

ПРИРОДА

№ 11, 2004 г.

Б.А.Трубников, О.Б.Трубникова

Пять великих распределений вероятностей

© “Природа”

Использование и распространение этого материала
в коммерческих целях
возможно лишь с разрешения редакции



Сетевая образовательная библиотека “VIVOS VOCO!”
(грант РФФИ 03-07-90415)

vivovoco.nns.ru
vivovoco.rsl.ru
www.ibmh.msk.su/vivovoco

Пять великих распределений вероятностей

Б.А.Трубников, О.Б.Трубникова

«Кажется, дождь собирается, кажется, дождь собирается», — приговаривал Пятачок, стараясь отогнать пчел от улья, куда хотел добраться Винни-Пух, чтобы полакомиться медом. Так Пятачок хотел передать пчелам информацию о том, что вероятность дождя велика. Спонятием вероятности, хотя бы на бытовом уровне, сталкивались все читатели нашего журнала, например играя в «орла» и «решку». Однако теория вероятностей — весьма точная наука, имеющая несколько замечательных достижений, с которыми мы и хотим познакомить читателей.

Азбука теории

Основные положения теории вероятностей достаточно просты. Если могут случиться два неких события «1» и «2» с вероятностями $w(1)$ и $w(2)$, причем эти события независимы («в огороде бузина, а в Киеве дядька»), то выполняются следующие соотношения:

— вероятность того, что произойдут оба независимых события, равна произведению $w_{1,2} = w(1)w(2)$;



Борис Андреевич Трубников, доктор физико-математических наук, профессор, главный научный сотрудник Института ядерного синтеза РНЦ «Курчатовский институт». Область научных интересов — теория плазмы, гидродинамика, происхождение космических лучей. Неоднократно публиковался в журнале «Природа».

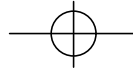


Оксана Борисовна Трубникова, младший научный сотрудник лаборатории экспериментальной эмбриологии им.Д.П.Филатова Института биологии развития РАН им.Н.К.Кольцова. Изучает временные зависимости процессов деления клеток.

— вероятность того, что произойдет либо первое, либо второе событие, равна сумме $w(1+2) = w(1) + w(2)$;

— вероятность того, что произойдет какое-либо одно событие из полного набора N всех возможных независимых событий, равная сумме всех вероятностей $W_N = w(1) + w(2) + w(3) + \dots + w(N)$, составляет единицу (или 100%, если вероятности исчисляются в процентах).

В качестве примера используем эти простые формулы для анализа движения молекул газа. Пусть $w(E_1)$ есть вероятность того, что молекула обладает полной энергией E_1 , и $w(E_2)$ есть вероятность того, что другая молекула обладает энергией E_2 . Тогда произведение $w(E_1)w(E_2) = w(E_1 + E_2)$ будет вероятностью того, что имеют место



СТАТИСТИКА

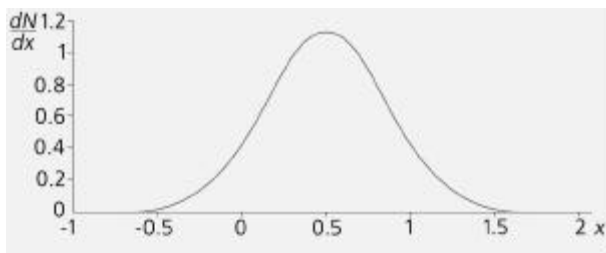


Рис.1. Пример нормированного на единицу распределения Гаусса с параметрами $B = 4$ и $\langle x \rangle = 0.5$.

сразу оба события.

Решением этого функционального уравнения, которое имеет вид $F(x + y) = F(x)F(y)$, является экспонента

$$w(E) = \exp(a - bE), \quad (1)$$

где a и b — постоянные, и полученную формулу называют распределением Максвелла—Больцмана. Если пренебречь потенциальной энергией, обусловленной силой тяжести, то полная энергия молекулы сводится к ее кинетической энергии $E_{кин} = mv^2/2$, и тогда формула $w = a \exp(-\beta v^2)$, где $\beta = bm/2$, описывает распределение молекул по скоростям.

Для простоты рассмотрим случай, когда молекулы газа движутся вдоль оси x со средней скоростью течения $\langle v \rangle$. Тогда число молекул dN , у которых скорость v заключена в интервале от v до $v + dv$, будет равно

$$dN = a \exp[-\beta(v - \langle v \rangle)^2] dv, \quad (2)$$

где a — множитель нормировки.

Забыв про молекулы, запишем сходную общую формулу вида

$$dN = A \exp[-B(x - \langle x \rangle)^2] dx, \quad (3)$$

где A , B — постоянные.

Эту функцию (рис.1) называют распределением вероятностей Гаусса, и она имеет много различных применений.

