А. Г. Дьячков, Асимптотика энтропии Шеннона для суммы независимых случайных величин, *Фундамент. и прикл. матем.*, 1996, том 2, выпуск 4, 1019–1028

Асимптотика энтропии Шеннона для суммы независимых случайных величин

А. Г. ДЬЯЧКОВ
Московский государственный университет им. М. В. Ломоносова
e-mail: dyachkov@nw.math.msu.su

Памяти Б. В. Гнеденко

УДК 621.391.1+519.2

Ключевые слова: энтропия, локальная предельная теорема.

Аннотация

С помощью классической локальной предельной теоремы Б. В. Гнеденко и уточнений этой теоремы исследуется асимптотическое разложение энтропии Шеннона для суммы независимых одинаково распределенных случайных величин.

Abstract

A. G. Dyachkov, Asymptotics of the Shannon entropy for sums of independent random variables, Fundamentalnaya i prikladnaya matematika **2**(1996), 1019–1028.

With the help of the local limit theorem we investigate the asymptotics of the Shannon entropy for sums of independent identically distributed random variables.

1 Постановка задач и результаты

Пусть $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ — независимые одинаково распределенные величины, каждая из которых имеет решетчатое распределение с максимальным шагом h. В дальнейшем, без нарушения общности, предполагается, что $h = 1$. Через

$$p_1(s) = \Pr\{\xi_k = s\}, \quad s = \ldots, -1,\ 0,\ 1, \ldots,$$

$$H(\xi_k) = -\sum_s p_1(s) \ln p_1(s) < \infty$$

обозначим распределение ξ_k и его энтропию Шеннона. Пусть сумма

$$\zeta_n = \sum_{k=1}^{n} \xi_k$$

Фундаментальная и прикладная математика 1996, **2**, № 4, 1019–1028.

имеет распределение

$$p_n(s) = \Pr\{\zeta_n = s\}, \quad s = \ldots, -1, 0, 1, \ldots.$$

Из свойства субаддитивности энтропии [1] вытекает следующая аддитивная оценка

$$H(\zeta_n) = -\sum_s p_n(s) \ln p_n(s) < nH(\xi_1). \tag{1}$$

Наша цель — исследовать асимптотическое ($n \to \infty$) разложение $H(\zeta_n)$. Пусть $M\xi_k^2 < \infty$. Введем $a = M\xi_k$, $\sigma^2 = M(\xi_k - a)^2$. Оценку (1) уточняет доказанная в работе [2]

Лемма 1. *Если $M\xi_k^2 < \infty$, то*

$$H(\zeta_n) \leqslant \frac{1}{2} \ln \left[2\pi e \left(n\sigma^2 + \frac{1}{12}\right)\right]. \tag{2}$$

С помощью стандартной предельной теоремы устанавливается главный член асимптотического разложения $H(\zeta_n)$.

Теорема 1. *Если $M\xi_k^2 < \infty$, то $H(\zeta_n) = \dfrac{1}{2} \ln n(1 + o(1))$.*

Используя асимптотическое разложение $p_n(s)$ [3], можно доказать, что оценка (2) является асимптотически точной, а именно, имеет место

Теорема 2. *Если $M\xi_k^4 < \infty$, то при $n \to \infty$*

$$H(\zeta_n) = \frac{1}{2} \ln n + \frac{1}{2} \ln \left(2\pi e \sigma^2\right) + o(1).$$

Для некоторых распределений $p_1(s)$, имеющих экспоненциальные моменты (испытания Бернулли, пуассоновское, геометрическое), можно написать асимптотические разложения

$$H(\zeta_n) = \frac{1}{2} \ln n + \frac{1}{2} \ln \left(2\pi e \sigma^2\right) + \sum_{i=1}^{\infty} \frac{c_i}{n^i}, \tag{3}$$

где коэффициенты c_i вычисляются через семиинварианты распределения $p_1(s)$.

