УДК 004.94

А. В. Шляева, И. В. Рудаков

ИСПОЛЬЗОВАНИЕ ОГРАНИЧЕННЫХ НЕПРЕРЫВНЫХ РАСПРЕДЕЛЕНИЙ ДЛЯ ПОСТРОЕНИЯ МОДЕЛЕЙ СЛУЧАЙНЫХ ВХОДНЫХ ДАННЫХ ПРИ ИМИТАЦИОННОМ МОДЕЛИРОВАНИИ СИСТЕМ

Рассмотрен вопрос построения моделей случайного входного воздействия для имитационных моделей систем. Обсуждается возможность использования ограниченных непрерывных распределений для моделирования входных воздействий в отсутствие экспериментальных данных, используя экспертные оценки параметров входного воздействия (минимальное, максимальное значения, мода). Приведено описание функциональности программного обеспечения для моделирования входных воздействий с использованием двухстороннего степенного и бета-распределений.

E-mail: shlyaeva@gmail.com; irudakov@yandex.ru

Ключевые слова: *имитационное моделирование, моделирование случайных входных воздействий, генерация случайных чисел, ограниченные непрерывные распределения.*

В настоящее время остается актуальной проблема исследования сложных дискретных систем (СДС), таких как компьютерные сети, вычислительные системы, дорожные потоки, поточные линии предприятий. Ввиду ограниченности возможностей экспериментального исследования больших систем была разработана методика их моделирования, которая позволила описать протекание процессов функционирования систем с помощью математических моделей и получить оценки характеристик исследуемых объектов по результатам экспериментов с моделями. Основным методом исследования характеристик сложных систем в настоящее время является компьютерное моделирование.

Имитационное моделирование служит одним из наиболее эффективных средств исследования характеристик СДС; с его помощью достаточно просто учитывать такие факторы, как нелинейные характеристики элементов системы, наличие многочисленных случайных воздействий и т. д.

Один из этапов создания имитационной модели сложной системы заключается в построении модели и генерации входных данных, зависящих от случайных факторов. Неверная формализация случайных входных воздействий приводит к получению недостоверных результатов моделирования и неверному принятию решений.

При построении модели случайных входных воздействий по имеющимся статистическим данным, представляющим собой реализацию случайной величины, как правило, осуществляется подбор теоретического распределения вероятностей и его параметров. При использовании в качестве модели случайного входного воздействия случайной величины, имеющей некоторое теоретическое распределение, может возникнуть следующая проблема: многие наиболее часто используемые теоретические распределения являются неограниченными (как гауссово распределение) или ограниченными слева (экспоненциальное, логнормальное, Вейбулла, Эрланга и т. д.). В таком случае существует ненулевая вероятность получения неправдоподобно больших значений при генерации случайных входных воздействий по выбранному теоретическому распределению. Один из подходов к решению этой проблемы состоит в применении ограниченных теоретических распределений в качестве моделей входных данных.

Ограниченные теоретические распределения также можно использовать в случае, если при моделировании входных данных отсутствует возможность получения выборок, описывающих эти данные. В таком случае прибегают к помощи экспертных оценок минимальных, максимальных и средних значений искомой величины, затем, исходя из экспертных оценок вида распределения, выбирают модель входного воздействия. В ходе моделирования поведения системы генерация требуемого случайного входного воздействия осуществляется в соответствии с выбранной моделью входного воздействия.

В настоящей статье описывается расширение функциональности программного обеспечения для моделирования и генерации случайных входных воздействий [1]: возможность построения модели без экспериментальных данных, по экспертным оценкам параметров искомого входного воздействия.

Ограниченные непрерывные распределения. Помимо традиционно используемых в области моделирования случайных входных воздействий треугольного и бета-распределения также рассмотрим ограниченное распределение Джонсона и двустороннее степенное распределение.

