \frac{1}{2}s^2\sigma^2 U_n\right\} \tag{7.10}$$

при всех $n \geq 1$ и $s \in \mathbb{R}$, а также

$$L_1(N_n/k_n, U_n) \to 0 \qquad (n \to \infty), \tag{7.11}$$
$$|v_n + c_n| \to 0 \qquad (n \to \infty).$$

Доказательство этого утверждения, к сожалению, требует привлечения методов, не затрагивавшихся в данном курсе. Заинтересованный читатель может самостоятельно получить его как простое следствие Теоремы 3 из работы (Korolev, 1995) (см. также Теорему 4.2.1 в книге (Gnedenko and Korolev, 1996)).

Предельная х.ф. (7.10) соответствует ф.р.

$$\mathsf{P}(Z < x) = \mathsf{E}\Phi\left(\frac{x - b - aU}{\sigma\sqrt{U}}\right) \tag{7.12}$$

где $b \in \mathbb{R}$ и с.в. U неотрицательна. К сожалению, семейство сдвиг-масштабных смесей нормальных законов вида (7.12) неидентифицируемо (см. раздел 4). Поэтому соотношению (7.12) могут удовлетворять разные с.в. U с разными распределениями. Именно поэтому в формулировке Теоремы 3 присутствует условие (7.11), описывающее "слабое сближение" последовательностей $\{N_n/k_n\}$ и $\{U_n\}$, а не допускающее более удобную интерпретацию условие $N_n/k_n \Rightarrow U$, которое достаточно для (7.11) и, стало быть, для (7.9).

Как говорилось выше, известно много попыток найти удобную и одновременно адекватную параметрическую форму распределения приращений процесса биржевых цен. Недавно О. Барндорфф-Нильсен (Barndorff-Nielsen, 1994) сообщил об установленном им и его коллегами чрезвычайно хорошем согласии так называемого гауссовского\\обратного гауссовского (GIG) распределения с данными о ценах на датских и немецких биржах. Плотность GIG-распределения имеет вид

$$g(x;\alpha,\beta,\mu,\delta) =$$
$$A(\alpha,\beta,\mu,\delta)\left[Q\left(\frac{x-\mu}{\delta}\right)\right]^{-1} K_1\left(\delta\alpha Q\left(\frac{x-\mu}{\delta}\right)\right)e^{\beta x}, \tag{7.13}$$

где K_1 — модифицированная бесселева функция третьего рода с индексом 1,

$$A(\alpha,\beta,\mu,\delta) = \frac{\alpha}{\pi}\exp\{\delta\sqrt{\alpha^2-\beta^2} - \beta\mu\}, \quad Q(x) = \sqrt{1+x^2}.$$

Хорошее согласие этого распределения с реальными данными на первый взгляд может показаться удивительным, поскольку плотность (7.13) имеет довольно громоздкий и труднообозримый вид. Однако с помощью Теоремы 7.3 мы убеждаемся, что на самом деле ничего удивительного здесь нет. Действительно, еще в 1977 г. было установлено, что GIG-распределение принадлежит к семейству обобщенных гиперболических законов, каждый из которых представляет собой сдвиг-масштабную смесь нормальных распределений (Barndorff-Nielsen, 1977), см. также (Barndorff-Nielsen, 1978) и (Barndorff-Nielsen et al., 1982). А именно, плотность (7.13) соответствует ф.р.

$$G(x;\alpha,\beta,\mu,\delta) = \mathsf{E}\Phi\left(\frac{x-\mu-\beta U}{\sqrt{U}}\right) \tag{7.14}$$

где с.в. U имеет обратное нормальное распределение с плотностью

$$p(u;\delta,\sqrt{\alpha^2-\beta^2}) =$$
$$\frac{\delta}{\sqrt{2\pi}}\exp\{\delta\sqrt{\alpha^2-\beta^2}\}u^{-3/2}\exp\left\{-\frac{1}{2}\left[(\alpha^2-\beta^2)u + \frac{\delta^2}{u}\right]\right\}. \tag{7.15}$$

Именно это обстоятельство послужило причиной того, что распределение (7.13) называется гауссовским\\обратным гауссовским. Отметим, что правая часть (7.14) с точностью до обозначений совпадает с правой частью (7.8). Свойства некоторых моделей финансовой математики, основанных на GIG-распределениях, обсуждаются в статье (Ширяев, 1995).

