УДК 621.391

В.Н. КРАСНИКОВ, А.Б. ЛЕЩЕНКО

Национальный аэрокосмический университет им. Н.Е. Жуковского "ХАИ", Украина

МОДЕЛИРОВАНИЕ ЗАКОНОВ РАСПРЕДЕЛЕНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН ПО ЭМПИРИЧЕСКИМ ДАННЫМ

Показан подход к представлению законов распределения случайных величин в виде рядов по системам ортонормированных многочленов, коэффициенты которых определяются через числовые характеристики эмпирических данных.

законы распределения, системы ортонормированных многочленов, эмпирические данные

Введение

Теоретической основой и практическим инструментарием анализа в технике, экономике и бизнесе являются математические модели и проводимые по ним расчеты.

Разнообразные реальные технические и экономические ситуации являются потенциальными объектами моделирования. Решение этих задач тесно связано с разработкой методов формирования случайных воздействий, выработкой систем как независимых, так и зависимых случайных величин с заданными статистическими характеристиками.

При рассмотрении моделей принятия решений в условиях неопределенности и риска можно дать рекомендации по применению моделей только для типовых случаев. При этом основная трудность заключается не в выполнении расчетов, а в построении самих моделей, адекватных рассматриваемым явлениям.

Случайные возмущения, действующие на техническую или экономическую систему, можно разделить на два вида, т.е. это возмущения, прикладываемые к объекту как явлению, и возмущения, которые возникают в системе управления объектом (явлением). При этом величины случайных возмущений могут иметь как постоянные, заранее не известные значения, так и переменные, описываемые некоторыми функциями времени. В этом случае используют статистические характеристики, позволяющие осуществить оценку исследуемой системы на основе вероятностных критериев.

Особо отметим, что совокупность возможных значений случайной величины и вероятностей того, что она примет эти возможные значения, образуют закон распределения случайной величины, являющийся исчерпывающей вероятностной характеристикой случайной величины.

Так, зная интегральный закон распределения, можно определить вероятность того, что случайная величина примет значение, находящееся в пределах любого интервала аргумента x.

Дифференциальный закон распределения можно рассматривать как предел отношения вероятность события, заключающегося в том, что случайная величина x примет значения, лежащие в интервале от x до $x+\Delta x$, к величине этого интервала $\Delta x \to 0$ (поэтому мы и говорим о плотности вероятности случайной величины).

В практике решения задач проектирования различных следящих систем, включая и экономические задачи, возникает необходимость формирования системы случайных величин с различными законами распределения.

Так, например, в задаче преобразования m случайных величин $\overline{X_1, X_m}$ в системе управления можно выделить две задачи [1 – 3].

Так, например в [2], на следящую систему управления объектом (рис. 1) с выходной координатой Y действует возмущение X, которое может быть представлено случайной величиной *с заданным законом распределения*.

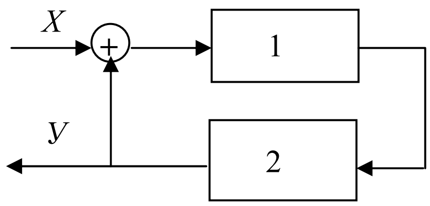

Рис. 1. Схема воздействия возмущения на следящую систему: 1 – следящая система; 2 – объект

Для проведения статистического исследования системы необходимо создать датчик, генерирующий случайную величину X с плотностью распределения $f(x)$. Однако реализация датчиков случайных величин с требуемым законом распределения $f(x)$ вызывает затруднения.

Решение возможно путем построения функционального преобразователя Z, позволяющего осуществить преобразование случайной величины Δ с заданным законом распределения $f(\delta)$ в случайную величину X с требуемым законом распределения $f(x)$ (рис. 2).

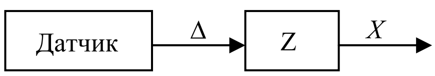

Рис. 2. Функциональная схема одноканального преобразования случайных величин

Таким образом, возможно построить систему независимых случайных величин $\overline{X_1, X_m}$ с плотностью распределения

$$f\left(x_1, x_2, ..., x_m\right) = \prod_{j=1}^{m} f_j\left(x_j\right).$$

В случае если следящая система подвергается воздействию зависимых случайных величин $\overline{X_1, X_m}$, то возникают проблемы как теоретического, так и практического порядка.