Если, например, мы много раз измерим как можно точнее некую величину x , то найдем, что в среднем она равна $\langle x \rangle$. Но каждый раз наблюдается некоторое отклонение $x - \langle x \rangle$ от среднего значения, и тогда dN — это число попыток, при которых измеренное значение лежит в интервале от x до $x + dx$.

А теперь займемся решением квартирного вопроса, который, по словам Булгакова, так испортил москвичей.

В село въехали гусары...

Предположим, что полк гусар, участвуя в маневрах, въехал в большое село, и N гусар надо посе-

лить на постоя в K квартир. Спрашивается: каким числом способов S это может сделать начальник-квартирмейстер (теперь их называют «риэлтерами», т.е. реализаторами квартирных возможностей).

Для наглядности можно изображать людей N точками в круге-доме, разделенном перегородками на K секторов. При расчете вероятностей возможны три разных варианта.

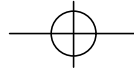
Одного человека можно поселить в одну из K квартир, очевидно, K способами, а N человек — K^N способами. Но среди этих способов есть тождественные, отличающиеся лишь перестановками людей; их мы считаем равноправными. Число таких перестановок равно факториалу $N!$, и если их исключить, то всего получается $C_1 = K^N/N!$ способов размещения.

Теперь усложним задачу, считая, что это не просто поход гусар, а штабные учения, на которые съехались N генералов и их надо разместить по K квартирам. Но генералы не могут жить вместе, и каждому генералу нужна отдельная квартира. Для такой возможности число квартир K должно превышать число людей N , так что $K - N$ квартир останутся пустыми. Нетрудно сообразить, что в этом случае число возможных способов расселения будет равно биномиальному коэффициенту $C_2 = K!/(K - N)!N!$, поскольку $K!$ — это число перестановок квартир, $N!$ — число перестановок генералов, а $(K - N)!$ — число перестановок пустых квартир.

Наконец, рассмотрим вариант, когда квартир мало, и гусары согласились на то, что в одну квартиру квартирмейстер может поселить сколько угодно людей. Возвращаясь к наглядному образу круга, в котором есть K секторов с K перегородками между ними и N точек (людей), всего имеем $N + K$ «субъектов и объектов». Полное число возможных перестановок и тех и других, очевидно, равно факториалу $(K + N)!$. Но, во-первых, среди них есть $K!$ взаимных перестановок только перегородок, и, во-вторых, $N!$ взаимных перестановок только людей не меняют общую схему расселения. И если исключить такие перестановки, то число возможных способов расселения в этом третьем случае будет равно дроби $C_3 = (K + N)!/K!N!$.

Полезно отметить: если число квартир значительно превышает число людей ($K \gg N$), последние два варианта сводятся к первому. Это легко увидеть, если во втором случае приближенно положить $K! \approx (K - N)!K^N$, а в третьем — $(K + N)! \approx K!K^N$, что и приведет к первому случаю расселения гусар.

Поясним также, что отличие чисел $C_{2,3}$ от C_1 здесь получено в результате простой игры в перестановки N точек и K перегородок в круге. На самом деле и сама игра, и три варианта понадобились нам, чтобы учесть фундаментальное свойство тождественности элементарных частиц мате-



рии — фермионов и бозонов. Это чисто квантовое свойство трудно передать на «языке гусар», поскольку никакой квартирмейстер не будет учитывать число возможных перестановок пустых квартир, а для наших дальнейших целей их важно принять во внимание.

Формула Стирлинга и квазиэнтропия

Будем считать числа N и K большими и для факториалов использовать известную приближенную формулу Стирлинга. Она получается, если сначала записать логарифм факториала в виде суммы $\ln M! = \ln 1 + \ln 2 + \dots + \ln M$, а затем заменить сумму на интеграл, что приближенно дает $M! \approx (M/e)^M$, где $e = 2.718281828\dots$ — основание натуральных логарифмов.

Тогда найденные выше три числа C_i ($i = 1, 2, 3$) можно записать в виде приближенных единообразных формул

$$(4) \quad C_i \cong f_i^{-K}, f_1 = \left(\frac{n}{e}\right)^n, f_2 = n^n(1-n)^{1-n}, f_3 = \frac{n^n}{(n+1)^{n+1}}.$$

Здесь мы ввели аргумент $n = N/K$, который во всех трех случаях означает среднее число людей (частиц), приходящихся на одну квартиру (состояние).

Поскольку числа C_i велики, удобнее оперировать с их натуральными логарифмами $S_i = \ln C_i = -K \ln f_i$. В статистической физике логарифм числа способов распределения частиц по состояниям («квартирам») принято называть энтропией (от греч. *entropia* — переворот, перестановка, превращение — введено Р.Клаузиусом в 1865 г.). Мы здесь условимся вместо термина «энтропия» применять более общий термин «квазиэнтропия», т.е. логарифм большого числа способов распределения множества каких-либо объектов (или субъектов) по каким-либо определенным признакам. Например, слов по частоте их встречаемости в тексте и пр.

