

ОБ ОДНОЙ ПРИРОДНОЙ ЗАКОНОМЕРНОСТИ

Ю.Г.Дмитриев, Ф.П.Тарасенко

Томский государственный университет
Отделение проблем информатики ТНЦ СО РАН

dmit@mail.tsu.ru, ftara@sibmail.com

Аннотация. Излагается математическая модель роста систем за счёт накопления. Из модели вытекает, что параметры частей таких систем связаны определёнными закономерностями, которые могут использоваться в практических целях. Показывается, что к рассматриваемому классу относятся реальные системы разной природы.

Ключевые слова: природные закономерности, растущие системы, квантили распределения.

1. Введение : элементы системной картины мира

Мы живем в бесконечно разнообразном и постоянно меняющемся мире. Стараясь не просто выжить, но и улучшить своё положение в нём, мы используем свои способности и особенности окружающей нас среды для достижения наших целей. Средством для этого служат *модели*, т.е. знание о той ситуации, в которой мы оказались. Несомненно, самой важной информацией для овладения процессом хоть какого-то контроля над происходящими с нами событиями, является информация о том, каковы наши возможности вносить изменения в неудовлетворяющую нас ситуацию. Такую информацию даёт нам *системная картина мира*, которая описывает его бесконечное разнообразие с помощью понятия *системы* (см., например, [1], гл. 2; [2], гл. 1)..

Системой называется некоторая сущность (реальная или абстрактная), обладающая некоторыми присущими именно ей свойствами, особенностями, функциями (называемыми *эмерджентными*, или *синергетическими*), отличающими её от всего остального (которое называется *окружающей средой*). Мир представляется неограниченной совокупностью различных систем, взаимосвязанных и взаимодействующих между собой, разномасштабных, соседствующих или содержащих в себе системы меньшего масштаба, или входящих как часть в большие системы, непрерывно изменяющихся с различными темпами и охватами изменений.

Столь сложную картину удастся описать с помощью всего трёх моделей: модели *черного ящика* (перечня существенных связей системы с окружающей средой), модели *состава* (перечня существенных частей системы), и модели *структуры* (перечня существенных связей между частями системы). Оценочное (т.е. субъективное) слово «*существенные*» необходимо в связи с тем, что в силу закона о всеобщей взаимозависимости в природе число связей неограниченно, но для конкретной цели рассмотрения системы решающее значение для нас имеют лишь самые сильные из связей.

Но и слово «*сильные*» -- оценочное, т.е. субъективное, с градациями. Например, если связь от одной сущности к другой является необходимой и достаточной, чтобы при осуществлении первой обязательно произошла вторая, -- мы говорим о «причинно-следственной» связи (например, смерть есть следствие рождения). Если связь необходима, но не достаточна (т.е. для реализации второй сущности нужны и другие сущности), мы говорим об отношении «производитель-продукт» (например, жёлудь есть производитель дуба). Если связь достаточна, но не необходима, мы говорим о связи косвенной (статистической, корреляционной, регрессионной и пр.). Например, связь между курением и раком лёгких.

Но и эти все понятия связаны с рассмотрением только однонаправленных связей двух сущностей. В реальности, в силу закона о всеобщей взаимозависимости, конкретные проявления определяются соотношением прямых и обратных связей. А так как между любыми двумя сущностями существует множество разнонаправленных и «разносильных» связей, нам приходится иметь дело с реальной (и неисчерпаемой) сложностью природы, которую мы описываем с помощью понятия *случайности*, и «распутываем», открывая законы природы.

Разнообразие систем упрощённо описывается в виде их классификаций. Согласно одной из классификаций, системы делятся на естественные и искусственные. Искусственные системы проектируются, конструируются и реализуются субъектом для того, чтобы их использовать как средства

достижения его целей. Естественные системы являются сущностями, возникающими в реальности в силу проявления тех взаимодействий, которые мы называем законами природы.

Обнаружение закономерностей природы является основной познавательной функцией фундаментальной науки, и этим занимаются учёные во всех областях знаний. Существуют даже специальные отрасли науки, исследующие вопрос, *как* выявлять и уточнять закономерности природы – от разработки очень общих подходов (теория познания), до конкретных технологических (теория измерений и подобию), технических (теория моделирования), или алгоритмических (анализ данных) вопросов.

