УДК 621.372 + 519.24

ИДЕНТИФИКАЦИЯ ЭМПИРИЧЕСКИХ РАСПРЕДЕЛЕНИЙ

© 2011 В. В. Губарев

Новосибирский государственный технический университет

Рассматриваются методы идентификации распределений случайных сигналов: существующие подходы и разные методы упорядочения и автоматического выбора аналитических моделей распределений. Приводятся примеры упорядочения по некоторым методам. Описываются особенности нового многоцелевого метода моделетеки идентификации распределений, измерения различных характеристик и имитации случайных сигналов, их вариативного и вектор-моделирования.

Случайные сигналы, функции, процессы, последовательности, распределения вероятностей, идентификация, упорядочение, выбор, моделетека, моделирование, измерение характеристик, имитация.

1. Введение. Постановка задачи

Есть множество задач, где необходимо измерять (для физических величин, сигналов) или оценивать (для математических величин, числовых данных) законы распределения вероятностей мгновенных значений величин [1–18]. Если при этом в качестве исходных (первичных) значений выступают эмпирические данные (результаты измерения мгновенных значений сигналов – физических параметров (величин) реальных объектов – или значения, полученные в измерительных экспериментах), то, как правило, полученные законы распределения называют эмпирическими. Если к тому же используются непараметрические методы, например, идентификация по методу «чёрного ящика», когда параметрическая модель распределения неизвестна, то зачастую для достижения заданной цели решения поставленной задачи необходимо или желательно иметь формализованную модель этого эмпирического распределения, например, его аналитическое описание. При этом желательно, а в ряде случаев необходимо сделать процесс аналитической аппроксимации автоматическим или автоматизированным.

Именно решение задачи автоматизации выбора аналитического описания моделей и определения их параметров и рассматривается в настоящей работе.

2. Цели описания

Прежде чем находить эмпирическое распределение, а тем более его аналитическое описание, необходимо чётко представить цель этих действий, согласовать её место в дереве целей решения той конечной задачи, для которой отыскивается аналитическое описание. В качестве таких целей могут быть [19]: гносеологические, логостические, созидательные, коммуникационные, управленческие, метрологические, имитационные и т.п. Помимо целей необходимо учитывать функции, выполняемые аналитическим описанием в рамках заданной цели [19]: передаточные, измерительные, описательные, интерпретаторские, предсказательные, критериальные и т.д. Затем выдвигаются требования и критерии качества описания, определяемые целью, функциями, приложениями результатов описания. Если есть необходимость, то надо учесть допустимость при этом управления качеством результатов [20]. Например, требования и критерии сильно отличаются, если нужно как можно более точно описать наиболее вероятные значения СФ или, наоборот, очень редкие значения, когда описываются хвосты распределений, а по ним надёжностные характеристики высоконадёжных объектов.

3. Основные подходы к автоматизации

Рассматривая вопросы автоматизации аналитического описания эмпирических распределений, необходимо прежде всего либо определиться с целью и функциями описания, либо строить многоцелевые, многофункциональные или адаптивные, самоперестраиваемые под конкретные цели, задачи, требования, средства. В настоящей работе рассмотрим общие приёмы, характерные для всех подходов, которые можно положить в основу автоматизации.

Решение задачи существенно зависит от априорных сведений о распределении. Лучшее решение можно получить, если априори известен вид распределения с учётом возможных отклонений его из-за условий получения исходных данных, а также «жизни» исследуемого объекта. Тогда задача автоматизации сводится к подзадаче оценивания параметров распределения и проверки априорных гипотез о верности вида модели распределения исходным данным и условиям их получения. Если же вид модели распределения априори не известен, т.е. приходится работать в условиях «чёрного ящика», то можно либо найти эмпирическое распределение непараметрическими методами, а потом осуществлять его идентификацию, аппроксимацию, либо предварительно выбрать вид модели распределения из априори построенного множества, а затем находить его параметры и проверять гипотезы. В обеих ситуациях, особенно в последнем случае, задача осложняется тем, что множество возможных распределений континуально. Выходом из этого является априорное формирование нескольких множеств из различных моделей, удовлетворяющих определенным требованиям полноты множеств и их минимальной избыточности [20, 21].

Простейшим вариантом автоматизации является применение селекции априори набранных моделей путём их перебора, определения параметров и проверки гипотез [18]. Помимо больших затрат времени этот метод имеет ещё ряд недостатков, среди которых необходимость повторного апостериорного выбора одной модели из совокупности нескольких моделей, выдержавших ранее апостериорный тест на адекватность (целям, функциям, критериям, условиям). Для этого необходимо привлекать внешние (дополнительные) критерии выбора.

Второй вариант – это априорное формирование банка моделей – набора (базы) моделей и систем управления априорным наполнением его (её) и апостериорным выбором моделей из него (неё). Рассмотрим его подробнее.