Очевидно, что для свободы выбора квартирмейстер-риэлтер хотел бы иметь в своем распоряжении как можно большее число способов размещения людей по квартирам, максимальную «квазиэнтропию». Распределение, для которого число способов реализации максимально, действительно самое интересное — оно будет наиболее вероятным. Но при поиске максимума C_i следует учитывать ряд ограничений.

Не числом единым

Далее будем считать, что квартиры различаются своим качеством, которое будем обозначать

буквой ϵ (от англ. *excellent* — превосходный), полагая, для конкретности, что оно характеризует бытовые удобства в данной квартире (в первую очередь число комнат, а также наличие телефона, ванны, холодильника, кондиционера и пр. — присутствие каждого из перечисленных удобств добавляет лишний балл к качеству ϵ). Разумно считать, что введенное выше среднее число людей в одной квартире зависит от аргумента ϵ , так что $n = N/K = n(\epsilon)$ (поскольку хорошие квартиры дорого стоят и доступны немногим).

Ранее мы рассматривали примеры с определенным набором людей и квартир (N_m, K_m). Эту совокупность назовем малым набором и, присвоив ему индекс m , запишем его квазиэнтропию в виде $S_m = -K_m \ln f^{(m)}$. Далее рассмотрим большой набор, состоящий из множества малых m -наборов. Очевидно, в большом наборе теперь будет $N = \sum N_m = \sum n_m K_m$ людей (частиц), и мы будем считать фиксированным их полное число N .

Большой набор необходимо ввести в рассмотрение по той причине, что теперь мы условились учитывать не только количество, но и качество квартир ϵ_m (число их бытовых удобств). И суммарное число бытовых удобств по всему большому набору $E = \sum \epsilon_m N_m = \sum \epsilon_m n_m K_m$ мы также будем считать заданным. В статистической физике частиц качеству ϵ_m квартиры соответствует энергия m -состояния, а n_m приобретает смысл среднего числа частиц в m -состоянии. Таким образом, задание фиксированного полного числа бытовых удобств соответствует в физике частиц постоянству их суммарной полной энергии.

Далее замечаем, что число способов C реализации большого набора равно произведению чисел способов реализации всех малых наборов. Поэтому наша логарифмическая квазиэнтропия $S = \ln C$ большого набора равна сумме квазиэнтропий всех малых наборов:

$$S = \sum S_m = -\sum K_m \ln f^{(m)}. \quad (5)$$

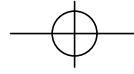
И мы хотели бы найти максимум этой величины при двух дополнительных ограничительных условиях — при постоянстве полного числа людей (частиц) и постоянстве суммарного числа всех бытовых удобств (в физике — полной энергии большого набора), т.е. при условиях

$$N = \sum n_m K_m = \text{const}, E = \sum \epsilon_m n_m K_m = \text{const}. \quad (6)$$

Путем Лагранжа — к трем распределениям

Максимум некоторой функции при двух дополнительных условиях можно найти, пользуясь так называемым методом неопределенных множителей Лагранжа. Для этого составляем комбинацию (здесь $i = 1, 2, 3$)

$$\Phi = S + |\alpha| N - |\beta| E = -\sum K_m (\ln f_i + \sigma_m n_m), \quad (7)$$



СТАТИСТИКА

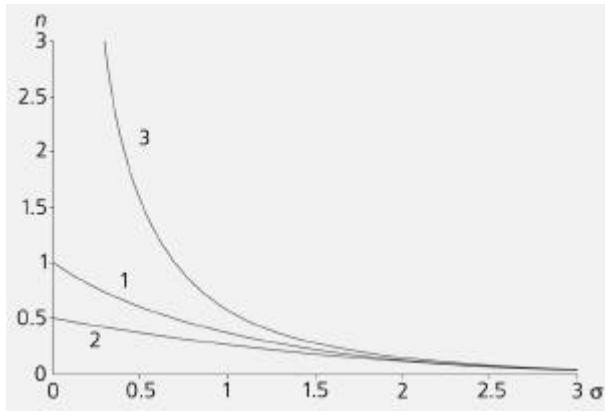


Рис.2. Распределения: $n_1(s)$ — Максвелла—Больцмана; $n_2(s)$ — Ферми—Дирака и $n_3(s)$ — Бозе—Эйнштейна.

где для краткости обозначено $\sigma_m = |\beta|\epsilon_m - |\alpha|$. И далее по рецепту Лагранжа решаем уравнения $-\partial\Phi/\partial n_m = 0$. Опуская для простоты индекс m , запишем эти уравнения в виде

$$(\ln f_i)' = -\sigma, \quad (i = 1, 2, 3), \quad (8)$$

где штрих означает производную по n . Нетрудно проверить (предоставим это читателю), что подставляя сюда три выражения (4) для трех функций $f_{1,2,3}(n)$ и вычисляя производные их логарифмов, найдем три распределения (рис.2)

$$n_1 = \frac{1}{e^\sigma}, \quad n_2 = \frac{1}{e^\sigma + 1}, \quad n_3 = \frac{1}{e^\sigma - 1}, \quad (9)$$

где $\sigma = |\beta|\epsilon - |\alpha|$, $|\beta| > 0$ и $|\alpha| > 0$ — положительные параметры и ϵ — качество квартир.