В связи с сообщением О. Барндорфф-Нильсена возникает вопрос, почему смешивающее распределение (7.15) обеспечивает хорошее совпадение модели с данными. Один возможный и почти тривиальный ответ заключается в том, что, как гласит одно из традиционных правил, известных в прикладной статистике, на практике четырех параметров достаточно для подгонки любой кривой к любым данным. Другими словами, семейство GIG-распределений настолько разнообразно, что в нем всегда найдется представитель, удовлетворительно согласующийся с данными. Менее тривиальные, но по-видимому, более правильные ответы можно найти, двигаясь по пути построения допредельных моделей для поведения с.в. N_n, фигурирующей в Теореме 7.3, с учетом или рассуждений, использованных самим О. Барндорфф-Нильсеном при описании класса гауссовских\\обратных гауссовских распределений, когда он изучал логарифмы размеров дробящихся частиц (Barndorff-Nielsen, 1977), или хорошо известной интерпретации обратного гауссовского распределения в терминах времени достижения заданного уровня винеровским процессом.

8 Математические модели неопределенности и информации

Материал, излагаемый в этом и следующем разделах, напрямую не связан с предельными теоремами для случайных сумм. Однако он представляется уместным в рамках курса "вероятностные модели", так как здесь будут сформулированы некоторые основные принципы построения математических (и прежде всего, вероятностных) моделей реальных объектов и процессов в условиях неопределенности. В разделах 6 и 7 мы уже видели, сколь важен правильный выбор распределения вероятностей в качестве математической модели тех или иных феноменов. Цель этого и следующего разделов – сформулировать несколько рецептов выбора распределений, которые в первую очередь следует рассматривать в качестве математических моделей в реальных ситуациях, характеризующихся полным или частичным отсутствием информации об изучаемом явлении. Мы также хотим привлечь внимание читателей к идеям Ю. В. Линника и А. Реньи, которые предложили рассматривать предельные теоремы теории вероятностей как одно из возможных описаний всеобщего принципа неубывания неопределенности в замкнутых системах, который, в частности, проявляет себя как второе начало термодинамики.

Начнем с построения математических моделей самих понятий информации и неопределенности. Хартли и Шеннон предложили связать понятия неопределенности и вероятности. Такой подход оказался вполне разумным и плодотворным.

Пусть A и B – события, вероятности которых положительны.

Определение 8.1. Информацией (по Шеннону), содержащейся в событии B относительно события A, называется

$$I(A|B) = \log \frac{\mathsf{P}(A|B)}{\mathsf{P}(A)}.$$

Если $B = A$, то, очевидно, $I(A|A) = -\log \mathsf{P}(A)$. Таким образом, мы приходим к следующему важному определению.

Определение 8.2. Информацией (по Шеннону), содержащейся в событии A, называется

$$I(A) = -\log \mathsf{P}(A). \tag{8.1}$$

Смысл этих определений легко пояснить, рассмотрев простейшие свойства введенных понятий. Эти свойства оказываются аналогичными тем, которые должны быть присущи информации с точки зрения здравого смысла.

1) Осуществление события, вероятность которого невелика, как правило, несет в себе больше информации, нежели осуществление события, вероятность которого значительна. Например, в начале нашего столетия представлялось практически невероятно обнаружить живой экземпляр считавшегося давно вымершим

вида кистеперых рыб. Когда же такая рыба – целакант –была поймана, фактически произошла революция в ихтиологии. В то же время, поимка любого экземпляра рыбы такого распространенного вида как, скажем, треска, может нести в себе информацию разве что о месте прохождения косяка этих рыб и никакой научной революции не вызывает.