Возможны, по крайней мере, три реализации этого варианта. Первый, прикладной, специализированный, «физический», изначально ориентирован на конкретную цель, назначение идентификации распределения. В этом случае априори набор моделей формируется так, чтобы наилучшим образом решалась конкретная прикладная задача. Например, при исследовании надёжности технических объектов в набор включаются такие модели распределений, для которых соответствующие им эмпирические характеристики надёжности (функция надёжности, интенсивность отказов и т.п.) будут адекватно (по заданному критерию) описывать исследуемые надёжностные характеристики [10], соответствовать ожидаемой полезности [16] или быть оптимальными либо экстремальными (наилучшими или наихудшими для решения задачи) по какому-либо критерию.

Вторая реализация – формальная, «математическая» (универсальная, многоцелевая, многофункциональная). В этом случае набор моделей создаётся исходя из формальных соображений, требований к нему и в силу этого допускает многоцелевое, многофункциональное применение на соответствующем этапе технического процесса движения к конечной цели решения итоговой задачи [3, 6, 20–22]. Одной из идей такой реализации является применение методов, близких по сути методам формирования библиотек и расположения книг в них (универсальная десятичная классификация, алфавитный и систематический каталоги и пр.). Ниже эта реализация будет рассмотрена подробнее.

Третья реализация – комбинация первой и второй. Например, на первом этапе быстро, формально по второму способу отбираются модели-претенденты без оценивания их параметров и проверки гипотез, а затем осуществляется селекция среди претендентов с оцениванием параметров, проверкой гипотез о распределениях или по критериям качества конечного результата.

4. Методы упорядочения и выбора моделей распределений

Рассмотрим некоторые методы априорного упорядочения и автоматического апостериорного выбора моделей эмпирических распределений.

В настоящее время наибольшее распространение получил метод пространства (в частности, плоскости) идентификаторов [3, 6, 16, 20, 21]). Его суть сводится к тому, что каждой параметрически заданной модели (например, одномерной плотности распределения $\mathrm{W}(x;a,\lambda;\boldsymbol{\alpha})$, где a – параметр положения, λ – масштаба, $\boldsymbol{\alpha}=(\alpha_1,...,\alpha_k)$ – формы (k – количество параметров формы), ставится в соответствие идентификатор $\boldsymbol{\beta}=(\beta_1,...,\beta_s)$, s – размерность пространства идентификаторов. Поскольку форма распределения инвариантна к параметрам a и λ, следует при упорядочении моделей иметь дело со стандартизованными моделями, у которых a = 0, λ = 1, либо идентификаторы $\beta_1,...,\beta_s$ строить не зависящими от a и λ. Другие требования (существования, однозначности, эффективности оценивания, устойчивости (робастности), многоцелевости, финитности, экономичности /вычислительной, измерительной, конструктивной/, инвариантности) описаны в [20, 21]. Там же описывается метод построения банков моделей μ в виде моделетеки, предполагающий полноту и минимальную избыточность моделей, входящих в неё (см. п. 5). Тогда для упорядочения моделей, не имеющих параметров формы (т.е. при k = 0), достаточно иметь всего два идентификатора β_1 и β_2, т.е. s = 2. В этом случае одномерные распределения будут располагаться на плоскости идентификаторов (β_1,β_2) в виде отдельных точек; с одним параметром формы (k = 1) – в виде линий, с двумя и более($k \geq 2$) в виде зоны ненулевой площади (рис. 1). Если же ввести три идентификатора β_1, β_2, β_3, то в таком пространстве идентификаторов $(\beta_1,\beta_2,\beta_3)$ распределения с $k \leq 1$ будут представляться точкой, с k = 2 – трёхмерной линией, с $k \geq 3$ – областью ненулевого объёма и т. д.

Априорное упорядочение в пространстве идентификаторов сводится к определению по модели μ, например, по модели $\mathrm{W}(x;a,\lambda;\boldsymbol{\alpha})$, конкретных значений (или зоны значений) идентификаторов $\beta_1,...,\beta_s$ и аналитическому, табличному или графическому упорядочению расположения значений $\beta_1,...,\beta_s$, соответствующих конкретным моделям $\mu_1,...,\mu_n$ (рис. 1). Автоматический априорный выбор модели μ_i, $i=\overline{1,n}$ по пространству идентификаторов $(\beta_1,...,\beta_s)$ сво-