В квантовой статистической физике величина n_m интерпретируется как среднее число частиц газа в m -состоянии, обладающем энергией ϵ_m , а комбинация σ записывается в виде $\sigma = (\epsilon - \mu)/kT$, где μ — химический потенциал газа, k — постоянная Больцмана и T — температура в градусах Кельвина.

Формулу для n_1 называют распределением Максвелла—Больцмана, для n_2 — распределением Ферми—Дирака для фермионов, а для n_3 — распределением Бозе—Эйнштейна для бозонов. Имеются в виду распределения газов из классических частиц (n_1) и квантовых ферми- (n_2) или бозе-частиц (n_3) по состояниям с разными энергиями ϵ .

При больших энергиях и температурах все три распределения совпадают (рис.2).

Какие города лучше?

Штабные учения с несколькими десятками генералов и даже полк гусар в несколько сотен сабель — все же не слишком большие множества, и в них могут не проявиться характерные статис-

тические закономерности, свойственные большим множествам.

Поэтому в качестве большого множества рассмотрим некую страну с населением N человек, проживающих в K городах разного типа. Предположим, что в некотором городке проживает m людей и в стране имеется n_m городков с таким же числом жителей. Тогда, очевидно, имеют место выражения с суммами

$$K = \sum n_m = \text{const}, \quad N = \sum m n_m = \text{const}. \quad (10)$$

Эти суммы весьма похожи на два дополнительных условия (6), встречавшихся в трех ранее рассмотренных распределениях квартир по их качеству (или же элементарных частиц материи, где требовалось сохранение полного числа частиц N и их суммарной энергии E).

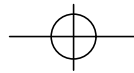
Сравнение пар уравнений (10) и (6) показывает, что показателем качества ϵ_m города следует считать просто число его жителей m . И если людей не расселяет по городам начальник-квартирмейстер, то сами люди, как правило, стремятся перебраться из мелких городов в более крупные, где больше возможностей найти работу и дать детям хорошее образование.

Лишенные разума молекулы газа как бы сами (но под «руководством» теории вероятностей) «стремятся» создать наиболее хаотическое распределение по энергиям с наибольшим числом возможных способов C его воплощения и, тем самым, с наибольшей энтропией $S = \ln C$. Таким же правилом руководствуется и квартирмейстер-риэлтер, стремящийся иметь как можно более высокую квазиэнтропию, т.е. логарифм числа возможных способов расселения N людей по K квартирам с учетом показателя их качества.

И наоборот, разумно считать, что $N = \sum m n_m$ людей, расселяясь по $K = \sum n_m$ разным городам, в отличие от «глупых» молекул сами (как бы даже бессознательно, поскольку не знают наших формул) стремятся создать наименее хаотическое, но наиболее удобное для них распределение по городам.

Число возможных способов его реализации должно иметь вид дроби $C = k_0/k_1 k_2$, где числитель, очевидно, равен $k_0 = N!$ — числу всех возможных взаимных перестановок местами проживания всех людей страны. При этом, однако, надо исключить случаи, когда m жителей одного города меняются местами, не покидая этот город. Число таких перестановок равно факториалу $m!$, а поскольку число городов с m жителями равно n_m , разумно считать, что первый множитель в знаменателе равен произведению $k_1 = \prod (m!)^{n_m}$.

Второй множитель k_2 в знаменателе можно получить из следующих, столь же очевидных соображений. Выше мы уже заметили, что показателем престижности города служит само число m его жителей. Далее замечаем, что $m n_m$ есть число жителей всех городов m -ого типа, т.е. городов с населением m . Число взаимных перестановок таких



жителей городов m -ого типа, очевидно, равно факториалу $(m n_m)!$, а для всех городов с разными m число перестановок всех жителей страны равно произведению факториалов $k_2 = \Pi(m n_m)!$.

Ему и следует приравнять второй множитель в знаменателе, чтобы исключить перестановки людей, живущих в равноценных городах и, следовательно, не имеющих причин, желания и времени заниматься переездами именно такого рода. Таким образом, для обсуждаемой здесь проблемы «кто где живет» число возможных способов распределения $N = \sum m n_m$ людей по $K = \sum n_m$ городам следует считать равным

$$C = \frac{k_0}{k_1 k_2} = \frac{N!}{\Pi[(m!)^{n_m} (m n_m)!]}. \quad (11)$$

Конкурентная борьба

Если число городов ограничено, люди явно начинают конкурировать друг с другом, и их можно назвать конкурентами по проблеме переездов в более крупные города. Используя формулу Стирлинга, перепишем формулу (11) для числа способов в приближенном виде

$$C \approx \frac{(N/e)^N}{\Pi(m^2 n_m / e^2)^{m n_m}}, \quad (12)$$

и далее найдем квазиэнтропию, т.е. логарифм этого числа способов

$$S \approx \sum m n_m (1 + \ln N - 2 \ln m - \ln n_m). \quad (13)$$

При этом мы учли, что число людей выражается формулой $N = \sum m n_m$.

