МАТЕМАТИЧНІ МОДЕЛІ ТА МЕТОДИ

УДК 519.21

О МЕТОДЕ ИДЕНТИФИКАЦИИ ЗАКОНА РАСПРЕДЕЛЕНИЯ СЛУЧАЙНОЙ ВЕЛИЧИНЫ

Л.О. Кириченко, А.В. Шкловец
(Харьковский национальный университет радиоэлектроники)

В работе предложена новая характеристика ограниченной случайной величины – индикатор распределения, определяемая как отношение квадрата длины интервала, на котором определена случайная величина, к её дисперсии. Показано, что данная характеристика зависит только от закона распределения. Моделирование случайных сигналов на компьютере и оценка индикатора распределения показало возможность определения закона распределения сигналов оцениванием данной характеристики.

индикатор распределения, закон распределения, случайный сигнал

Постановка проблемы. При разработке интеллектуальных средств измерений, контроля и диагностики возникают задачи автоматического распознавания объекта исследований, представленного временным рядом наблюдений. В случае, когда исходные данные носят случайный характер, распознавание заключается в идентификации вида распределения вероятностей этих данных, так как знание распределения позволяет выбрать оптимальные алгоритмы дальнейшей обработки информации.

Анализ литературы. Классическим методом идентификации закона распределения случайной величины (СВ) является построение гистограммы. При использовании этого метода необходимо участие человека для определения закона распределения. Другим методом является анализ таких характеристик как математическое ожидание, дисперсия, мода и медиана и сравнение с нормальным распределением относительного расположения значений этих характеристик на числовой оси. Эти методы подробно описаны в работах [1 – 3]. В работе [4] рассмотрены эмпирические методы фрактального анализа случайных сигналов. Основываясь на некоторых идеях работы [4] авторами настоящей статьи предложена новая числовая характеристика непрерывной СВ ξ, определенной на ограниченном интервале $[a,b]$ – индикатор распределения $\mathrm{Id}[\xi]$.

Цель работы – показать, что данная характеристика зависит лишь от закона распределения СВ, является идентификатором типа распределения и ее оценка позволяет определять закон распределения по выборочным данным.

Определение индикатора типа распределения. Под типом распределения будем считать множество СВ ξ с дисперсией σ^2 и математическим ожиданием m, получаемых из СВ ξ_0 с нулевым математическим ожиданием и единичной дисперсией линейным преобразованием $\xi = \sigma\xi_0 + m$. Если ξ_0 имеет плотность распределения $p_{\xi_0}(x)$, то плотность преобразованной СВ ξ, определенной на интервале $[a_\xi, b_\xi]$ будет равна $p_\xi(x) = \frac{1}{\sigma} p_{\xi_0}\left(\frac{x-m}{\sigma}\right)$. Индикатор типа распределения $\mathrm{Id}[\xi]$ рассчитывается по формуле

$$\mathrm{Id}[\xi] = \frac{(b-a)^2}{D[\xi]},$$

где (b – a) – длина интервала; $D[\xi]$ – дисперсия СВ. Можно доказать следующее утверждение.

Утверждение. Отношение квадрата длины интервала, на котором определена случайная величина, к её дисперсии остается постоянным для всех СВ одного типа.

Доказательство.

$$\mathrm{Id}[\xi] = \mathrm{Id}[\sigma\xi_0 + m] = \frac{\left(\sigma b_{\xi_0} - m - (\sigma a_{\xi_0} - m)\right)^2}{\int\limits_{\sigma a_{\xi_0} - m}^{\sigma b_{\xi_0} - m} (x-m)^2 \frac{1}{\sigma} p_{\xi_0}\left(\frac{x-m}{\sigma}\right) dx} =$$

$$= \frac{\sigma^2 \left(b_{\xi_0} - a_{\xi_0}\right)^2}{\sigma^2 \int\limits_{\sigma a_{\xi_0} - m}^{\sigma b_{\xi_0} - m} \left(\frac{x-m}{\sigma}\right)^2 p_{\xi_0}\left(\frac{x-m}{\sigma}\right) d\frac{x-m}{\sigma}} = \frac{\left(b_{\xi_0} - a_{\xi_0}\right)^2}{\int\limits_{a_{\xi_0}}^{b_{\xi_0}} y^2 p_{\xi_0}(y) dx} = \mathrm{Id}[\xi_0].$$