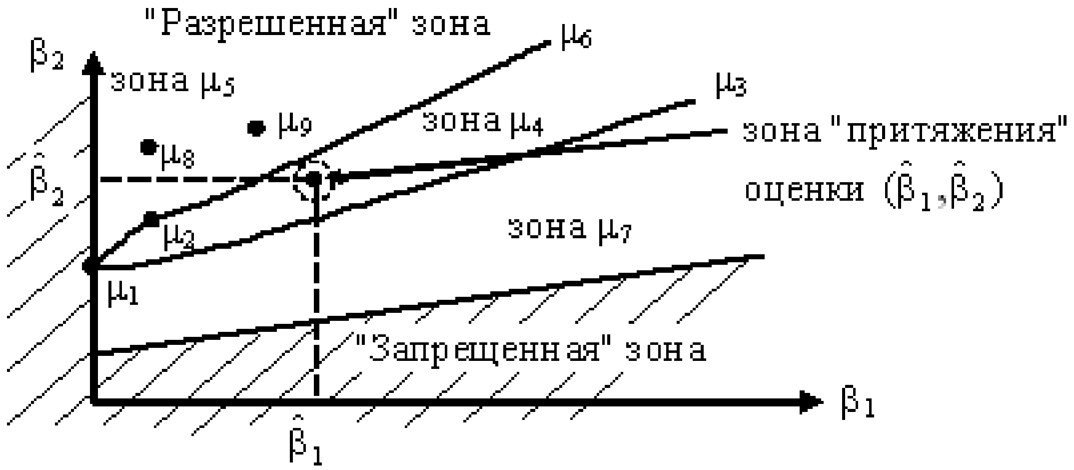

Рис. 1. Пример апостериорного выбора модели по плоскости идентификаторов (α_1, α_2): модели μ_1, μ_2, μ_8, μ_9 не имеют параметров формы; модели μ_3, μ_6 имеют один параметр формы; модели μ_4, μ_5, μ_7 имеют два или более параметров формы; по $(\hat\alpha_1,\hat\alpha_2)$ выбирается параметрическая модель μ_4

дится к поиску номера i (типа, вида, названия) модели μ_i, которой соответствуют заданные значения $\beta_1,...,\beta_s$. Апостериорный же выбор сводится, во-первых, к нахождению одним из непараметрических методов значений оценок $\hat{\beta}_1,...,\hat{\beta}_s$ идентификаторов $\beta_1,...,\beta_s$, во-вторых, к назначению области притяжения $V(\hat{\beta}_1,...,\hat{\beta}_s)$, построенной с учётом погрешностей (или доверительных зон) оценок $\hat{\boldsymbol{\beta}}$ идентификаторов $\boldsymbol{\beta}$ (рис. 1), в-третьих, к выбору той модели (или соседних моделей) μ, апостериорное значение идентификаторов $\boldsymbol{\beta}$, т.е. оценок $\hat{\boldsymbol{\beta}}$, которой (которых) попали в зону притяжения V и, наконец, к оценке параметров и проверке гипотез согласия выбранной модели с эмпирическими данными.

Примеры таких упорядочений приведены на рис. 2 (дополнительно см. [20, 21], здесь и далее нумерация распределений взята из [21]).

Как указывалось выше, одним из требований к идентификаторам $\boldsymbol{\beta}$ является их существование для упорядочиваемых моделей. Если же это условие не выполняется (так, на плоскости моментов не могут быть отражены распределения, не имеющие моментов, например, Коши), то либо надо переходить к другим идентификаторам (рис. 2 в), либо к другим методам упорядочения. Среди них определённый интерес представляют методы типа плоскости распределений (рис. 2 г), плоскости мер, расстояний или близости, канонических представлений, разложений в ряд, затянутости хвостов и другие [11, 20, 21].

В тех случаях, когда необходимо упорядочить семейство моделей (распределений) или на втором этапе описанных выше процедур необходима дальнейшая конкретизация выбранной модели в рамках выбранного семейства, следует использовать упорядочение распределений по иерархии, т.е. по конкретным значениям какого-то или каких-то параметров формы $\alpha_1,...,\alpha_k$ модели. Примеры такого упорядочения приведены в [21].

Наконец, для решения задач имитации, построения механизма образования случайностей или закона распределения, ожидаемого после функционального преобразования процессов (сигналов) с известным распределением и т.п., можно использовать упо-