Как уже пояснялось ранее, в отличие от поведения «глупых» молекул с максимальной хаотизацией их энергетических состояний, у сообщества всех разумных живых существ — не только людей, но и животных и даже бактерий океанского планктона — можно ожидать стремления к минимуму такой квазиэнтропии. По-видимому, это можно рассматривать как закон жизни сообществ живых существ — биоценозов: их организация не может быть хаотической, а должна быть максимально упорядоченной и построенной по строго иерархическим, но естественным принципам. С аналогичной ситуацией мы сталкиваемся в кибернетике, где информация рассматривается как отрицательная энтропия и считается, что полный хаос лишен информации.

Минимум квазиэнтропии (экстремум, как и максимум) должен достигаться при выполнении двух дополнительных условий (10). Поэтому по рецепту Лагранжа вновь составляем комбинацию с неопределенными множителями a и b :

$$\begin{aligned} \Phi &= S + a K + b N = \\ &= \sum m n_m [m (1 + \ln N - 2 \ln m - \ln n_m) + \sigma], \end{aligned} \quad (14)$$

где $\sigma = a + b m$ не зависит от n_m . И тогда нетрудно проверить (предоставим это читателю), что уравнения Лагранжа $\Phi'_{n_m} = 0$ приводят к соотношению $\ln(m^2 n_m) = b + \ln N + (a/m)$. Из него мы сразу получим для числа городов с населением m

$$n_m = \frac{A}{m^2} \exp \frac{a}{m}, \quad (15)$$

где $A = N e^b$ — постоянная нормировки. Часто представляет интерес так называемый интегральный спектр, т.е. число городов с населением больше данного:

$$N(m) = \int_m^{m_{\max}} n_m dm = \frac{A}{a} \left(\exp \frac{a}{m} - \exp \frac{a}{m_{\max}} \right). \quad (16)$$

Мы пришли к так называемому распределению конкурентов, характерному для объектов, которые соревнуются между собой за обладание неким ограниченным ресурсом [1]. В пределе $a \rightarrow 0$ это распределение переходит в чисто гиперболическое

$$N(m) = A \left(\frac{1}{m} - \frac{1}{m_{\max}} \right). \quad (17)$$

В качестве иллюстрации формулы (17) на рис. 3 приведены распределения городов США по их населенности, выявленные за 140 лет общего роста населения страны. Схорошей точностью они следуют гиперболическому закону с зависящим от времени t коэффициентом $N(m) = A(t)/m$.

Для некоторых других множеств «конкурентов»

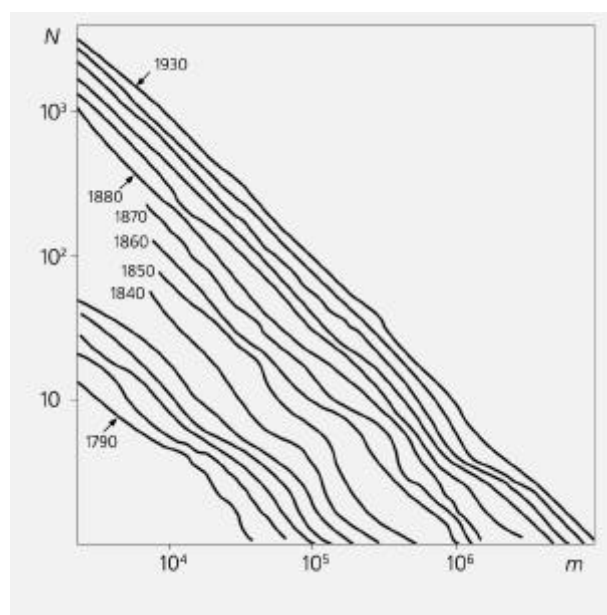
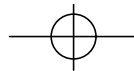


Рис.3. Интегральные спектры распределения числа городов США по числу их жителей с 1790 по 1930 г. через каждые 10 лет.