2) Если события A и B независимы, то, очевидно, осуществление события B не дает никакой информации о событии A. Действительно, в этом случае мы имеем
$$I(A|B) = \log 1 = 0.$$

3) Если события A и B независимы, то их одновременное осуществление несет в себе столько же информации, сколько содержится в каждом из них в отдельности. Действительно, в этом случае мы имеем
$$I(AB) = -\log \mathsf{P}(AB) = -\log(\mathsf{P}(A)\mathsf{P}(B)) = \\ -\log \mathsf{P}(A) - \log \mathsf{P}(B) = I(A) + I(B).$$

Логарифмическое определение информации (8.1) восходит к Хартли (Hartley, 1928). Основание логарифма не играет определяющей роли и существенно лишь для выбора единицы измерения информации. Обычно используют логарифмы по основанию 2, для которых единица информации содержится в событии, вероятность которого 1/2. Такая единица информации называется бит (от английского bit (кусочек) или как аббревиатура термина BInary digiT (двоичный разряд)). Единица информации, порожденная натуральными логарифмами, называется нат (nat).

Пусть $\mathcal{E}$ – эксперимент, в котором может осуществиться лишь один из n исходов $A_1, \ldots, A_n$. Обозначим $\mathsf{P}(A_i) = p_i$ (очевидно, что $p_1 + \ldots + p_n = 1$). Тогда мы можем считать информацию, полученную в результате этого эксперимента, случайной величиной, принимающей значения $I(A_1), \ldots, I(A_n)$ соответственно с вероятностями $p_1, \ldots, p_n$. Обозначим эту случайную величину $Q(\mathcal{E})$. Введем следующую интегральную информационную характеристику $\mathcal{E}$.

Определение 8.3. Энтропией $H(\mathcal{E})$ эксперимента $\mathcal{E}$ назывется величина

$$H(\mathcal{E}) = \mathsf{E}Q(\mathcal{E}) = -\sum_{i=1}^{n} p_i \log p_i. \tag{8.2}$$

Энтропия эксперимента может служить мерой его неопределенности, что подтверждается совпадением следующих свойств формально введенной величины $H(\mathcal{E})$ с ожидаемыми с позиций здравого смысла свойствами разумной меры неопределенности.

1) $H(\mathcal{E}) \geq 0$, причем равенство достигается тогда и только тогда, когда существует $i \in \{1, \ldots, n\}$ такое, что $p_i = 1$.

2) Пусть $\mathcal{E}_0$ – эксперимент с n равновероятными исходами. Тогда $H(\mathcal{E}) \leq H(\mathcal{E}_0)$, каким бы ни был эксперимент $\mathcal{E}$ с таким же числом n возможных исходов.

3) Пусть $\mathcal{E}_1$ – эксперимент с $n-1$ исходом, построенный из эксперимента $\mathcal{E}$ с помощью объединения двух исходов, скажем, A_i и A_j $(i \neq j)$, и пусть $\mathcal{E}_2$ – эксперимент с исходами A_i и A_j, вероятности которых (в рамках $\mathcal{E}_2$) соответственно равны $p_i/(p_i+p_j)$ и $p_j/(p_i+p_j)$. Тогда

$$H(\mathcal{E}) = H(\mathcal{E}_1) + (p_i + p_j)H(\mathcal{E}_2).$$

4) Энтропия $H(\mathcal{E})$ зависит не от $A_1, \ldots, A_n$, а от $p_1, \ldots, p_n$, будучи симметрической функцией переменных $p_1, \ldots, p_n$.

5) Энтропия $H(\mathcal{E})$ – непрерывная функция от $p_1, \ldots, p_n$.

Доказательство этих свойств, равно как и обсуждение других свойств можно найти, например, в книге (Голдман, 1957). Д. К. Фаддееву удалось показать, что система свойств 1-5 однозначно определяет функционал (8.2) (Фаддеев, 1956) (см. также (Яглом и Яглом, 1973)).