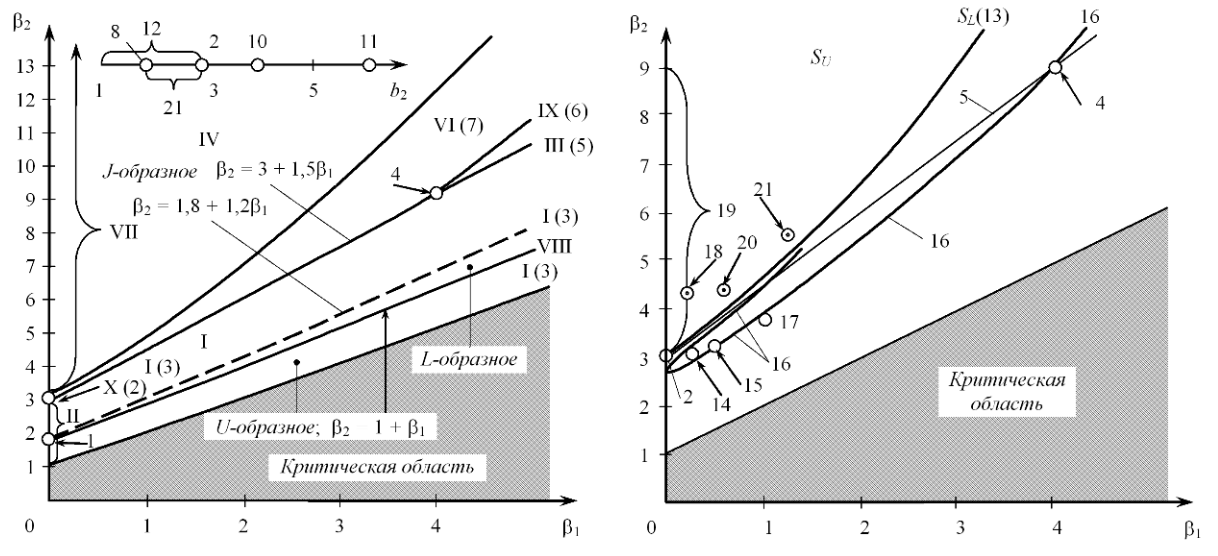

Рис. 2 а. Пример графического упорядочения по методу плоскости моментов:
I-X – Пирсона, 1 – равномерное, 2 – нормальное, 3 – бета-I, 4 – экспоненциальное, 5 – гамма, 6 – Парето, 7 – бета-II (Фишера), 8 – арксинуса, 9 – Симпсона, 10 – логистическое, 11 – Лапласа, 12 – Кэптейна-1; 13 – логнормальное, S_L – Джонсона, 14 – Максвелла, 15 – Рэлея, 16 – Вейбулла, 17 – полунормальное, 18 – Колмогорова, 19 – Стьюдента, 20 – Реньи, 21 – двойное показательное (экстремальных значений), 22 – τ-Крамера [21]; β_1 – квадрат коэффициента асимметрии; β_2 – неприведённый коэффициент эксцесса

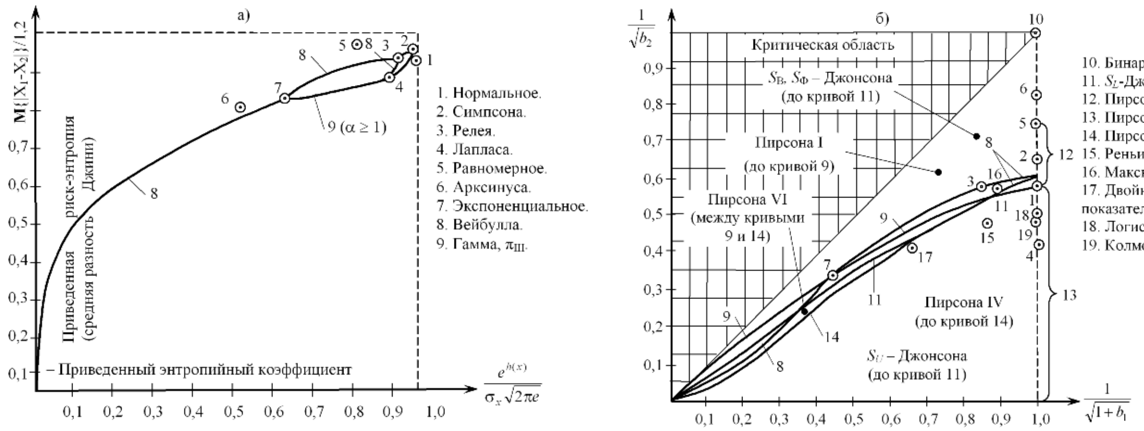

Рис. 2 б. Пример графического упорядочения распределений по методу плоскости информационных характеристик (а) и преобразованной плоскости моментов (б): $h(x) = -\mathrm{M}\{\ln W_X(x)\}$ *– дифференциальная энтропия,* $W_X(x)$ *– плотность распределения вероятностей,* $\mathbf{M}\{\cdot\}$ *– оператор математического ожидания*

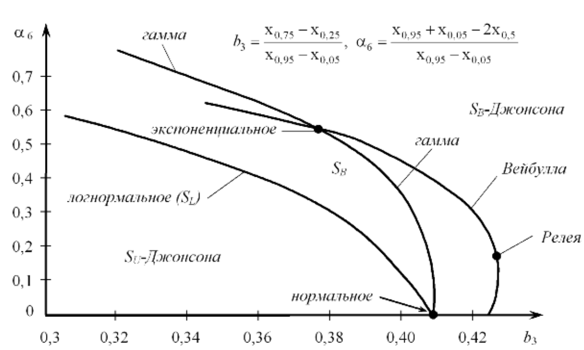

Рис. 2 в. Пример графического упорядочения распределений по методу плоскости квантилей; x_p *– квантиль порядка p*