Теоретически можно получить такие типы распределений, для которых индикаторы распределений совпадают. Например, СВ, заданные на одном интервале и имеющие зеркально симметричные плотности распределения, будут обладать одинаковым индикатором. В данном случае для различения СВ необходимо рассчитывать дополнительную характеристику, например, коэффициент асимметрии или моду. Однако на практике такие ситуации возникают достаточно редко.

В табл. 1 приведены значения индикатора $Id[\xi]$ для некоторых распространённых типов распределений [5].

Таблица 1

Значения индикатора распределения

Тип распределения	$p_\xi(x)$	Интервал	$D[\xi)$	$Id[\xi]$
Равномерное	$p(x)=\dfrac{1}{b-a}$	[a, b]	$\dfrac{(b-a)^2}{12}$	12
Симпсона	$p(x)=\dfrac{2}{b-a}-\dfrac{2}{(b-a)^2}\lvert a+b-2x\rvert$	[a, b]	$\dfrac{(b-a)^2}{24}$	24
Арксинуса	$p(x)=\dfrac{1}{\pi\sqrt{x(t-x)}}$	[0, t]	$\dfrac{t^2}{8}$	8

Оценивание индикатора распределения. В практических приложениях индикатор распределения можно использовать для идентификации типа (закона) распределения СВ по значениям временного ряда длиной n. Оценкой $Id[\xi]$ является

$$\hat{Id}=\frac{(X_{max}-X_{min})^2}{S^2},$$

где X_{max} и X_{min} – максимальное и минимальное значения ряда; S – среднее квадратичное отклонение.

Оценка $\hat{Id}$ является состоятельной

$$\lim_{n\to\infty}\hat{Id}=\frac{\lim\limits_{n\to\infty}(X_{max}-X_{min})^2}{\lim\limits_{n\to\infty}S^2}=\frac{(b-a)^2}{D^2}=Id[\xi]$$

и смещённой влево

$$\hat{Id}\le Id,\quad \text{так как } (X_{max}-X_{min})\le(b-a).$$

Для расчёта индикатора распределения $Id[\xi]$ было проведено численное моделирование значений СВ с законами распределения вероятности, приведенными в табл. 1. Значения СВ были представлены в виде временного ряда длиной n. На рис. 1 приведены графики зависимости $\hat{Id}$ от n. На рисунке видна сходимость оценки $\hat{Id}$ к теоретическим значениям. Из состоятельности оценки $\hat{Id}$ следует среднеквадратическая сходимость ряда

$\hat{Id}(n)$ к $Id[\xi]$. Среднее значение n_1, для которого на интервале $[n_1, n]$ выполняется условие $\frac{1}{n-n_1}\sum_{i=n_1}^{n}\left(Id-\hat{Id}_i\right)^2 < \varepsilon$, характеризует скорость сходимости. Значения n_1, рассчитанные для разных ε, приведены в табл. 2.