Очевидно, вместо экспериментов с n исходами мы можем рассматривать дискретные случайные величины, принимающие какие-либо значения $x_1, \ldots, x_n$ с вероятностями $p_1, \ldots, p_n$, считая, что значение x_i случайной величины X наблюдается в результате осуществления исхода A_i эксперимента $\mathcal{E}$. Поэтому по аналогии с (8.2) мы можем определить энтропию простой случайной величины X как

$$H(X) = -\sum_{i=1}^{n} p_i \log p_i,$$

или, вводя плотность $p(x)$ распределения X относительно считающей меры

$$p(x) = \begin{cases} p_i, & x = x_i, \ i = 1, \ldots, n; \\ 0, & x \notin \{x_1, \ldots, x_n\}, \end{cases}$$

как

$$H(X) = -\mathsf{E} \log p(X). \tag{8.3}$$

По формальной аналогии с (8.3), если X – абсолютно непрерывная случайная величина с лебеговой плотностью $p(x)$, то определим энтропию $H(X)$ величины X так же, как (8.3). Однако, аналогия таких определений энтропии дискретной и абсолютно непрерывной случайных величин оказывается чисто формальной. Как известно, каждая случайная величина X как функция элементарного исхода может быть представлена в виде поточечного предела последовательности простых случайных величин X_n. Тогда, конечно же, $X_n \Rightarrow X$. Но если мы попытаемся использовать предельный переход при неограниченно увеличивающемся числе значений аппроксимирующих простых случайных величин для того, чтобы получить энтропию предельной абсолютно непрерывной случайной величины X в виде (8.3), мы неизбежно

потерпим неудачу, так как предел энтропий аппроксимирующих простых случайных величин оказывается бесконечным. Тем не менее, выражение в правой части (8.3) оказывается пределом для "стандартизованных" энтропий аппроксимирующих простых случайных величин в следующем смысле. По сути энтропия (8.3) абсолютно непрерывной случайной величины определяет среднюю информацию, содержащуюся в X по сравнению с "бесконечно большой аддитивной постоянной". Чтобы пояснить сказанное, разобьем область значений величины X на непересекающиеся интервалы Δ_i равной длины δ. Определим соответствующую дискретную аппроксимацию X_δ случайной величины X, полагая X_δ равной некоторому фиксированному элементу интервала Δ_i, если X попадает в этот интервал. Тогда $X_\delta \Rightarrow X$ $(\delta \to 0)$, и при некоторых условиях регулярности на плотность $p(x)$ величины X для некоторых точек $x_i^* \in \Delta_i$ мы имеем

$$\begin{aligned} H(X_\delta) = -\sum_i \mathsf{P}(X \in \Delta_i) \log \mathsf{P}(X \in \Delta_i) = -\sum_i p(x_i^*)\delta \log(p(x_i^*)\delta) = \\ -\sum_i p(x_i^*)\delta \log p(x_i^*) - \sum_i p(x_i^*)\delta \log \delta = \\ -\sum_i p(x_i^*)\delta \log p(x_i^*) - \log \delta. \end{aligned} \tag{8.4}$$

Первое слагаемое в правой части (8.4) представляет собой интегральную сумму Дарбу для величины $H(X)$, определяемой соотношением (8.3). Поэтому мы можем записать

$$H(X) = \lim_{\delta \to 0} [H(X_\delta) + \log \delta].$$

В соответствии с Определением 8.2, величину $-\log \delta$ можно интерпретировать как информацию, содержащуюся в событии, вероятность которого равна δ. Таким образом, эта величина может характеризовать рост неопределенности X_δ, вызванный квантованием.

Понятие информации можно формализовать многими способами. Наряду с подходом Шеннона, рассмотренным выше, в дальнейшем мы также будем рассматривать информацию по Фишеру $J_0(X)$, которая, как известно из курса математической статистики, определяется для случайной величины X с непрерывно дифференцируемой плотностью $p(x)$ как

$$J_0 = \mathsf{E}\left[\frac{p'(X)}{p(X)}\right]^2.$$

9 Информационные принципы построения вероятностных моделей

Энтропии (8.2) и (8.3) оказываются тесно связанными по форме с термодинамической энтропией замкнутой физической системы, определенной Больцманом как

$$H = -\sum_{i=1}^{N} \frac{m_i}{M} \log \frac{m_i}{M},$$

где M – число молекул в системе, а m_i – число молекул со скоростью $v \in [v_i, v_i + \delta)$, $i = 1, \ldots, N$. Знаменитый второй закон термодинамики устанавливает, что $\Delta H \geq 0$, то есть энтропия замкнутой системы не убывает во времени. Формальное совпадение определений энтропии в физике и теории информации выдвигает вопрос о том, существуют ли в теории вероятностей аналоги второго закона термодинамики.