И всё же обнаружение ещё одной закономерности, да ещё и общесистемной, универсальной, имеющей место для широкого круга систем различной природы, -- можно считать несомненной удачей исследователей.

Мы предлагаем вашему вниманию интересное наблюдение, обобщающее реальные проявления организации природных, естественных систем, произведённое группой томских специалистов в разных областях: Ф.Н.Алексеев (геология), Ю.Г.Дмитриев (прикладная математика, -- лидер группы), Ю.К.Устинов (теоретическая математика), П.Ф.Тарасенко (информатика), Ф.П.Тарасенко (кибернетика).

2. Математическая модель накопления вещества в растущей системе

Многие системы образуются путём наращивания масс своих частей, путём накопления, присоединения к ним соответствующей субстанции извне. (Типичными примерами могут служить образование месторождений полезных ископаемых или изменение численности населения городов.)

Рассмотрим математическую модель растущей системы, с успехом применяемую для описания разнообразных реальных систем [4]. Пусть система состоит из N частей (ячеек накопления), в которые одновременно с течением времени накапливаются количества (массы) вещества. Процесс накопления в k -ой ячейке можно описать уравнением

$$x_k^{i+1} = x_k^i + q_k^{i+1} = x_k^i + z_k^{i+1} \cdot g_k(x_k^i), \quad i = \overline{0, L-1}, \quad (1)$$

где x_k^0 - начальное значение x ; q_k^{i+1} - прирост массы, z_k^i - независимые случайные величины с математическим ожиданием $M z_k^i = a_k > 0$ и дисперсией $Var z_k^i = b_k^2$, а $g_k(x_k^i)$ - некоторые функции, задающие механизм накопления и свойства приоритетности ячеек накопления: $g_k(x) > 0$ при $x > 0$ и

$$g_1(x) > g_2(x) > \dots > g_N(x). \quad (2)$$

Предположим, что в некоторый L -й момент времени мы наблюдаем состояние растущей системы, выражаемое накопленными массами $x_1^L, x_2^L, \dots, x_k^L, \dots, x_N^L$, которые упорядочены по убыванию в силу условия (2). Схема накопления (1) при увеличении L приводит к асимптотическому распределению величины x_k^L в форме распределения Кептейна [5, р.72] с параметрами $\{La_k, Lb_k^2, G_k(x)\}$, где

$$G_k(x) = \int_{x_k^0}^x \frac{du}{g_k(u)}.$$

Пример 1. Пусть $g_k(x) = \alpha_k = Const$. Тогда накопленная масса x_k^L имеет при большом L нормальное распределение с соответствующим средним и дисперсией.

Пример 2. Пусть $g_k(x) = \alpha_k \cdot x$ (приращение пропорционально накопленной массе). Тогда x_k^L будет иметь логнормальное распределение.

Рассмотрим схему накопления (1) растущей системы в векторном виде:

$$x^{i+1} = x^i + z^{i+1} \cdot g(x^i), \quad i = \overline{0, L-1}. \quad (3)$$

Здесь $x^i = (x_1^i, \dots, x_N^i)$, $g(x^i) = (g_1(x_1^i), \dots, g_N(x_N^i))$; вектор $z^{i+1} = (z_1^{i+1}, \dots, z_N^{i+1})$ имеет моменты $Mz^{i+1} = a = (a_1, \dots, a_N)$ и $Cov(z_l, z_j) = \{b_{lj}\} = B$, $l, j = \overline{1, N}$. Умножение векторов принимается покомпонентное:

$$z^{i+1} \cdot g(x^i) = (z_1^{i+1} \cdot g_1(x_1^i), \dots, z_N^{i+1} \cdot g_N(x_N^i)), \quad i = \overline{0, L-1}. \quad (4)$$

Полагая векторы $\{z^i\}$ в (3) независимыми и одинаково распределёнными, на основании многомерной теоремы Леви -- Линдеберга [5], после соответствующих преобразований получим, что

координаты преобразованного вектора

$$G^*(x) = (G_1^*(x_1), \dots, G_N^*(x_N)), \quad x = x^L = (x_1^L, \dots, x_N^L) = (x_1, \dots, x_N).$$

можно считать независимыми и асимптотически нормально распределенными $N(0,1)$ с кумулятивной функцией распределения $\Phi(\cdot)$. Здесь

$$G^*(x) = \frac{G(x) - aL}{\sqrt{L/B}}, \quad G(x) = (G_1(x_1), \dots, G_N(x_N)),$$

$$G_k(x_k) = \int_{x_k^0}^{x_k} \frac{du}{g_k(u)} \approx b_k \cdot \sqrt{L} \cdot G_k^*(x_k) + a_k \cdot L. \quad (5)$$