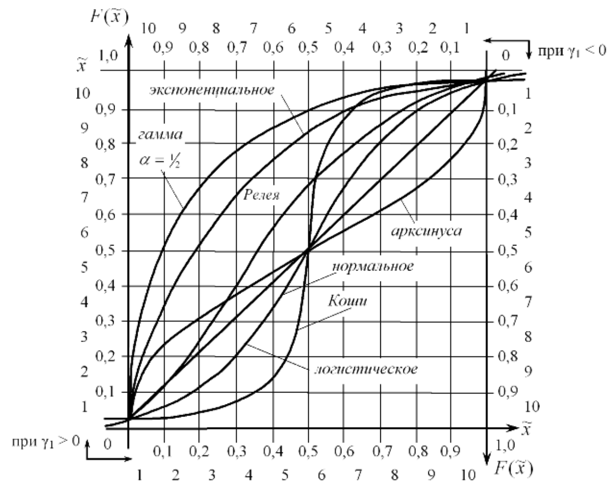

Рис. 2 г. Пример графического упорядочения распределений по методу плоскости распределений; γ_1 *– коэффициент асимметрии*

рядочение распределений по видам функциональных преобразований. Пример таких преобразований приведён в таблице 1 (см. другие примеры в [21]).

5. Моделетека

Как уже упоминалось, одна из проблем упорядочения и апостериорного выбора распределений, в качестве которых, как правило, рассматриваются дискретные и абсолютно непрерывные, заключается в том, что возможное множество распределений является континуальным, т.е. не представимым взаимно-однозначно с помощью конечного (и даже счётного) набора идентификаторов $\beta_1,...,\beta_s$. Это приводит к тому, что в одном и том же s-мерном пространстве идентификаторов $\boldsymbol{\beta}$ могут иметь место множественные пересечения моделей, имеющих тот же набор значений $\beta_1,...,\beta_s$. Во избежание этого можно использовать идеи моделетеки [20, 21]. Моделетека (от model – модель /фр. – modele, ит. modello, лат. modulus – способ, образец/ и древн. греч. τηκη – theka – хранилище) есть упорядоченное множество моделей, удовлет-

Таблица 1. Пример упорядочения моделей через функциональные преобразования

Распределение исходных независимых базовых величин $X_1, X_2, \ldots, X_n$	Функциональное преобразование (формула алгоритма имитации последовательности выборочных значений)	Распределение полученной величины Y	
Стандартное /заданное на интервале (0,1)/ равномерное распределение $R(0,1)$	1. $Y = a + \lambda X$	Равномерное $R(a,\lambda)$	36
	2.1. $Y = a - \lambda \ln X$	Экспоненциальное Ý(a,λ)	18
	2.2. $\begin{cases} Y_1 = a - \lambda X_3 \ln(X_1 X_2), \\ Y_2 = a + \lambda(X_3 - 1)\ln(X_1 X_2) \end{cases}$		
	3. $Y = a - \lambda \ln(X_1 X_2 \ldots X_n)$	Гамма (Эрланга) $G(a,\lambda;\mathrm{n})$	16
	4. $Y = a + \lambda\sqrt{-\ln X}$	Рэлея $\mathrm{Re}(a,\lambda)$	23
	5. $Y = a + \lambda(-\ln X)^{1/\alpha}$	Вейбулла ÂÃ$(a,\lambda;\alpha)$	25
	6. $Y = a + \lambda \mathrm{ctg}\,(\pi x)$	Коши $K(a,\lambda)$	40
	7.1. $Y = a + \lambda \times \begin{cases} \ln(2x), & 0 \le x < \frac{1}{2}, \\ -\ln[2(1-x)], & \frac{1}{2} \le x \le 1 \end{cases}$	Лапласа $L(a,\lambda)$	4
	7.2. $Y = a + \lambda \ln X_1 / X_2$		
	8. $Y = a + \lambda sign\left(X_1 - \frac{1}{2}\right) \ln X_2$	Лапласа $L(a,\lambda)$	4
	9.1. $Y = a + \lambda S(x)$, $S(x) = \begin{cases} \sqrt{2x}, & 0 \le x < \frac{1}{2}, \\ 2 - \sqrt{2(1-x)}, & \frac{1}{2} < x \le 1 \end{cases}$	Симпсона $C(a,\lambda)$	69
	9.2. $Y = a + \frac{\lambda}{2}(X_1 + X_2)$		
	10.1. $Y = a + \lambda \cos(\pi X)$	Арксинуса $A(a,\lambda)$	35
	10.2. $Y = a + \lambda \sin(\pi X)$		
	11. $Y = a + \lambda \ln[(1 - X)/X]$	Логистическое $\Lambda(a,\lambda)$	54
	12. $Y = a + \lambda\left[(1 - X)^{\frac{1}{1-\alpha}} - 1\right]$	Пирсона IX $\pi_{IX}(a,\lambda;\alpha)$	49
	13. $Y = a + \lambda\left(X^{\frac{1}{1-\alpha}} - 1\right)$	Пирсона VIII $\pi_{XIII}(a,\lambda;\alpha)$	30
	14.1. $Y_1 = a + \lambda\sqrt{-2\ln X_1}\cos(2\pi X_2)$	Нормальное $N(a,\lambda)$; Y_1, Y_2 – независимые	5
	14.2. $Y_2 = a + \lambda\sqrt{-2\ln X_1}\sin(2\pi X_2)$		