Теорема 3. *Пусть ξ_k принимает значения 0 и 1 соответственно с вероятностями p и $q = 1 - p$, $0 < p < 1$. Тогда при $n \to \infty$ разложение (3) имеет вид*

$$\begin{aligned} H(\zeta_n) = {} & \frac{1}{2} \ln n + \frac{1}{2} \ln(2\pi e q p) - \frac{1}{12n} \left[\frac{q^2}{p} + \frac{p^2}{q} - 1\right] - \frac{1}{24n^2} \left[\frac{p^2}{q^2} + \frac{q^2}{p^2}\right] + \\ & + \frac{1}{360n^3} \left[\frac{3p^4 - 12p^3 - 12p^2 + 3p - 1}{q^3} + \frac{3q^4 - 12q^3 - 12q^2 + 3q - 1}{p^3}\right] + \\ & + O(n^{-4}). \end{aligned} \tag{4}$$

В теореме 4 исследуется случай n независимых случайных величин, каждая из которых имеет пуассоновское распределение.

Теорема 4. *Пусть ξ_k имеет пуассоновское распределение с параметром λ_0, т. е. $p_1(s) = \lambda_0^s e^{-\lambda_0}/s!$, $s = 0, 1, 2, \ldots$, и пусть $\lambda = n\lambda_0$. Тогда при $n \to \infty$*

$$H(\zeta_n) = \frac{1}{2}\ln\lambda + \frac{1}{2}\ln(2\pi e) - \frac{1}{12\lambda} - \frac{1}{24\lambda^2} - \frac{19}{360\lambda^3} + O\left(\frac{1}{\lambda^4}\right).$$

В последней теореме исследуется случай суммы n независимых случайных величин, где каждая из них имеет геометрическое распределение, т. е. исследуется асимптотика энтропии отрицательного биномиального распределения.

Теорема 5. *Пусть ξ_k имеет геометрическое распределение, т. е.*

$$p_1(j) = pq^{j-1}, \quad j = 1, 2, \ldots,$$

где $0 < p < 1$, $q = 1 - p$. Тогда при $n \to \infty$

$$H(\zeta_n) = \frac{1}{2}\ln n + \frac{1}{2}\ln\left(2\pi e\frac{q}{p^2}\right) + n^{-1}\left[\frac{1}{12} - \frac{q+1}{6pq} + \frac{1}{4pq} - \frac{1}{4q} + \frac{q}{12p} - \right.$$
$$\left. - \frac{q^2(q+1)}{3p} + \frac{q(q+1)}{2p} - \frac{q^2}{p} + \frac{3q^3}{4p} - \frac{q}{4} - \frac{q(q+1)}{3} + \frac{3q^2}{4} - \frac{p}{12}\right] + O\left(n^{-2}\right).$$

Теоремы 1 и 2 с помощью локальных теорем [4] обобщаются и на случай неодинаково распределенных случайных величин. Используя локальные предельные теоремы для плотностей, можно доказать аналоги сформулированных утверждений в непрерывном случае.

Часть результатов этой работы была анонсирована в тезисах [7].

2 Доказательство теорем

Доказательство теоремы 1. Введем обозначения

$$\eta_k = \frac{\zeta_n - na}{\sigma\sqrt{n}}, \quad y_{ns} = \frac{s - na}{\sigma\sqrt{n}}, \quad B_n = \sigma\sqrt{n} \text{ и } \phi(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}x^2\right).$$

Из классической формулировки локальной предельной теоремы вытекает следующее утверждение. Пусть $A > 0$ — фиксированное число, тогда для значений s, таких что $\|y_{ns}\| \leqslant A$, найдется положительная последовательность ε_n (зависит от n и A), такая что $\varepsilon_n \to 0$ при $n \to \infty$ и справедливы неравенства

$$\phi\left(y_{ns}\right)\left(1 - \varepsilon_n\right) \leqslant B_n p_n(s) \leqslant \phi\left(y_{ns}\right)\left(1 + \varepsilon_n\right). \tag{5}$$