Треугольное распределение. Наиболее часто его используют как приблизительную модель в отсутствие данных. Плотность треугольного распределения имеет следующий вид:

$$f(x) = \begin{cases} \dfrac{2\,(x-a)}{(b-a)\,(c-a)}, & a < x \leqslant c; \\ \dfrac{2\,(b-x)}{(b-a)\,(b-c)}, & c < x \leqslant b, \end{cases}$$

где a, b — минимальное и максимальное значения распределения; c — значение моды. Функцию распределения запишем как

$$F(x) = \begin{cases} \dfrac{(x-a)^2}{(b-a)(c-a)}, a < x \leqslant c; \\ 1 - \dfrac{(b-x)^2}{(b-a)(c-a)}, c < x \leqslant b. \end{cases}$$

При выборе параметров следует руководствоваться экспертной оценкой минимального, максимального и наиболее вероятного значений переменной.

Бета-распределение. Еще одно известное и широко используемое в области моделирования случайных входных воздействий ограниченное распределение — бета-распределение, которое предоставляет более гибкие возможности моделирования по сравнению с треугольным [2]. Плотность бета-распределения:

$$f(x) = \frac{x^{\alpha_1 - 1}(1-x)^{\alpha_2 - 1}}{B(\alpha_1, \alpha_2)},$$

где $B(\alpha_1, \alpha_2) = \int\limits_0^1 t^{\alpha_1 - 1}(1-t)^{\alpha_2 - 1}\,dt$ — бета-функция; α_1 и α_2 — параметры формы.

Распределение является ограниченным на отрезке $[0, 1]$, но масштабированием можно получить ограниченное бета-распределение на отрезке $[a, b]$.

Значения параметров формы можно выбирать по экспертным оценкам среднего и моды [1]:

$$\hat{\alpha}_1 = \frac{(\mu - a)(2c - a - b)}{(c-\mu)(b-a)} \text{ и } \hat{\alpha}_2 = \frac{(b-\mu)\,\hat{\alpha}_1}{\mu - a}.$$

Таким образом, эксперту необходимо задать оценку моды и математического ожидания.

Для оценки параметров можно также применять метод моментов. В этом случае эксперту следует указать оценки математического ожидания и среднеквадратического отклонения:

$$\hat{\alpha}_1 = \mu\left(\frac{\mu(1-\mu)}{\sigma} - 1\right) \text{ и } \hat{\alpha}_2 = (1-\mu)\left(\frac{\mu(1-\mu)}{\sigma} - 1\right).$$

Ограниченное распределение Джонсона. В качестве модели можно также рассматривать ограниченное распределение Джонсона, причем это распределение, как и бета-распределение, может быть бимодальным. Идея семейства распределений Джонсона основана на таком преобразовании $f(x)$ исходной случайной величины X, которое позволит принять результат преобразования как нормированную случайную величину, имеющую Гауссово распределение.

Функция распределения Джонсона имеет вид

$$F(x) = \Phi\left(\gamma + \delta f\left(\frac{x-\xi}{\lambda}\right)\right),$$

где $\Phi(x)$ — функция распределения вероятностей стандартного нормального распределения; γ, δ — параметры формы; λ — масштабный параметр; ξ — параметр положения, $f(x) = \ln\left[\dfrac{x}{1-x}\right]$ в случае рассматриваемого ограниченного распределения:

$$F(x) = \Phi\left(\gamma + \delta \ln\left(\frac{x-\xi}{\lambda+\xi-x}\right)\right),$$

причем $\lambda + \xi \geqslant b$, $\xi \leqslant a$.

Параметры распределения Джонсона по выборке независимых данных можно оценить методом моментов [3]. В этом случае эксперту необходимо указать оценки среднего, дисперсии, асимметрии и эксцесса, поэтому использование распределения Джонсона для описания случайного входного воздействия без экспериментальных данных затруднительно.

Двухстороннее степенное распределение. Данное распределение (two-sided power distribution, TSP) является обобщением треугольного распределения [6]. Плотность распределения имеет вид:

$$f(x) = \begin{cases} \dfrac{n}{b-a}\left(\dfrac{x-a}{c-a}\right)^{n-1}, a < x \leqslant c; \\ \dfrac{n}{b-a}\left(\dfrac{b-x}{b-c}\right)^{n-1}, c < x \leqslant b, \end{cases}$$

где $n > 0, n \in R$;

функция распределения —

$$F(x) = \begin{cases} \dfrac{c-a}{b-a}\left(\dfrac{x-a}{c-a}\right)^{n}, \ a < x \leqslant c; \\ 1 - \dfrac{b-c}{b-a}\left(\dfrac{b-x}{b-c}\right)^{n}, \ c < x \leqslant b. \end{cases}$$

Параметры a, b и c двухстороннего степенного распределения аналогичны используемым в треугольном распределении (минимальное, максимальное значения и мода). В зависимости от параметра n плотность распределения имеет различный вид. При $n = 2$ двухстороннее степенное распределение сводится к треугольному распределению, при $n = 1$– к равномерному распределению. При $0 < n < 1$–это бимодальное распределение с модами в точках a и b и антимодой в c, при $n > 1$ — унимодальное распределение с модой в точке c.