СТАТИСТИКА

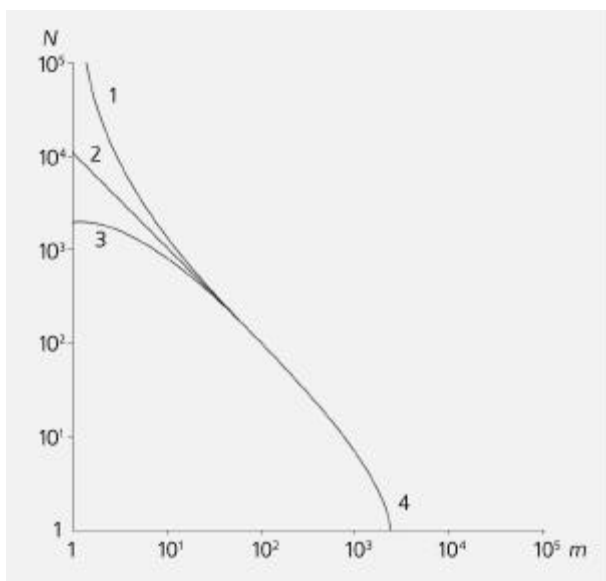


Рис.4. Три примера интегральных спектров общего распределения конкурентов, построенных по формуле (16). Выбрана нормировка $A = 10^4$. Параметр a определяет левый конец спектра, параметр $m_{\max}$ — правый конец. Показаны варианты $a = 5$ (1); 0.1 (2); -5 (3) и завал правого конца при $m_{\max} = 10^{3.5}$.

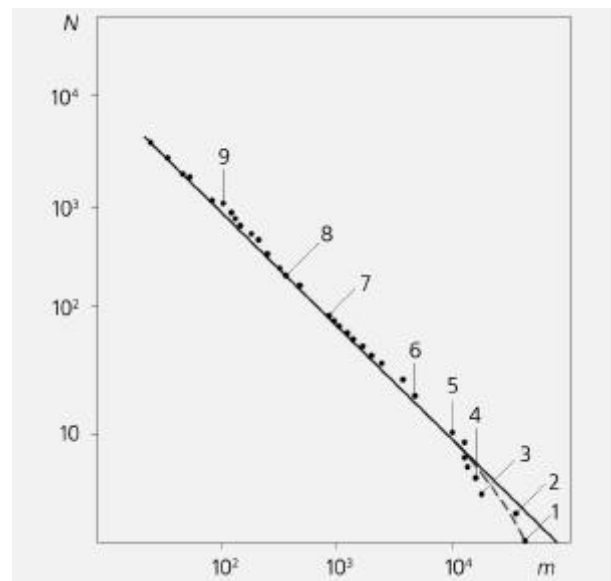


Рис.5. Интегральный спектр распределения слов по частоте их встречаемости в тексте из 1 056 382 слов. 1 — в, 2 — 4, 3 — не, 4 — на, 5 — а, 6 — как, 7 — каждый, 8 — огромный, 9 — красота.

тов» истинные (наблюдаемые) интегральные спектры могут несколько отличаться от чисто гиперболических. В этих случаях для сравнения наблюдаемых спектров с предсказываемыми «теорией конкурентов» следует пользоваться полной формулой (16), содержащей три параметра A , a и $m_{\max}$. Подбором параметра a можно подправить левый конец спектра, а подбором параметра $m_{\max}$ — правый конец, как это показано на рис. 4.

Формулы (15–16) применимы не только к городам, но и к другим множествам «конкурентов», частными случаями которых рассматривались и ранее многими авторами. Так, гиперболическое распределение слов по частоте их встречаемости в тексте обычно называют законом Эсту—Ципфа—Мандельброта (рис.5). Распределение богатей по их богатствам в древних государствах (да и в современных) принято называть законом Парето. Распределение ученых по числу опубликованных ими статей — законом Альфреда Лотки. Чтобы не замыкаться на человеческой деятельности, проиллюстрируем указанную зависимость картинками для океанического биоценоза (рис.6) и малых космических тел (рис.7). Около 500 различных примеров гиперболических распределений приведено в книге [2]. Поэтому распределение конкурентов вполне заслуженно можно поставить в один ряд с четырьмя великими распределениями. Обычно обсуждаются лишь эмпирические закономерности конкретных множеств конкурентов,

хотя попытки теоретического обоснования распределения конкурентов также предпринимались и ранее. Краткий обзор этих усилий приведен, в частности, в [3]. Но, по нашему мнению, многие из этих попыток не были столь последовательными, как изложенный выше подход, основанный, по существу, на полной аналогии с термодинамикой частиц, достигнутой путем введения общего понятия качества и квазиэнтропии множеств конкурентов. Ряд примеров множеств конкурентов приведен в наших работах [1, 3, 4], где рассмотрено около 20 примеров таких распределений.