Окончание табл. 1

Распределение исходных независимых базовых величин $X_1, X_2, \ldots, X_n$	Функциональное преобразование (формула алгоритма имитации последовательности выборочных значений)	Распределение полученной величины Y	
Стандартное /заданное на интервале (0,1)/ равномерное распределение $R(0,1)$	15. $Y = a + \lambda\left[(1-X)^{\frac{1}{\beta}} - 1\right]^{\frac{1}{\alpha}}$	Берра-II $Á_{II}(a,\lambda;\alpha,\beta)$	60
Стандартное /при a = 0, λ = 1/ нормальное $N(0,1)$	16. $Y = a + \lambda X_1 / X_2$	Коши $K(a,\lambda)$	40
	17. $Y = a + \lambda\sqrt{X_1^2 + X_2^2 + X_3^2}$	Максвелла $M(a,\lambda)$	22
	18. $Y = a + \lambda\left(X_1^2 + X_2^2 + \ldots + X_n^2\right)/2$	Гамма /хи-квадрат/ $G(a,\lambda;n/2)$	17
	19. $Y = a + 0{,}5\left(X_1^2 + X_2^2\right)$	Экспоненциальное $Ý(a,\lambda)$	18
	27.1. $Y = a + \lambda \exp\left\{-\frac{1}{2}\left(X_1^2 + X_2^2\right)\right\}$	Равномерное на (a,λ) : $R(a,\lambda)$	36
	27.2. $Y = a + \lambda \mathrm{arctg}\left(X_1 / X_2\right)$		
	28. $Y = a + \lambda \exp\{(X-\alpha)/\beta\} =$ $a + \lambda' \exp\{X/\beta\}, \quad \lambda' = \lambda e^{-\alpha/\beta}$	S_L-Джонсона (логнормальное) $S_L(a,\lambda;\alpha;\beta)$	9
	29. $Y = a + \frac{\lambda}{2}\left[1 + th\left(\frac{X-\alpha}{\beta}\right)\right]$	S_B-Джонсона $S_B(a,\lambda;\alpha,\beta)$	10
	30. $Y = a + \lambda sh[(X-\alpha)/\beta]$	S_U-Джонсона $S_U(a,\lambda;\alpha,\beta)$	11
	31. $Y = a + \lambda\Phi[(X-\alpha)/\beta]$, Ф – стандартная нормальная функция распределения	S_Φ-Джонсона $S_\Phi(a,\lambda;\alpha,\beta)$. при α = 0, β = 1 – равномерное $R(a,\lambda)$	12
	32. $Y = a + \frac{\lambda}{2}\sum_{i=1}^{n}(a_i + X_i)^2$, $\alpha = \frac{a_1^2 + a_2^2 + \ldots + a_n^2}{2}$	Нецентральное Гамма $G(a,\lambda;\frac{n}{2},\alpha)$	79
Гамма $G(0{,}1;\alpha)$	42. $Y = a + \lambda X^{1/\beta}$	$\Im$-распределение $\Im(a,\lambda;\alpha;\beta)$	13

воряющих требованиям полноты, минимальной избыточности, уровня описания и исследованности в приложении к конкретной предметной области. Характерные особенности моделетеки: упорядоченность моделей, их полнота и отсутствие избыточности; примерно одинаковая сложность моделей; сходство в детализации описания моделей; изученность свойств всех моделей; наличие для каждой модели портфолио (все характеристики модели, аттестованные алгоритмы оценивания в разных условиях параметров моделей и параметрического оценивания всех их характеристик, погрешностей оценивания; аттестованные алгоритмы имитации выборочных значений (в разных условиях); иерархическая связь в рамках семейств (переход в частные случаи при фиксированных значениях параметров модели) и формул связи между описываемыми ими случайными элементами (величинами, векторами, функциями); примеры успешного применения моделей для решения теоретических и практических задач; упрощение автоматизации априорного и апостериорного выбора моделей; упрощение исследования погрешности классификации (от неправильного выбора модели); упрощение решения многофункциональных задач, в частности, параметрическим измерением нескольких характеристик; упрощение робастных измерений; компактность моделей.

Чтобы избежать неоднозначности или свести её к ожидаемой, следует априори оформить несколько моделеток. Например, типовых надёжностных, наиболее широко применяемых в приложениях; всех распределений Пирсона либо Джонсона или Бородачева и им подобных [21], а апостериорный выбор сделать многоэтапным: на первом этапе выбирать моделетеку, на втором – модель-претендента из выбранной моделетеки, на третьем – отбор одной из моделей-претендентов по дополнительному критерию селекции. Обратим внимание на упомянутые выше особенности моделетеки, которые дают основанным на ее использовании методам определённые преимущества при решении многофункциональных задач идентификации, измерения характеристик, имитации, прогнозирования, распознавания и др. [20].