Из рассмотренных распределений для моделирования случайных входных воздействий при отсутствии выборки наибольший интерес представляют: бета-распределение, треугольное распределение и двухстороннее степенное распределение, так как они позволяют довольно просто и интуитивно понятно использовать экспертные оценки параметров случайных входных воздействий.

Генерация случайных входных воздействий. ***Генерация случайных величин, распределенных по треугольному и двухстороннему степенному распределениям.*** Для генерации случайной величины, имеющей двухстороннее степенное распределение, применяют метод обратного преобразования:

$$X = \begin{cases} a + \sqrt[n]{U\left(b-a\right)\left(c-a\right)^{n-1}},\ 0 < U \leqslant \dfrac{c-a}{b-a}; \\ b - \sqrt[n]{\left(1-U\right)\left(b-a\right)\left(b-c\right)^{n-1}},\ \dfrac{c-a}{b-a} < U \leqslant 1, \end{cases}$$

где U — случайная величина, имеющая равномерное распределение на отрезке [0,1].

Например, для генерации случайной величины, характеризующейся треугольным распределением, значения вычисляют по формуле:

$$X = \begin{cases} a + \sqrt{U\left(b-a\right)\left(c-a\right)},\ 0 < U \leqslant \dfrac{c-a}{b-a}; \\ b - \sqrt{\left(1-U\right)\left(b-a\right)\left(b-c\right)},\ \dfrac{c-a}{b-a} < U \leqslant 1. \end{cases}$$

Генерация случайных величин, имеющих бета-распределение. Общий метод генерации случайных величин, имеющих бета-распределение, заключается в использовании генераторов случайных чисел, имеющих гамма-распределение:

$$X = \frac{Y_1}{Y_1 + Y_2},$$

здесь Y_1, Y_2 — случайные величины, подчиняющиеся гамма-распределению с параметрами $\alpha_1, 1$ и $\alpha_2, 1$ соответственно.

Построение модели случайных входных воздействий без экспериментальных данных. Рассмотрим виды плотностей двухстороннего степенного и бета-распределений для смещенного распределения. Пусть распределения находятся на отрезке [0, 1] и распределение смещено влево (мода расположена в точке $c = 0{,}2$). Зададим параметр n двухстороннего степенного распределения: тогда оно полностью определяется параметрами a, b, c и n. Для вычисления параметров бета-распределения используем указанные выше оценки параметров по моде (точка c= 0,2) и математическому ожиданию, равному математическому ожиданию двухстороннего степенного распределения:

$$\mu = \frac{a + (n-1)\,c + b}{n+1}.$$

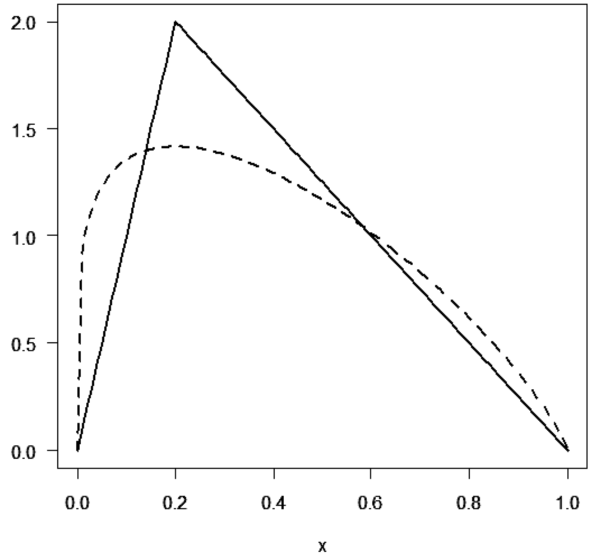

Рис. 1. Плотности вероятностей двухстороннего степенного ($n = 2$) и бета-распределения (1,2; 1,8), сплошная и пунктирная линии соответственно