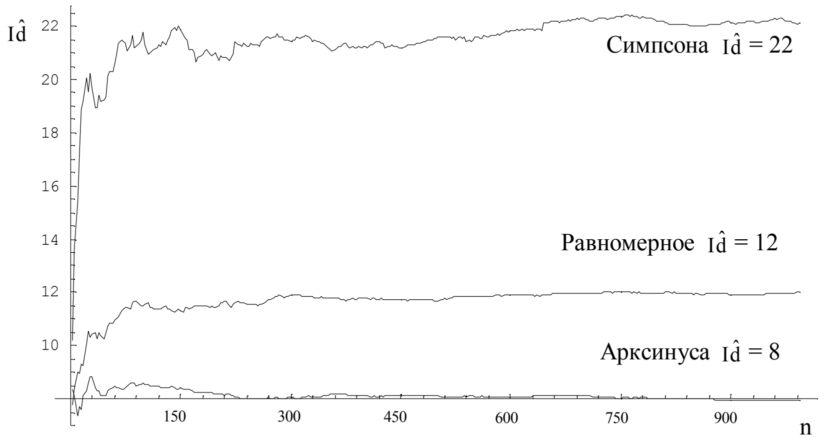


Рис. 1. Зависимость $\hat{Id}$ от n

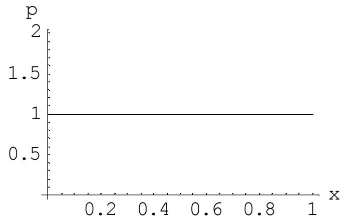
Таблица 2

Сходимость оценок

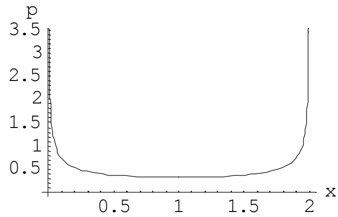
Закон распределения	ε = 0,3	ε = 0,5	ε = 1
Равномерный	53	48	30
Арксинус	19	16	12
Симпсона	351	315	204

Скорость сходимости непосредственно связана с значениями функции плотности распределения в окрестности границ интервала $[a, b]$.

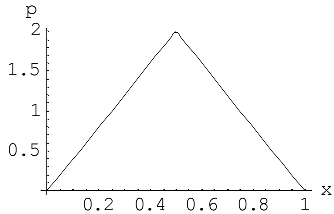
Из рис. 2 видно, что вероятность выпадения чисел вблизи концов интервала $[a, b]$ для распределения арксинуса больше, чем у равномерного распределения, поэтому размах $\left(X_{max} - X_{min}\right)$ быстрее стремится к своему истинному значению. Для распределения Симпсона вероятность выпадения крайних значений на интервале мала, по сравнению с вероятностью выпадения средних значений, поэтому сходимость $\hat{Id}(n)$ к $Id[\xi]$ происходит медленно. Значения n_1, приведенные в табл. 2, можно рассматривать как длину временного ряда, необходимую для идентификации закона распределения с требуемой точностью.



а



б



б

Рис. 2. Плотности распределений:
а – равномерного;
б – арксинуса;
в – Симпсона

Выводы. В работе предложен и теоретически обоснован метод идентификации закона распределения непрерывной ограниченной на интервале СВ. Введённый индикатор распределения Id[ξ] является числовой характеристикой закона распределения. Численные исследования показали, что предложенный метод позволяет определить закон распределения СВ по выборочным данным. Результаты работы могут быть использованы для разработки средств измерений, контроля, диагностики и многих задач распознавания объектов.

ЛИТЕРАТУРА

1. *Афифи А., Эйзен С. Статистический анализ. Подход с использованием ЭВМ. – М.: Мир, 1982. – 486 с.*
2. *Тюрин Ю. Н., Макаров А. А. Статистический анализ данных на компьютере. – М: Инфра, 1997. – 528 с.*
3. *Новицкий П.В., Зограф И.А. Оценка погрешностей результатов измерений. – Л.: Энергоатомиздат, 1991. – 304 с.*
4. *Кликушин Ю.Н. Фрактальная шкала для измерения распределений вероятности // Интернет-публикация. – М.: Журнал Радиоэлектроники. – 2000. – № 3.*
5. *Губарев В.В. Вероятностные модели: Справочник в 2-х частях. – Новосибирск: НЭТИ, 1992. – 422 с.*

Поступила 10.05.2005

Рецензент: доктор технических наук, профессор Е.В. Бодянский,
Харьковский национальный университет радиоэлектроники.