Особый интерес идея моделетеки представляет в вариативном (поливариантном) моделировании [20] – методе исследования объектов, основанном на замене исследуемого объекта-оригинала набором разнообразных моделей его, на одновременной (совместной) работе с ними и переносе полученных результатов на объект-оригинал. При этом желательно набор моделей формировать так, чтобы они составляли вектор-модель. Вектор-модель – это система из минимального набора родственных по назначению как можно более простых и близких по сложности моделей, отражающих в совокупности всё интересующее исследователя многообразие существа (сути), закономерностей, свойств и особенностей состояния, строения и функционирования (поведения, жизни) объекта-оригинала на требуемом (согласно её назначению) уровне и обеспечивающих появление системного свойства эмерджентности (эмергентности).

В [20] перечислены основные причины, побуждающие применять вариативное моделирование и основные методы вариативного моделирования. Это экспертное формирование множества моделей; последовательный перебор, в том числе с использованием самоперестройки, адаптации; упорядоченный выбор (по типу моделетеки); отбор заведомо наилучших, хорошо проявивших себя в собственной практике, и т. д.

При этом рекомендуется использовать различные формализации типа целевого поиска, морфологического анализа, таблиц сопряжённости и т. д. Один из вариантов представлен в таблице 2.

6. Заключение

В работе рассмотрены формализованные процедуры, облегчающие автоматизацию априорного упорядочения и апостериорного выбора модели по имеющимся исходным данным. По аналогии с рассмотренными методами идентификации одномерных распределений можно ввести методы упорядочения и выбора многомерных распределений, спектральных плотностей мощности (а через них корреляционных функций), функций регрессии и т. п. [20, 21]. При этом не следует забывать о следующем: допусти-

Таблица 2. Метод матриц (таблиц) сопряженности

Задачи моделирования или требования к моделям	Свойства моделей					
	μ_1	μ_2	μ_3	μ_4	μ_5	μ_6
Задача 1 (требование 1)	–	+	–	+	+	+
Задача 2 (требование 2)	+	–	+	–	–	+
Задача 3 (требование 3)	+	–	+	–	+	–
Задача 4 (требование 4)	–	+	+	+	–	+
Задача 5 (требование 5)	+	–	–	+	+	+

Примечание. Пример выводов из матрицы сопряжённости: поставленные задачи 1, 2, 4, 5 могут быть решены (требования 1, 2, 4, 5 удовлетворены) с помощью множества μ, состоящего из одной модели μ_6; задачи 1-5 с помощью $\mu_{(1)} = (\mu_1, \mu_2)$, либо $\mu_{(2)} = (\mu_3, \mu_4)$, либо $\mu_{(3)} = (\mu_3, \mu_5)$, … .

мость и зачастую желательность описания объекта множеством моделей, связь степени полезности выбранной модели с решаемыми на её основе задачами, финитность любой модели и необходимость разной степени общности или детальности её для разных целей исследования объекта и решаемых задач; необходимость учёта дерева целей, условий и ресурсов для их достижения; зависимость вида модели от природы, реальности и качества исходных данных и условий их получения; невозможность идеального соответствия модели объекту; интерпретируемость параметров модели в терминах решаемой на её основе прикладной или исследовательской задачи; неполнота наших априорных знаний об объекте. Следует при этом всегда помнить, что излишняя формализация и автоматизация процедур сбора, обработки и анализа данных, выбора модели приводит к потере значимости таких факторов, как опыт, интуиция исследователя, способность его выявлять и отсеивать абсурдные («нефизические») данные, выделять структуры, выдвигать и проверять гипотезы и т.д.

Библиграфический список

1. Баскин, Э. М. Новый подход для анализа регрессионных моделей / Э. М. Баскин // Надежность. - 2008. - № 1 (24). – С. 10–29.

2. Губарев, В. В. О законах распределения иммунологических показателей / В. В. Губарев, В. П. Лозовой, Е. Н. Наумова, Т. В. Елисеева // Иммунология. - 1989. – № 2. – С. 50–53.

3. Карпов, И. Г. Аппроксимация вероятностных распределений интенсивности оптического излучения, распределяющегося в турбулентной атмосфере / И, Г. Карпов, А. С. Мариненко, Ю. В. Якушев // Радиотехника. - 2003. – № 10. – С. 45–48.

4. Кулямин, Г. П. Пространственно-временные характеристики обратного рассеяния от земной поверхности / Г. П. Кулямин, Е. А. Горошко, Е. В. Тарнавский // Успехи современной радиоэлектроники. - 2004. – № 4. – С. 60–70.

5. Курбацкий, В. Г. Методическое обеспечение средств измерений электромагнитных полей эллиптического характера на объектах электроэнергетики / В. Г. Курбацкий, А. В. Струмеляк // Электрика. - 2009. – № 12. – С. 36–40.