Очевидно
$$H\left(\zeta_n\right) \geqslant -\sum_{s \in M(A)} p_n(s) \ln p_n(s),$$
где
$$M\left(A\right) = \left\{s\colon\ |y_{ns}| \leqslant A\right\}.$$
Используя нижнюю оценку для $p_n(s)$ из (5), получаем

$$H(\zeta_n) \geqslant -\sum_{s\in M(A)} \frac{e^{-\frac{y_{ns}^2}{2}}}{\sqrt{2\pi n\sigma^2}}\left(1-\varepsilon_n\right)\ln\frac{e^{-\frac{y_{ns}^2}{2}}}{\sqrt{2\pi n\sigma^2}}\left(1-\varepsilon_n\right) =$$
$$= \frac{1}{2}\left(1-\varepsilon_n\right)\sum_{s\in M(A)} y_{ns}^2 \frac{e^{-\frac{y_{ns}^2}{2}}}{\sqrt{2\pi n\sigma^2}} + \left(1-\varepsilon_n\right)\ln\left\{\frac{\left(2\pi n\sigma^2\right)^{\frac{1}{2}}}{1-\varepsilon_n}\right\}\times$$
$$\times\sum_{s\in M(A)}\frac{e^{-\frac{y_{ns}^2}{2}}}{\sqrt{2\pi n\sigma^2}} = \frac{1}{2}\left(1-\varepsilon_n\right)S_1 + \left(1-\varepsilon_n\right)S_2\ln\frac{\left(2\pi n\sigma^2\right)^{\frac{1}{2}}}{\left(1-\varepsilon_n\right)},$$

где
$$S_1 = S_1(n,A) = \sum_{s\in M(A)} y_{ns}^2 e^{-\frac{y_{ns}^2}{2}}\left(2\pi n\sigma^2\right)^{-\frac{1}{2}} \underset{n\to\infty}{\longrightarrow} \int_{-A}^{A} z^2\phi(z)dz,$$
$$S_2 = S_2(n,A) = \sum_{s\in M(A)} e^{-\frac{y_{ns}^2}{2}}\left(2\pi n\sigma^2\right)^{-\frac{1}{2}} \underset{n\to\infty}{\longrightarrow} \psi(A) = \int_{-A}^{A}\phi(z)dz.$$

Отсюда
$$\varliminf_{n\to\infty}\frac{2H(\zeta_n)}{\ln n} \geqslant \psi(A). \tag{6}$$
Поскольку при $A\to\infty$ функция $\psi(A)\to 1$, то в силу произвольности числа A в (6), имеем
$$\varliminf_{n\to\infty}\frac{2H(\zeta_n)}{\ln n} \geqslant 1. \tag{7}$$
Из леммы 1 следует, что
$$\varlimsup_{n\to\infty}\frac{2H(\zeta_n)}{\ln n} \leqslant 1. \tag{8}$$
Из (7) и (8) вытекает доказываемое утверждение. Теорема 1 доказана.

Доказательство теоремы 2. Из локальной теоремы (см. [3], § 51) вытекает, что если слагаемое имеет конечный третий момент, то при $n\to\infty$ равномерно по s справедливо соотношение

$$B_n p_n(s) = \phi(y_{ns})\left\{1 + O\left(\frac{y_{ns}^3}{\sqrt{n}}\right)\right\} + o\left(n^{-1}\right), \tag{9}$$

где используются обозначения, введенные при доказательстве теоремы 1. Если s, n меняются так, что $|y_{ns}| \leqslant \sqrt[3]{\ln n}$, то в силу (9) найдется положительная последовательность $\varepsilon_n \to 0$ при $n \to \infty$, для которой справедливы неравенства

$$\phi(y_{ns})(1-\varepsilon_n) \leqslant B_n p_n(s) \leqslant \phi(y_{ns})(1+\varepsilon_n). \tag{10}$$

Введем множество $S_n = \left\{s\colon\ |y_{ns}| \leqslant \sqrt[3]{\ln n}\right\}$. Поскольку

$$H(\zeta_n) \geqslant -\sum_{s\in S_n} p_n(s)\ln p_n(s), \tag{11}$$

то из (10) следует, что

$$H(\zeta_n) \geqslant \sum_{s\in S_n} p_n(s)\ln\left[\frac{\left(2\pi n\sigma^2\right)^{\frac{1}{2}}}{1+\varepsilon_n} e^{\frac{y_{ns}^2}{2}}\right] = a_n + b_n, \tag{12}$$