Еще одна причина для поисков взаимосвязи между результатами теории вероятностей и вторым законом термодинамики заключается в следующем. Этот физический закон является проявлением всеобщего принципа неубывания неопределенности замкнутых систем, необязательно описываемых на "физическом" уровне. Так как, вообще говоря, теория вероятностей – это прикладная наука, изучающая математические модели реальных явлений с учетом неопределенных или случайных факторов, должны быть и математические, точнее говоря, вероятностные модели самого́ принципа неубывания неопределенности.

Перед тем как мы перейдем к основным выводам этого раздела и сформулируем некоторые общие принципы и проблемы, основанные на этих выводах, напомним экстремальные энтропийные свойства некоторых распределений вероятностей.

Лемма 9.1. (i) *Пусть Y_1 – случайная величина с равномерным распределением на интервале $[-a, a]$, $a > 0$. Тогда $H(X) \leq H(Y_1)$ для любой случайной величины X такой, что $\mathsf{P}(-a \leq X \leq a) = 1$.*

(ii) *Пусть Y_2 – случайная величина с показательным распределением, $\mathsf{E}Y_2 = \alpha > 0$. Тогда $H(X) \leq H(Y_2)$ для любой случайной величины X такой, что $\mathsf{P}(X \geq 0) = 1$, $\mathsf{E}X = \alpha$.*

(iii) *Пусть Y_3 – нормально распределенная случайная величина с параметрами $0, \sigma^2$. Тогда $H(X) \leq H(Y_3)$ для любой случайной величины X с $\mathsf{E}X = 0$ и $\mathsf{E}X^2 = \sigma^2$.*

Доказательство этого утверждения можно найти, например, в (Голдман, 1957).

Таким образом, равномерное распределение имеет максимальную энтропию среди всех распределений с ограниченным носителем, показательное распределение имеет максимальную энтропию среди всех распределений, сосредоточенных на неотрицательной полуоси и имеющих конечное математическое ожидание, нормальное распределение имеет максимальную энтропию среди всех распределений с конечными дисперсиями.

Распределения, перечисленные в Лемме 9.1, играют самые важные роли во многих предельных теоремах теории вероятностей. Действительно, в полном соответствии со своим названием, центральная предельная теорема играет центральную роль в теории вероятностей и математической статистике. Разнообразные модификации этой теоремы описывают сходимость к нормальному закону распределений многих случайных величин, аддитивно зависящих от других "атомических" случайных величин. Показательное распределение, во-первых, играет важную роль, выступая в качестве предельного закона в предельных теоремах для случайных величин, также получаемых из "атомических" случайных величин, но с помощью операции взятия максимума. Во-вторых, будучи тесно связанным с распределением Пуассона (показательно распределенные длины интервалов между событиями определяют пуассоновский процесс), оно также неявно участвует и в предельныхтеоремах для сумм случайных величин. Таким образом, сопоставление предельных законов, появляющихся в предельных теорем теории вероятностей с распределениями, перечисленными в Лемме 9.1, приводит нас к важному предположению о том, что *всеобщий принцип неубывания неопределенности в замкнутых системах проявляет себя в теории вероятностей в виде предельных теорем, в которых предельный переход осуществляется при неограниченном возрастании числа "атомических" случайных величин, участвующих в математической модели.*

В подтверждение этого предположения мы приведем несколько информационных формулировок классических предельных теорем. В частности, мы сформулируем центральную предельную теорему в терминах плотностей. Мы базируемся на результатах работы (Barron, 1986).