6. Лабутин, С. А. Анализ сигналов и зависимостей / С. А. Лабутин, М. В. Пугин. – Новгород: НГТУ. - 2001. – 158 с.

7. Лесных, Н. Б. Законы распределения случайных величин в геодезии: монография / Н. Б. Лесных – Новосибирск: СГГА. - 2005. – 129 с.

8. Михеева, И. В. Вероятностно-статистические модели свойств почв / И. В. Михеева. – Новосибирск: Изд-во СО РАН, 2001. – 200 с.

9. Молев, Ю. И. Статистический метод определения влияния параметров колейнос-

ти зимних дорог на уровень безопасности дорожного движения / Ю. И. Молев // Известия вузов. Машиностроение. - 2005. – № 10. – С. 46–56.

10. Надев, А. И. Математическая модель эксплуатационной надёжности интеллектуальных датчиков / А. И. Надев, Р. А. Юсупов, Ю. К. Свечников, Д. Р. Юсупов // Измерительная техника. - 2004. – № 1. – С. 8–12.

11. Новицкий, П. В. Оценка погрешностей результатов измерений / П. В. Новицкий, И. А. Зограф. – Л.: Энергоатомиздат, Л.О, 1985. – 248 с.

12. Сергеев, Г. А. Статистические методы исследования природных объектов / Г. А. Сергеев, Д. А. Януш. – Л.: Гидрометеоиздат, 1973. – 300 с.

13. Сидоров, В. А. Статистическое описание электрической прочности последовательно соединенных одинаковых изоляционных элементов / В. А. Сидоров, Д. Ф. Алферов, Е. Д. Алферова // Электричество. - 2005. – № 2. – С. 10–17.

14. Смоляк, С. А. Сопоставление вероятностных распределений по критерию ожидаемой сравнительной полезности / С. А. Смоляк // Экономика и математические методы. - 2005. – Т. 41. – № 4. – С. 91-101.

15. Труш, В. Д. ЭВМ в нейрофизиологических исследованиях / В. Д. Труш, А. В. Кориневский. – М.: Наука, 1978. – 237 с.

16. Уланова, Е. С. Методы корреляционного и регрессионного анализа в агрометеорологии / Е. С. Уланова, В. Н. Забелин. – Л.: Гидрометеоиздат, 1990.–208 с.

17. Хан, Г. Статистические модели в инженерных задачах / Г. Хан, С. Шапиро. – М.: Мир, 1969. – 395 с.

18. Яшин, А. В. Выбор решения задачи идентификации законов распределения случайных погрешностей средств измерений / А. В. Яшин, М. А. Лотонов // Измерительная техника. - 2003. – № 3. – С. 3–5.

19. Губарев, В. В. Информатика в рисунках и таблицах / В. В. Губарев. – Новосибирск: НГТУ, 2003. – 198 с.

20. Губарев, В. В. Алгоритмы спектрального анализа случайных сигналов: Монография / В. В. Губарев. – Новосибирск, 2005. – 660 с.

21. Губарев, В. В. Вероятностные модели. Справочник в 2-х частях / В. В. Губарев. – Новосибирск: НЭТИ. - 1992. – Ч. 1. – С. 1–197, Ч. 2. – С. 198–422.

22. Кликушин, Ю. Н. Методы и средства идентификационных измерений сигналов: Монография / Ю. Н. Кликушин, К. Т. Кошеков. – Петропавловск: Изд-во СКГУ им. М. Козыбаева, 2007. – 186 с.

IDENTIFICATION OF EMPIRICAL DISTRIBUTIONS

© 2011 V. V. Gubarev

Novosibirsk State Technical University

Methods of identification of random signal distributions are discussed: the existing approaches and different methods of ordering and automatic selection of analytical distribution models. Examples of ordering using some of the methods are given. The peculiarities of modeletheka as a new multipurpose method of distribution identification, various characteristics measurement and imitation of random signals, their variative and vector-modelling are described.

Random signals, functions, processes, sequences, probability distributions, identification, ordering, selection, modeletheka, modelling, characteristics measurement, imitation.

Информация об авторе

Губарев Василий Васильевич, доктор технических наук, профессор, заведующий кафедрой вычислительной техники, Новосибирский государственный технический университет. E-mail: gubarev@vt.cs.nstu.ru. Область научных интересов: вероятностное моделирование, анализ, идентификация, имитация, прогнозирование случайных сигналов, данных; информатика.

Gubarev Vasily Vasilyevitch, Honoured Science Worker of the Russian Federation, Honoured Worker of Higher School of the Russian Federation, doctor of technical sciences, professor, head of the department of computer engineering, Novosibirsk State Technical University. E-mail: gubarev@vt.cs.nstu.ru. Area of research: probability modelling, analysis, identification, imitation, forecasting of random signals and data, informatics.