Подчеркнем, что функции $G_1, \dots, G_N$ преобразуют случайные массы $x_1^L, x_2^L, \dots, x_k^L, \dots, x_N^L$ в нормально распределенные величины, но с разными параметрами. Поскольку "механизма накопления" один для всей растущей системы, то хотелось бы описать эти величины одним распределением вероятностей. Общее распределение естественно искать среди нормальных с параметрами μ и σ :

$$[G_k(x_k) - \mu] / \sigma \approx N(0,1)$$

Один из способов выбора параметров исходит из того принципа, что каждая случайная величина, стремясь проявиться в эксперименте наилучшим образом, "показывает" свое распределение в квантилях $\{x_k\}$ уровней

$$p_k = (2N - 2k + 1) / 2N, \quad k = \overline{1, N}. \quad (6)$$

Поэтому, рассматривая реализации $G_k(x_k)$ как оценки квантилей уровней p_k из нормального распределения $N(\mu, \sigma)$, параметры μ и σ находят путем минимизации суммы

$$\sum_{k=1}^N \{ \Phi[(G_k(x_k) - \mu) / \sigma] - p_k \}^2. \quad (7)$$

Другой способ оценивания μ и σ основан на соотношении

$$\Phi((G_k(x_k) - \mu) / \sigma) = p_k + \varepsilon_k, \quad (8)$$

где ε_k - случайная ошибка, $p_k = \frac{N+1-k}{N+1}$, $k = \overline{1, N}$.

После преобразований из (8) получаем

$$G_k(x_k) = \mu + \sigma \Phi^{-1}(p_k) + \sigma (\Phi^{-1}(p_k))' \varepsilon_k + o_p(\varepsilon_k^2) = \mu + \sigma \Phi^{-1}(p_k) + \varepsilon_k^*.$$

Таким образом, значения ячеек накопления $G_k(x_k)$ удовлетворяют функции регрессии вида

$$G_k(x_k) = \mu + \sigma \Phi^{-1}(p_k) + \varepsilon_k^*. \quad (9)$$

Оценка параметров μ и σ в уравнении (9) может осуществляться методом наименьших квадратов (МНК).

Отметим, что формулы (5) и (9) позволяют рассматривать регрессионные модели для описания зависимости между накопленными величинами частей в растущей системе и для осуществления прогнозов.

3. Практическая проверка теории

Суждение о правильности любой модели издавна выносится на основе близости между теоретическим предсказанием и реально произошедшим событием. Теоретическая модель, изложенная в предыдущем разделе, была приложена к реальным данным о нескольких системах различной природы.

Для проверки логнормальности размеров частей системы был применён W-тест Шапиро-Уилкса. Результаты приведены в Таблице 1.

Таблица 1.

№	Тип системы	Вероятность
1	Колонии бактерий в чашке Петри	0.810
2	Население административных областей во Франции	0.997
3	Население городов во Франции	0.999
4	Население городов в Армении	0.997
5	Население городов в Англии	0.999
6	Биомассы биогруппы деревьев в роще	0.999
7	Биомассы деревьев в биогруппе	0.860
8	Биомассы частей дерева N 20	0.817
9	Биомассы частей дерева N 42	0.935
10	Биомассы частей дерева N 47	0.965
11	Биомассы частей дерева N 61	0.955
12	Биомассы частей дерева N 67	0.719
13	Массы планет Солнечной системы	0.760

Результаты Таблицы 1 позволяют сделать следующие заключения.

— Логнормальность данных подтверждается с высокой вероятностью, что соответствует предположению о пропорциональности приращений накопленной массе.

— Закономерности накопления одинаковы для всех уровней систем (биогруппы деревьев в роще; деревья в биогруппе; крупные ветви одного дерева; мелкие веточки одной ветви; население регионов государства; население городов в регионе; и т.д.) По-видимому, природа организована единообразно на всех уровнях. Этот феномен замечен и в других областях (теория фракталов в математике, мультиноды С.Бара в менеджменте, конструкции помехоустойчивых кодов, и т.д.).

— Интригующим является вопрос о том, какие другие сетевые системы существуют, и каковы их структурные закономерности.