где

$$a_n = \frac{1}{2}\sum_{s\in S_n} y_{ns}^2 p_n(s), \tag{13}$$

$$b_n = \ln\left[\frac{\left(2\pi n\sigma^2\right)^{\frac{1}{2}}}{1+\varepsilon_n}\right]\sum_{s\in S_n} p_n(s). \tag{14}$$

Очевидно, при $n\to\infty$

$$a_n = \frac{1}{2}(1+o(1)). \tag{15}$$

Для оценки b_n используем неравенство Чебышева, т. е.

$$\Pr\left\{|X-MX| \leqslant \varepsilon\right\} \geqslant 1 - \frac{M\left(|X-MX|^4\right)}{\varepsilon^4},$$

откуда в силу конечности четвертого момента получаем

$$\sum_{s\in S_n} p_n(s) \geqslant 1 - \frac{M\left(|\zeta_n - na|^4\right)}{n^2\sigma^4(\ln n)^{4/3}} = 1 + O\left((\ln n)^{-4/3}\right).$$

Следовательно,

$$b_n \geqslant \left[1 + O\left((\ln n)^{-4/3}\right)\right]\ln\frac{(2\pi n\sigma^2)^{1/2}}{1+\varepsilon_n}. \tag{16}$$

Из (15) и (16) при $n\to\infty$ получим

$$H(\zeta_n) \geqslant \frac{1}{2}\ln n + \frac{1}{2}\ln(2\pi e\sigma^2) + o(1). \tag{17}$$

Соотношения (2) и (17) устанавливают теорему 2.

Для вывода теорем 3–5 нам понадобится

Лемма 2. *Пусть $\xi_1 \geqslant 0$ имеет конечный экспоненциальный момент, т. е.*

$$g(\alpha) = \sum_{s=0}^{\infty} p_1(s) e^{\alpha s} < \infty, \ \text{где } \alpha > 0.$$

Тогда для любого фиксированного натурального числа $N \geqslant 1$ при $n \to \infty$ справедливо асимптотическое разложение

$$\sum_{s=0}^{\infty} p_n(s) \ln s! = an \ln n + an \ln \frac{a}{e} + \frac{1}{2} \ln n + \frac{1}{2} \ln 2\pi a + \sum_{i=1}^{N} \frac{c_i}{n^i} + O\big(n^{-(N+1)}\big), \quad (18)$$

где коэффициенты c_i, $i = \overline{1, N}$, можно выразить через моменты распределения ξ_1.

Доказательство леммы 2. Зафиксируем некоторое число δ, $0 < \delta < a$, и представим интересующую нас сумму из левой части (18) в виде

$$\sum_{s=0}^{\infty} p_n(s) \ln s! = \sum\nolimits_n^{(1)} + \sum\nolimits_n^{(2)} + \sum\nolimits_n^{(3)},$$

где

$$\sum\nolimits_n^{(1)} = \sum_{s:\, s > (a+\delta)n} p_n(s) \ln s!,$$
$$\sum\nolimits_n^{(2)} = \sum_{s:\, s < (a-\delta)n} p_n(s) \ln s!,$$
$$\sum\nolimits_n^{(3)} = \sum_{s:\, |s-na| \leqslant \delta n} p_n(s) \ln s!.$$

Далее мы покажем, что для любого натурального $N \geqslant 1$ при $n \to \infty$ справедливы асимптотические соотношения

$$\left.\begin{aligned} &\sum\nolimits_n^{(1)} = O\big(n^{-(N+1)}\big), \\ &\sum\nolimits_n^{(2)} = O\big(n^{-(N+1)}\big), \\ &\sum\nolimits_n^{(3)} = (an + 1/2) \ln an - an + \frac{1}{2} \ln 2\pi + \sum_{i=1}^{N} \frac{c_i}{n^i} + O\big(n^{-(N+1)}\big), \end{aligned}\right\} \quad (19)$$

из которых очевидным образом вытекает утверждение леммы 2.