Отметим, что такие множества должны состоять из достаточно большого числа объектов, и (следуя А.В.Бялко [5]) их удобно называть «широкими распределениями», в отличие от сравнительно «узких» распределений гауссовского типа. В них по существу отсутствует понятие среднего значения, вокруг которого группировалась бы основная масса объектов. Этим свойством множеств конкурентов резко отличаются от множеств, подчиняющихся нормальному распределению Гаусса $n_m \sim \exp[-a(m - \langle m \rangle)^2]$. В чисто гиперболическом распределении (при $a \rightarrow 0$) полная «масса» множества равна интегралу $M = \int m n_m dm = A \ln(m_{\max}/m_{\min})$. Как видим, она одинаковым логарифмическим образом определяется и верхним, и нижним пределами спектра. Эта слабая логарифмическая зависимость от обоих пределов ука-

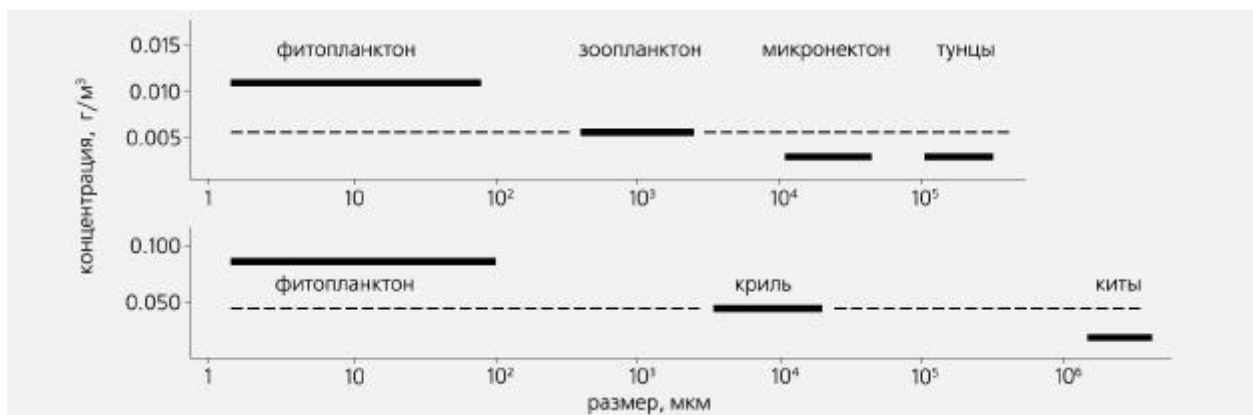
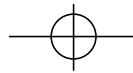


Рис.6. Распределение обитателей Мирового океана по размерам (данные канадских ученых). По вертикальной оси указана концентрация «живого вещества» с размерами из данного интервала (размер оценивается по диаметру шара, равного по объему рассматриваемой особи). Все организмы распределены так, что на каждый десятичный интервал приходится практически одинаковый суммарный вес. Это соответствует гиперболическому распределению (17).

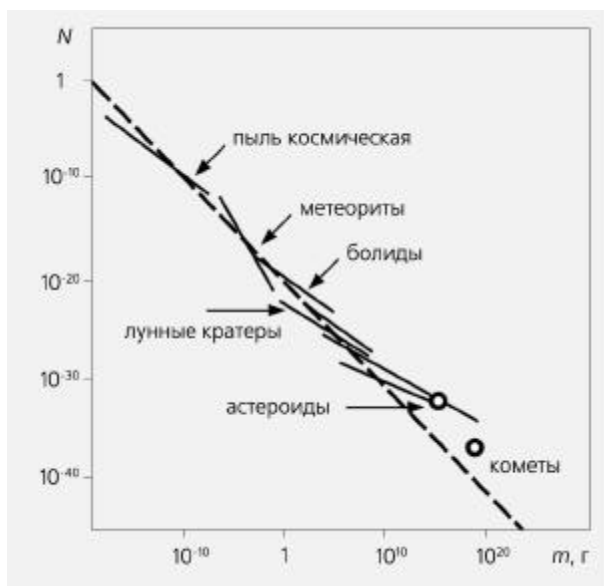


Рис.7. Интегральный спектр распределения малых космических тел по их массам, различающимся на 30 порядков. N — число данных тел в 1 см^3 .

зывает на устойчивость таких распределений. Подобное свойство оказывается особенно важным для различных социально-экономических проблем, в целом называемых «макроэкономикой» (распределение доходов, структура различных фирм и пр.).

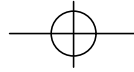
Во всех случаях можно указать пару дополнительных условий, при которых вычисляется экстремум квазиэнтропии. Для перечисленных примеров ими будут: длина текста и длина словаря;

сумма всех финансов и число владельцев; число всех выбранных для анализа ученых и массив статей; число всех рыб (на кубометр воды) и число рыб с определенным весом в стандартных логарифмических интервалах; число всех космических тел, отобранных для анализа (число видов тел здесь не учитывалось, что соответствует пределу $a \rightarrow 0$ в формулах).

Во что играет природа?

Конечно, к формуле (16), содержащей три параметра, можно относиться и просто как к удобной аппроксимационной формуле (как говорится — не более того). Однако полезно подчеркнуть, что и все четыре «главных» для неживой природы распределения также содержат по три параметра. Для распределения Гаусса — это коэффициент нормировки, среднее значение $\langle x \rangle$ и $1/\beta$ — среднеквадратичное отклонение (полуширина кривой Гаусса). Для распределений $p_{1,2,3}$ — это параметры α , β и коэффициент нормировки.