Таким образом, при исследовании следящей системы с большим числом зависимых случайных воздействий $\overline{X_1, X_m}$ возникает более общая задача, а именно, найти некоторое функциональное преобразование

$$\mathrm{Z}\left(\delta_1, \delta_2, ..., \delta_m, x_1, x_2, ..., x_m\right) = 0,$$

осуществляющее переход от системы случайных величин $\overline{\Delta_1, \Delta_m}$ с заданной плотностью распределения $f_3\left(\delta_1, \delta_2, ..., \delta_m\right)$ к системе случайных величин $\overline{X_1, X_m}$ с требуемой плотностью распределения $f\left(x_1, x_2, ..., x_m\right)$.

Решение описанных выше двух задач возможно, если будет найден вид функции Z, позволяющей преобразовать случайные числа имеющегося в наличие датчика случайных чисел в случайные числа с требуемой плотностью распределения $f(x)$. Очевидна искусственность подхода. Мы хотели показать тот факт, что при вероятностном подходе в различных задачах практических исследований знание закона распределения случайной величины играет исключительную роль.

Формулировка задачи и ее решение

Решение многих задач в значительной степени упрощается, если в качестве гипотезы относительно распределения используется нормальный закон распределения (в силу его «аналитичности», прежде всего).

В связи с растущими требованиями достижений более высоких точностей при оценивании параметров измерений, например, или обработки эмпирических данных в экономике и социологии, стоит задача отыскания алгоритмов оценивания непосредственно по выборочным данным без привлечения гипотетических законов распределения. Поэтому имеет практический интерес представление распределений случайных величин в виде рядов по системам ортонормированных многочленов, коэффициенты которых определяются через числовые характеристики эмпирических данных.

В данной работе будут использованы системы ортонормированных многочленов $\{Q_n(x)\}_{n=0}^{\infty}$ в пространстве $L_\sigma^2(T)$, где $x \in T \subseteq R^m$, $\sigma(x)$ – некоторая функция распределения, т.е. $\sigma(x)$ – функция ограниченной вариации $\underset{T}{\operatorname{var}}\,\sigma < +\infty$.

Система многочленов $\{Q_n(x)\}_{n=0}^{\infty}$ называется ортонормированной, если

$$\int_T Q_i(x) Q_j(x) d\sigma(x) = \delta_{ij} = \begin{cases} 1, i = j; \\ 0, i \neq j. \end{cases} \qquad (1)$$

Отметим, что $i = (i_1, i_2, ..., i_m)$, $j = (j_1, j_2, ..., j_m)$ – мультииндексы.

Для наглядности изложения воспользуемся случаем одномерного $T \subseteq R^1$. Примерами системы многочленов (1) являются:

– система многочленов Лежандра на интервале $[-a, a]$, ортогональных относительно веса $d\sigma(x) = dx$:

$$P_j^a(x) = \sqrt{\frac{2j+1}{2a}} \cdot \frac{1}{2^j} \sum_{i=0}^{j/2} \frac{(-1)^j (2j-2i)!}{i!(j-i)!(j-2i)!} \times$$
$$\times \frac{1}{a^{j-2i}} \cdot x^{j-2i};$$

– система многочленов Лагерра на интервале $(0, +\infty)$, ортогональных относительно веса $d\sigma(x) = e^{-x} dx$:

$$L_j(x) = (-1)^j \sum_{i=0}^{j} \frac{j!(-1)^i}{(i!)^2 (j-i)!} \cdot x^i ;$$

– система многочленов Эрмита на интервале $(-\infty, +\infty)$, ортогональных относительно веса $d\sigma(x) = e^{-x^2} dx$:

$$H_j(x) = \sqrt{\frac{j!}{\sqrt{\pi} \cdot 2^j}} \sum_{i=0}^{j/2} \frac{(-1)^j \, 2^{j-2i}}{i!(j-2i)} \cdot x^{j-2i} .$$