Обнаруженная закономерность может быть выражена в виде линейной регрессивной зависимости между логарифмом накопленных масс частей системы и квантилями нормального распределения (см. предыдущий раздел). В качестве иллюстрации приведем несколько примеров (Рис. 1-3). Проявление обсуждаемой закономерности выглядит весьма убедительно. Интересно, что выбросы на краях выборки на Рис.2 соответствуют Москве и Чукотке. Это говорит о том, что накопление в этих ячейках выпадает из общей закономерности. Интересно выяснить, каковы конкретные отличия механизмов накопления в них от общесистемного.

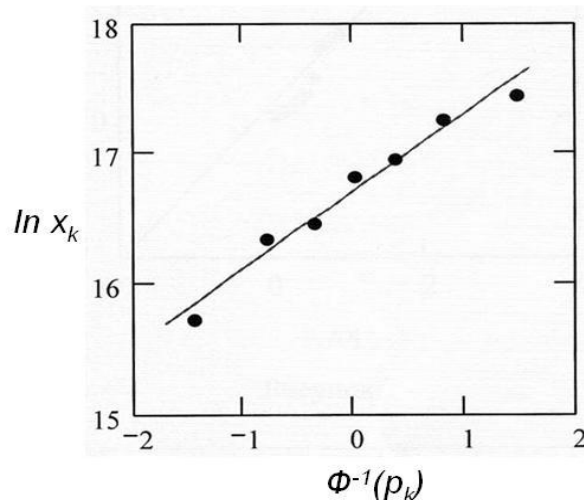


Рис.1. Регрессия логарифма численности населения восьми округов Российской Федерации по квантилям нормального распределения

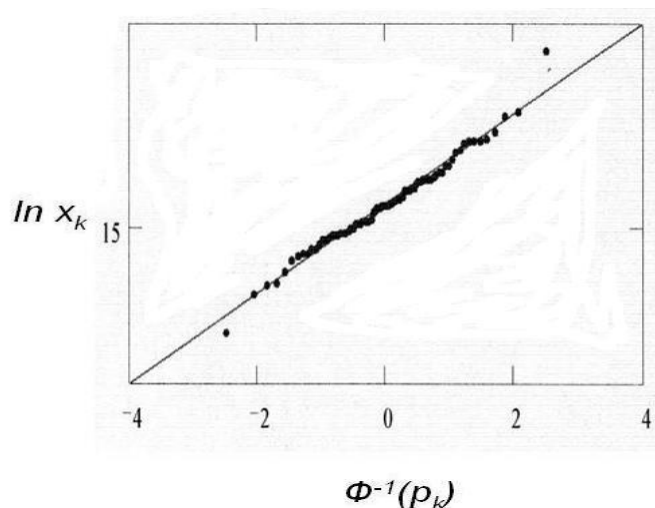


Рис.2. Регрессия логарифма налогов в 89 регионах России по квантилям нормального распределения.

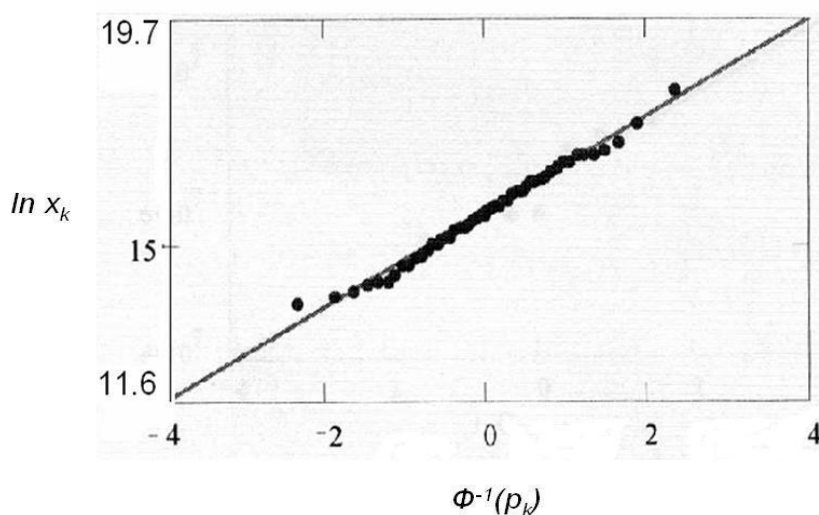


Рис.3. Регрессия логарифма налогов в 50 штатах США по квантилям нормального распределения.