Заметим также, что все три главных для физики частиц распределения $p_{1,2,3}$ в сущности, получаются из «игры в перестановки» всего лишь с двумя числами N и K , точнее со слагающими их частями p_m , комбинируемыми тремя разными способами. Образно можно сказать, что именно простота и естественность их трех комбинаций с перестановками приводят к столь фундаментальной роли этих трех распределений в неживой природе. Аналогично этому и четвертое распределение — распределение конкурентов — получается из «игры в перестановки» с теми же двумя числами и теми же составляющими их частями p_m . По нашему мнению, это позволяет считать по-



СТАТИСТИКА

следнее столь же основополагающим для природы, как и три первых распределения. По-видимому, природа (или теория вероятностей) просто не может «придумать» какую-либо еще одну новую «игру в перестановки» с теми же числами N , K и их частями p_m .

Здесь, конечно, имеется в виду новая «игра», которая имела бы какой-нибудь разумный смысл. Например, простая принудительная (не обоснованная разумными соображениями) перестановка числителей со знаменателями во всех трех фундаментальных случаях $p_{1,2,3}$ не изменит их окончательных результатов, поскольку это привело бы лишь к изменению знаков логарифмических членов в энтропии, но ее экстремум (с заменой максимума на минимум) дал бы те же распределения $p_{1,2,3}$ (Быть может, наши читатели сами попробуют изобрести новую игру в перестановки с теми же числами p_m ?)

Именно указанная фундаментальность распределения конкурентов (как результат игры больших чисел), по-видимому, и приводит к не-

обычайно широкой его распространенности как в живой, так и в неживой природе. С этой точки зрения, например, часто используемое так называемое логнормальное распределение (распределение Гаусса с логарифмическими аргументами) можно рассматривать лишь как удачную аппроксимацию, поскольку оно не выводится из каких-либо общих положений, подобных использованным нами. Иногда считают (см., например [6]), что распределение Гаусса (3) описывает ситуацию с множествами аддитивных величин (с суммой вероятностей), а логнормальное распределение относится к множествам мультипликативных величин. Но ведь в логнормальном распределении фигурируют не просто логарифмы, а квадраты логарифмов в экспоненте, и эти логарифмы введены искусственно, просто для «охвата» широких интервалов аргументов. Поэтому в логнормальное распределение нельзя разумным образом ввести учет двух дополнительных условий (6) и (10), составленных из сумм. Так что мы вполне справедливо поставили в ряд со знаменитыми

Литература

1. Трубников Б.А. Закон распределения конкурентов // Природа. 1993. №11. С.3—13.
2. Кудрин Б.И. Введение в технетику. Томск, 1993.
3. Трубников Б.А., Румынский И.А. // Докл. АН СССР. 1991. Т.321. №2. С.270.
4. Трубников Б.А. О законе распределения конкурентов // Природа. 1995. №11. С.48—50.
5. Бялко А.В. Конструктивность закона конкуренции // Природа. 1993. №11. С.14—19.
6. Карасев Б.В. Логарифмически-нормальное распределение // Природа. 1995. №11. С.41—48.

Космические исследования

Солнце вредит спутникам

4 ноября 2003 г. на Солнце произошла мощная вспышка, вызвавшая сильнейший поток частиц высокой энергии. Она была завершающей в серии ей подобных, происходивших в двух чрезвычайно активных областях Солнца¹. Эти явления заметно повлияли на физическую обстановку в окрестностях Земли, что зафиксировали, в частности, приборы на борту спутника «GOES». Потоки высокоэнергетичных частиц были настолько интенсивными, что бортовые приборы ИСЗ «ACE»

¹ Что случилось с Солнцем? // Природа. 2004. №10. С.80.

(«Advanced Composition Explorer»), предназначенного для изучения химического состава околоземной среды, на время зашкалило, и измерения прервались.

Из Японии пришли сообщения, что с ее спутником «ADEOS-2» была прервана связь; временно она была нарушена и с японским спутником «Coda-ma», используемым для ретрансляции солнечно-геомагнитных данных с других ИСЗ. Аналогичные трудности испытывало оборудование спутника «SMART» («Small Missions for Advanced Research in Technology»), запущенного на переходную орбиту Европейским космическим агентством.

У ряда коммерческих геостационарных спутников-ре-

трансляторов телевизионных сигналов снизилось энергопитание из-за повреждения солнечных батарей частицами высоких энергий.

Earth System Monitor. 2003. V.14. №1. P.3 (CША).

Следить за «погодой» на Солнце

Европейское космическое агентство намерено в начале 2006 г. вывести на околоземную орбиту миниатюрный спутник «PROBA-2» («Project for On-board Autonomy» — «Проект бортовой автономии»), которому предстоит стать родоначальником спутников, способных автономно принимать решения в зависимости от обстановки в окружа-