Все три вида приведенных выше многочленов будем записывать в виде

$$Q_j(x) = \sum_{i=0}^{j} b_{ij} x^i . \qquad (2)$$

По набору эмпирических данных случайных величин $x_1, x_2, ..., x_n$ вычисляются моменты

$$S_\nu = \frac{1}{n} \sum_{\mu=1}^{n} x_\mu^\nu, \; \nu = 0, 1, 2, ..., k , \qquad (3)$$

полагая, что

$$S_\nu = \int_T x^\nu f(x) d\sigma(x) .$$

Метод приближенного вычисления плотности распределения $f(x)$ случайной величины x_μ заключается в ее приближенном представлении

$$f(x) \approx \sum_{j=0}^{k} a_j Q_j(x) . \qquad (4)$$

В этом случае

$$a_j = \int_T f(x) Q_j(x) d(x) =$$
$$= \sum_{i=0}^{j} b_{ij} \int_T f(x) x^i dx = \sum_{i=0}^{j} b_{ij} S_i ,$$
$$j = 0, 1, 2, ..., k , \qquad (5)$$

Таким образом, функция $f(x)$ может быть восстановлена по формулам (4) и (5).

Практическую проверку приближенного вычисления плотности распределения можно осуществить путем сравнения предполагаемого закона распределения с получаемым.

1. Для заданных случайных чисел, распределенных по равномерному закону на интервале $[-1, 1]$, например, используем многочлены Лежандра:

$$P_j(x)=\sqrt{\frac{2j+1}{2}}\cdot\frac{1}{2^j}\sum_{i=0}^{j/2}\frac{(-1)^j(2j-2i)!}{i!(j-i)!(j-2i)!}\cdot x^{j-2i},$$

где $b_{ij}=\sqrt{\frac{2j+1}{2}}\cdot\frac{1}{2^j}\cdot\frac{(-1)^j(2j-2i)!}{i!(j-i)!(j-2i)!}$ (см. (2)).

Далее рассчитываем по формуле (3) моменты для заданного ν, определяем коэффициенты a_j по (5) и находим $f(x)$ по (4).

2. Для случайных чисел, распределенных по закону Пуассона, например, на интервале $[0,+\infty]$, используем многочлены Лагерра.

Функция плотности распределения определяется выражением

$$f(x)=e^{-\lambda x}\sum_{j=0}^{k}a_jL_j^{\lambda}(x),$$

а коэффициент a_j – равенством

$$a_j=\int_0^{\infty}f(x)L_j^{\lambda}(x)dx=\sum_{i=0}^{j}b_{ij}^{\lambda}S_i,$$

где S_i – определенные эмпирические моменты.

3. Для случайных чисел, распределенных по закону Гаусса, используем многочлены Эрмита.

Функция плотности распределения определяется выражением

$$f(x)=e^{-x^2}\sum_{j=0}^{k}a_jH_j(x),$$

где в наших обозначениях $H_j(x)=\sum_{i=0}^{j}b_{ij}x^i$, а коэффициенты $a_j=\sum_{i=0}^{j}b_{ij}S^i$.

Заключение

На основе использования систем ортонормированных многочленов Лежандра, Лагерра и Эрмита представлено распределение случайных величин в виде рядов, коэффициенты которых определяются через эмпирические данные на интервалах, соответственно, $[-1,1]$, $[0,+1]$ и $[-\infty,+\infty]$.

Полученные результаты позволяют создавать алгоритмы оценивания параметров измерений в области радиоэлектроники, а также при решении задач обработки данных в социологии и экономики.

Литература

1. Лившин Н.А., Пугачёв В.Н. Вероятностный анализ систем автоматического управления. Т. 1. – М.: Сов. радио, 1963. – 887с.

2. Проектирование следящих систем / Под общ. ред.. Е.П. Попова. – М.: Машиностроение, 1977. – 357 с.

3. Математическое обеспечение сложного эксперимента: в 5 т. / Под общ. ред. И.И. Ляшко. – К.: Наук. думка, 1990. – Т. 5. Проблемы построения математического и программного обеспечения измерительно-вычислительных комплексов. – 368 с.

Поступила в редакцию 4.11.2005

Рецензент: д-р физ.-мат наук, проф. А.Г. Николаев, Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ», Харьков.