4. Заключение

Представленная информация о регулярности, связывающей объемы частей системы, растущей путём накопления, является не открытием (обнаружением) закономерности, а лишь уточнением (усилением) закона природы, который уже давно обнаружен.

Здесь мы выходим на уровень обсуждения природы случайности. С одной стороны, вероятность некоторого события явным образом зависит от степени информированности субъекта, наблюдающего данное событие, о конкретных особенностях этого события. В теории вероятностей и математической статистике это отражено понятием *условной* вероятности: при изменении некоторых условий вероятность события изменяется. А учёт факта осведомленности субъекта о реальности привел к тому, что понятие условной вероятности расширяется до введения понятия *субъективной* вероятности (когда в число условий включаются не достоверные знания, а субъективные предположения, верность которых неизвестна).

В этой логике случайность выглядит как сугубо субъективное отношение человека с природой: чем больше я узнаю о том, с чем я взаимодействую, тем меньше для меня неопределенности – случайности! – в том, что будет происходить.

Однако, современные представления о мире выражаются во мнении, что случайность – не только результат ограниченности знаний субъекта, но и некое объективное свойство природы. Это ярко проявляется в квантовой физике с её волновыми функциями, принципом неопределенности Гейзенберга; в генетике с законами наследственности; в частотно-временной неопределенности в теории информации, в нелинейной динамике с её «странными аттракторами». У А.Н.Колмогорова есть замечательная статья, в

которой показывается, что однородный кусок породы раскалывается в камнедробилке на куски случайных объемов, подчиняющихся логнормальному распределению.

И приведенные в предыдущем параграфе примеры подкрепляют мнение, что понятие случайности имеет не только субъективную составляющую (незнание), но и отражает объективное свойство природы (случайность самой реальности). Это свойство можно сформулировать так:

Природа реализует возможные исходы равномерно по собственной мере.

Уточним смысл этой фразы. Происходящее явление рассматривается как реализация случайного события, т.е. вместо этого могло произойти любое из возможных событий, вероятности которых определены функцией распределения. Суть обсуждаемой закономерности состоит в том, что реализовавшиеся события «рассыпаны» на этом интервале возможностей равномерно по вероятностной мере. Ярче всего это наблюдается на примере реализаций наблюдений над некоторой непрерывной случайной величиной. Упорядоченные по величине наблюдения (порядковые статистики) разбивает всю область возможных значений на интервалы, средние вероятности которых равны, а математические ожидания самих порядковых статистик равны квантилям уровней $k/(N+1)$. Это явление имеет место и для многомерных величин (равновероятность выборочных блоков)[3]. В данном случае эта закономерность выражается в виде эквидистантности квантилей, представляющих накопленные массы частей системы.

Другое свойство накопления в рассматриваемой схеме состоит в том, что накопленные массы частей системы имеют закон распределения Кептейна с различными параметрами для каждой части (эти параметры не обязательно сдвиг и масштаб, и мы можем не получить сдигово-масштабное семейство распределений). Это обстоятельство затрудняет построение и анализ регрессионных моделей для описания зависимости между массами частей системы. Однако, третьим параметром является функция $G_k(x)$, которая описывает механизм накопления в системе. Ее знание позволяет преобразовать накопленные массы так, что все разнообразие законов распределения масс сводится к одному закону – нормальному, зависящему всего от двух параметров. А это приводит нас к линейной регрессии, т.е. к простейшему описанию зависимости масс через прямую линию. Таким образом, даже внешне различные природные процессы обладают внутренней общностью; за видимым разнообразием природы скрывается её внутреннее единство.

ЛИТЕРАТУРА

1. Ф.П.Тарасенко. Прикладной системный анализ. Изд-во Томского университета, 2004 г., 185 с.
2. Р.Л.Акофф. Менеджмент в XXI веке. Изд-во Томского университета, 2006 г., 418 с.
3. Ф.П.Тарасенко. Непараметрическая статистика. Изд-во Томского университета, 1976 г., 300 с.
4. Yu.G.Dmitriev, F.P.Tarasenko. Towards a Modeling of Natural Systems. "Cybernetics and Systems", 2009, Vol. 40, No. 3, pp. 236-248.
5. В.С. Королюк и др. Справочник по теории вероятностей и математической статистике. М. Изд-во «Наука», 1985 г.