Введем функцию действительного параметра α:

$$\mu(\alpha) = \ln g(\alpha) = \ln \sum_{s=0}^{\infty} p_1(s) e^{\alpha s},$$

которая, в силу условия леммы, определена в окрестности $\alpha = 0$. Нетрудно проверить, что $\mu(\alpha)$ выпукла вниз, $\mu(0) = 0$ и $\mu'(0) = a$. Поэтому найдутся числа $\alpha_1 > 0$ и $\alpha_2 < 0$, для которых справедливы неравенства

$$\alpha_1(a+\delta) > \mu(\alpha_1); \quad \alpha_2(a-\delta) > \mu(\alpha_2). \tag{20}$$

Воспользовавшись сначала очевидной оценкой $\ln s! < s \ln s$, а затем монотонностью функции $(s\ln s)e^{-\alpha_1 s}$ при больших значениях s, получаем, что для всех достаточно больших n справедливы неравенства

$$\sum_n^{(1)} < \sum_{s:\, s>(a+\delta)n} p_n(s) e^{\alpha_1 s} \frac{s \ln s}{e^{\alpha_1 s}} <$$
$$< (a+\delta)n \ln\bigl[(a+\delta)n\bigr] \exp\Bigl\{-n\bigl[\alpha_1(a+\delta) - \mu(\alpha_1)\bigr]\Bigr\},$$

где учли также, что для нашего распределения $p_n(s)$ сумма

$$\sum_{s=0}^{\infty} p_n(s) e^{\alpha_1 s} = \exp\bigl\{n\mu(\alpha_1)\bigr\}.$$

Аналогичным образом доказывается, что при любом $\alpha_2 < 0$ для всех достаточно больших n справедливо неравенство

$$\sum_n^{(2)} < (a-\delta)n \ln\bigl[(a-\delta)n\bigr] \exp\Bigl\{-n\bigl[\alpha_2(a-\delta) - \mu(\alpha_2)\bigr]\Bigr\}.$$

Теперь для того, чтобы завершить доказательство первых двух соотношений (19), достаточно в доказанных неравенствах выбрать числа α_i с учетом (20).

Перейдем к выводу третьего соотношения (19). Пусть s удовлетворяет неравенству $|s-an| \leqslant \delta n$. Если $n \to \infty$, то, применяя асимптотическую формулу Стирлинга, получим для любого натурального $N \geqslant 1$ равномерно по s справедливость асимптотических соотношений

$$\ln s! = F_N(s) + O\bigl(n^{-(N+1)}\bigr), \tag{21}$$

где

$$F_N(s) = \left(s + \frac{1}{2}\right) \ln s - s + \frac{1}{2}\ln 2\pi + \sum_{k=1}^{N} \frac{B_k}{s^k}. \tag{22}$$

Напомним, что числа B_k вычисляются, исходя из асимптотического разложения гамма-функции от натурального аргумента [5]. В частности, $B_1 = 1/12$, $B_2 = 0$, $B_3 = -1/360$ и т. д. В силу (21) при $n \to \infty$ имеем

$$\sum_n^{(3)} = \sum_{s:\, |s-an| \leqslant \delta n} p_n(s) F_N(s) + O\bigl(n^{-(N+1)}\bigr).$$

При каждом s, $|s-an| \leqslant \delta n$, запишем значение $F_N(s)$ по формуле Тейлора с остаточным членом в форме Лагранжа

$$F_N(s) = F_N(an) + F'_N(an)(s-an) + \sum_{j=2}^{K} \frac{F_N^{(j)}(an)}{j!}(s-an)^j + R(K;s),$$

$$R(K;s) = \frac{F_N^{(K+1)}(C_s)}{(K+1)!}(s-an)^{K+1},$$

где число C_s лежит между s и an. В силу (22) при $i \geqslant 2$ производная i-го порядка функции $F_N(s)$ имеет вид

$$F_N^{(i)}(s) = \sum_{j=i-1}^{N+i} f_i(j)s^{-j}, \tag{23}$$

где $f_i(j)$, $j = \overline{i-1, N+i}$, — некоторые действительные коэффициенты.