Пусть X – случайная величина с конечной дисперсией. Вместо шенноновской энтропии, введенной в (8.3), при доказательстве центральной предельной теоремы удобнее иметь дело с относительной энтропией, определяемой следующим образом. Если $p(x)$ – плотность случайной величины X (относительно меры Лебега), то относительной энтропией величины X назовем

$$D(X) = \int p(x) \log \left[\frac{p(x)}{\phi(x)}\right] dx$$

где ϕ – нормальная плотность с теми же математическим ожиданием и дисперсией, что и p. В противном случае положим $D(X) = \infty$. Легко видеть, что если $\mathsf{E}X = a$, $\mathsf{D}X = \sigma^2$, то

$$\begin{aligned} D(X) &= \int p(x) \log p(x)\, dx - \int p(x) \log \phi(x)\, dx = \\ &-H(X) + \int p(x) \left[\frac{1}{2}\log(2\pi\sigma^2) + \frac{(x-a)^2}{2\sigma^2}\right] dx = \\ &-H(X) + \frac{1}{2}\left[\log(2\pi\sigma^2) + 1\right], \end{aligned}$$

откуда $H(X) = \frac{1}{2}[\log(2\pi\sigma^2) + 1] - D(X)$, а $\frac{1}{2}[\log(2\pi\sigma^2) + 1]$, очевидно, – энтропия нормального распределения с дисперсией σ^2. Следовательно, относительная энтропия $D(X)$ представляет собой разность между энтропией нормального распределения и энтропией случайной величины X. В силу вогнутости логарифма, D неотрицательна и равна нулю тогда и только тогда, когда $p(x) = \phi(x)$ почти всюду. (Как следствие, мы можем получить, что нормальный закон имеет максимальную энтропию, то есть пункт (iii) Леммы 9.1.)

Пусть Y – случайная величина с непрерывно дифференцируемой плотностью $g(y)$, $\mathsf{E}Y = a$, $\mathsf{D}Y = \sigma^2 < \infty$. Определим стандартизованную фишеровскую информацию как

$$J(Y) = \sigma^2 \mathsf{E}\left[\frac{g'(Y)}{g(Y)} - \frac{\phi'(Y)}{\phi(Y)}\right]^2,$$

где ϕ – нормальная плотность с теми же средним и дисперсией, что и g. Фишеровская информация $J_0(Y) = \mathsf{E}[g'(Y)/g(Y)]^2$ удовлетворяет соотношению

$$\sigma^2 J_0(Y) = J(Y) + 1.$$

Поскольку $J \geq 0$, и равенство здесь возможно лишь тогда, когда $g = \phi$, нормальный закон доставляет минимум фишеровской информации при фиксированной дисперсии (откуда вытекает неравенство Рао-Крамера $J_0 \geq \sigma^2$). Величины D и J инвариантны относительно параметров сдвига и масштаба.

Шенноновскую энтропию и фишеровскую информацию связывает следующее утверждение.

Лемма 9.2. *Относительная энтропия $D(X)$ является интегралом от фишеровской информации в том смысле, что, если X – случайная величина с конечной дисперсией, то*

$$D(X) = \int_0^1 J(\sqrt{t}X + \sqrt{1-t}Y)\frac{dt}{2t}, \tag{9.1}$$

где Y – независимая от X нормально распределенная случайная величина с теми же средним и дисперсией, что и X.

Доказательство см. в (Barron, 1986).

В терминах J и D центральная предельная теорема формулируется следующим образом. Пусть $\{X_j\}_{j\geq 1}$ – независимые одинаково распределенные случайные величины с нулевым средним и дисперсией σ^2. Пусть $S_n = (X_1 + \ldots + X_n)/\sqrt{n}$ – нормированная сумма этих величин. Пусть Y – нормально распределенная случайная величина с нулевым средним и дисперсией σ^2, независимая от $\{X_j\}_{j\geq 1}$. Следующая теорема устанавливает сходимость фишеровских информаций.

Теорема 9.1 (Barron, 1986). *Пусть $S'_n = \sqrt{t}S_n + \sqrt{1-t}Y$ при фиксированном $t \in [0,1]$. Тогда $J(S'_n) \geq J(S'_{2n})$ при каждом $n \geq 1$. Более того,*

$$\lim_{n\to\infty} J(S'_n) = 0. \tag{9.2}$$

Эквивалентно, последовательность фишеровских информаций $J_0(S'_n)$ сходится к границе Рао-Крамера σ^2.

Поскольку $J(X) \geq 0$, причем равенство достигается только тогда, когда плотность случайной величины X нормальна, соотношение (9.2) означает, что распределение случайной величины $\sqrt{t}S_n + \sqrt{1-t}Y$ становится все "более и более" нормальным при $n \to \infty$ и следовательно, в силу теоремы Леви-Крамера о разложимости нормального закона только на нормальные компоненты, все "более и более" нормальным становится распределение величины S_n.

Теорема 9.2 (Barron, 1986). *При каждом $n \geq 1$ справедливо неравенство $D(S_{2n}) \leq D(S_n)$. Более того, относительная энтропия стремится к нулю*

$$\lim_{n\to\infty} D(S_n) = 0 \tag{9.3}$$

тогда и только тогда, когда при каком-нибудь n величина $D(S_n)$ конечна. Эквивалентно, шенноновская энтропия $H(S_n)$ сходится к энтропии нормального закона $\frac{1}{2}[\log(2\pi\sigma^2) + 1]$, если она конечна при каком-нибудь n.

Следует отметить, что сходимость, устанавливаемая в теореме 9.2, сильнее, чем слабая сходимость.

Если мы посмотрим на предельные теоремы теории вероятностей (например, для случайных сумм независимых случайных величин) с информационной точки зрения и согласимся с тезисом о том, что эти теоремы описывают различные проявления всеобщего принципа неубывания неопределенности в теории вероятностей, мы сразу же обратим внимание, что список возможных предельных законов ни в коей мере не исчерпывается теми, которые упомянуты в Лемме 9.1. Простейшими примерами служат смеси нормальных или других безгранично делимых законов, упоминаемые в данном курсе лекций. Это может означать следующее. Во-первых, системы, моделируемые с помощью случайных сумм, не являются замкнутыми в том смысле, что помимо слагаемых есть еще один источник неопределенности – случайные индексы. При этом агрегируется лишь неопределенность, обусловленная стохастичностью слагаемых, поскольку предполагается, что их число в определенном смысле неограниченно возрастает, в то время как неопределенность, обусловленная стохастичностью индексов, не обязана увеличиваться, поскольку, вообще говоря, никаких структурных предположений об индексах не делается. Во-вторых, в соответствии с нашим основным предположением, следует ожидать, что предельные законы должны обладать экстремальными энтропийными свойствами при определенных условиях.

Поэтому мы приходим к следующим проблемам.

1. Найти условия, при которых конкретные предельные законы, появляющиеся в предельных теоремах теории вероятностей, обладают экстремальными энтропийными свойствами. Первыми кандидатами для того, чтобы быть подвергнутыми такому исследованию, конечно же, являются устойчивые законы или законы, предельные для экстремальных порядковых статистик. Если эта проблема будет (хотя бы частично) решена, то будет получено еще одно подтверждение нашего основного предположения и, что не менее важно, будет уточнена роль этих распределений как математических моделей в тех или иных ситуациях.

2. Дать информационные доказательства предельных теорем теории вероятностей, например, о сходимости распределений сумм независимых одинаково распределенных случайных величин к ненормальным устойчивым законам или о сходимости распределений случайных сумм.

3. Поскольку распределения с максимальной энтропией обладают целым рядом благоприятных характеристических свойств (таких как свойство отсутствия последействия показательного закона, эквивалентность некоррелированности и независимости нормально распределенных случайных величин и т. д.), целесообразно найти количественные оценки устойчивости упомянутых характеризаций в терминах информационных характеристик типа (9.5).

С практической точки зрения необходимо отметить следующее. При построении математических моделей широко используется принцип максимума энтропии, согласно которому при отсутствии информации для описания реального явления выбирается та модель, которая предполагает наибольшую возможную неопределенность описываемой ею ситуации. В соответствии со сказанным выше, этот принцип может быть обобщен и уточнен.