Положим $K = 2N+3$. Поскольку $|s-an| \leqslant \delta n$, то в силу (23) найдется не зависящее от n и s число $R^*(N,\delta)$, такое что

$$\bigl|R(2N+3;s)\bigr| \leqslant R^*(N;\delta)\frac{(s-an)^{2(N+2)}}{n^{2N+3}}.$$

Следовательно,

$$\Biggl|\sum_{s:\,|s-an|\leqslant\delta n} p_n(s)R(2N+3;s)\Biggr| \leqslant \frac{R^*(N;\delta)m_n(2N+4)}{n^{2N+3}},$$

где через

$$m_n(k) \triangleq \sum_{s=0}^{\infty} p_n(s)(s-an)^k, \quad k = 1,2,\dots$$

обозначен k-й центральный момент распределения $p_n(s)$. Известно [6], что для распределения суммы независимых случайных величин $p_n(s)$ момент $m_n(k)$ можно записать в виде многочлена от переменной n степени $\lfloor k/2 \rfloor$ с коэффициентами, зависящими от моментов распределения $p_1(s)$. Это означает, что при $n \to \infty$

$$m_n(k) = O\bigl(n^{\lfloor k/2 \rfloor}\bigr).$$

По аналогии с выводом первых двух соотношений (19) нетрудно понять, что для любых фиксированных $j = 0,1,2,\dots$ и $N = 1,2,\dots$ при $n \to \infty$

$$\sum_{s:\,|s-an|>\delta n} p_n(s)(s-an)^j = O\bigl(n^{-(N+1)}\bigr).$$

Вышеприведенные формулы доказывают, что при $n \to \infty$ исследуемая нами сумма для любого фиксированного натурального $N \geqslant 1$ удовлетворяет соотношению

$$\sum_n{}^{(3)} = F_N(an) + \sum_{j=2}^{2N+3} \frac{F_N^{(j)}(an)}{j!} m_n(j) + O\big(n^{-(N+1)}\big),$$

которое с учетом (22) – (23) завершает доказательство (19). Лемма 2 доказана.

Замечание. Легко понять, что коэффициенты c_i, $i = \overline{1, N}$, асимптотического разложения (18) следует вычислять с помощью приведения подобных членов при одинаковых степенях n в выражении

$$F_N(an) + \sum_{j=2}^{2N+2} \frac{F_N^{(j)}(an)}{j!} m_n(j),$$

где суммирование по j ведется от 2 до $2N + 2$.

Доказательство теоремы 3. Для биномиального распределения

$$p_n(s) \equiv B_n(s, p) = C_n^s p^s (1-p)^{n-s}$$

имеем

$$\begin{aligned} H(\zeta_n) &= -\sum_{s=0}^{n} p_n(s) \ln C_n^s p^s (1-p)^{n-s} = \\ &= nh(p) - \ln n! + \sum_{s=0}^{n} p_n(s) \ln s! + \sum_{s=0}^{n} p_n(s) \ln(n-s)! = \\ &= nh(p) - \ln n! + \sum_{s=0}^{n} B_n(n, p) \ln s! + \sum_{s=0}^{n} B_n(s, 1-p) \ln s!. \end{aligned}$$

Отсюда с помощью асимптотической формулы Стирлинга для $\ln n!$ и леммы 2, применяемой при $N = 3$, после приведения подобных членов получаем требуемое соотношение (4). Теорема 3 доказана.

Теоремы 4 и 5 доказываются аналогичным образом.

Литература

[1] Р. Г. Галлагер. Теория информации и надежная связь. — М.: Сов. Радио, 1974.

[2] A. G. Dyachkov. On a search model of false coins // Colloq. Math. Soc. J. Bolyai. Topics in Information Theory. — 1975. — P. 163–170.

[3] Б. В. Гнеденко, А. Н. Колмогоров. Предельные распределения для суммы независимых случайных величин. — М.: Гостехиздат, 1949.

[4] В. В. Петров. Суммы независимых случайных величин. — М.: Наука, 1972.

[5] В. Н. Сачков. Комбинаторные методы дискретной математики. — М.: Наука, 1977.

[6] М. Дж. Кендал, А. Стьюарт. Теория распределений. — М.: Наука, 1966.

[7] А. Г. Дьячков. О детектирующих матрицах // 4$^{\underline{\text{ый}}}$ Международный симпозиум по теории информации. Ч. 1. — Ленинград: 1976.

Статья поступила в редакцию в апреле 1996 г.