1. При недостатке информации круг поисков подходящих вероятностных моделей должен быть расширен и должен включать в себя разнообразные комбинации законов, обладающих свойством максимума энтропии, например, их смеси, которые могут соответствовать моделям открытых систем, не включающим все факторы, описывающие поведение этих систем. В самом деле, распределения с максимальной энтропией являются адекватными моделями в идеальных ситуациях, предполагающих замкнутость моделируемой системы. В то же время, по само́му своему определению никакая математическая модель не может учесть все факторы, воздействующие на изучаемую систему. Поэтому никакая моделируемая система не может считаться замкнутой и соответствие какого-либо распределения, упомянутого в Лемме 9.1, экспериментальным данным скорее является исключением, чем правилом. Однако вполне естественно в качестве начального приближения, то есть в качестве грубой модели, рассматривать законы с максимальной энтропией, а затем, вовлекая в модель все новые и новые факторы, усложнять ее. При этом естественно ожидать, что распределения с максимальной энтропией должны каким-либо образом участвовать в более сложных моделях.

2. Наоборот, если наблюдается хорошее согласие вероятностной модели, то есть распределения с экспериментальными данными, то следует изучить его связь с распределениями с максимальной энтропией. Например, следует изучить возможность представления такого распределения в виде смеси законов с максимальной энтропией. Хорошей иллюстрацией этого принципа служит пример с GIG-распределением, упомянутый в разделе 7. Такая взаимосвязь, если ее найти, может позволить исследователю предложить подходящую формулировку предельной теоремы, которая в свою очередь, может способствовать лучшему осознанию изучаемого явления.

ЛИТЕРАТУРА

1) С. А. Айвазян, В. М. Бухштабер, И. С. Енюков и Л. Д. Мешалкин. *Прикладная статистика. Классификация и снижение размерности.* М.: "Финансы и статистика", 1989.

2) В. Е. Бенинг и В. Ю. Королев. Асимптотическое поведение обобщенных неординарных процессов Кокса. — Вестник Московского университета, сер. Вычислительная математика и кибернетика, 1997, N 4, с. 3-16.

3) Б. В. Гнеденко и А. Н. Колмогоров. *Предельные распределения для сумм независимых случайных величин.* М.-Л.: ГИТТЛ, 1949.

4) Б. В. Гнеденко и Х. Фахим. Об одной теореме переноса. — Доклады АН СССР, 1969, т. 187, N 1, с. 15-17.

5) С. Голдман. *Теория информации.* М.: ИЛ, 1957.

6) В. М. Золотарев. Одномерные устойчивые распределения. М.: Наука, 1983.

7) О. К. Исаенко и В. Ю. Урбах. Разделение смесей распределений вероятностей на их составляющие. — в сб.: *Итоги науки и техники. Сер. Теория вероятн., матем. статист., теорет. киберн.*,М.: ВИНИТИ, 1976, с. 37-58.

8) В. Ю. Королев. Сходимость случайных последовательностей с независимыми случайными индексами. — Теория вероятн. и ее примен. 1994, т. 39, вып. 2, с. 313-333.

9) В. М. Круглов и В. Ю. Королев. *Предельные теоремы для случайных сумм.* М.: изд-во МГУ, 1990.

10) В. М. Круглов. Смеси вероятностных распределений. — Вестник Московского университета, сер. Вычислительная математика и кибернетика, 1991, N 2, с. 3-15.

11) Ю. В. Линник. Теоретико-информационное доказательство центральной предельной теоремы. — Теория вероятн. и ее примен., 1959, т. 4, вып. 3, с. 311-321.

12) М. Лоэв. *Теория вероятностей.* М.: ИЛ, 1962.

13) П. В. Новицкий и И. А. Зограф. *Оценка погрешностей результатов измерений.* Л.: Энергоатомиздат, 1991.

14) С. Т. Рачев и Л. Рушендорф. Модели и расчеты контрактов с опционами. — Теория вероятн. и ее примен., 1994, т. 39, вып. 1, с. 150-190.

Оглавление