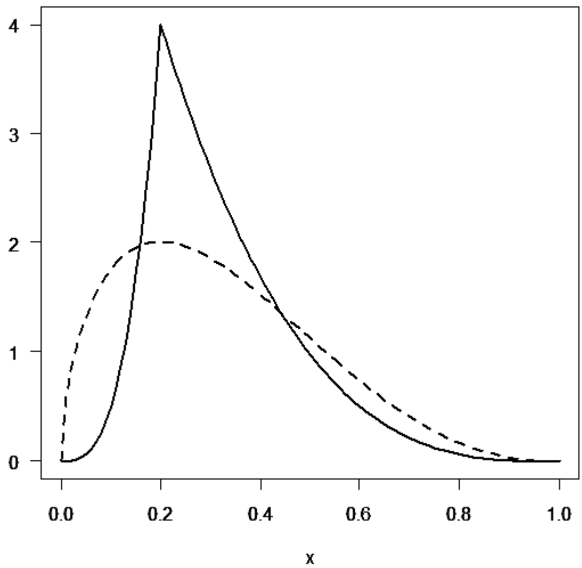

Рис. 2. Плотности вероятностей двухстороннего степенного ($n = 4$) и бета-распределения (1,6; 3,4) , сплошная и пунктирная линии соответственно

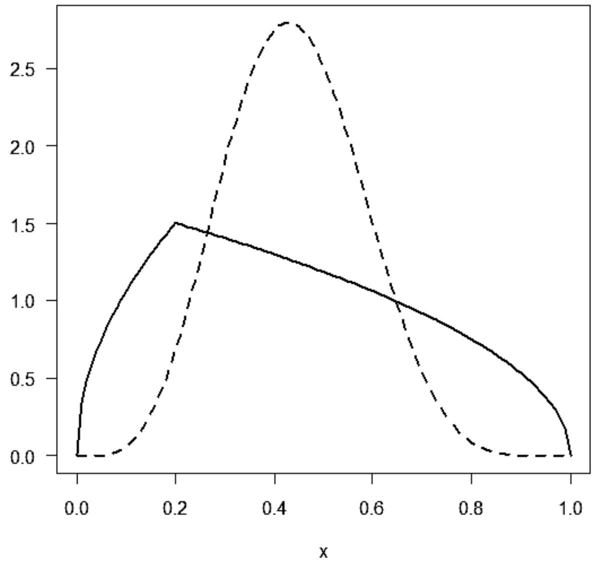

Рис. 3. Плотности вероятностей двухстороннего степенного ($n = 1{,}5$) и бета-распределения (5,5; 7), сплошная и пунктирная линии соответственно

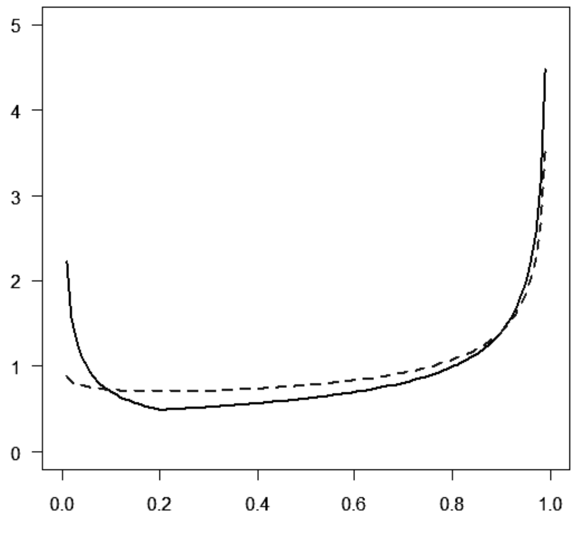

Рис. 4. Плотности вероятностей двухстороннего степенного ($n = 0{,}5$) и бета-распределения (0,9; 0,6) , сплошная и пунктирная линии соответственно

На рис. 1–4 приведены виды плотностей вероятностей двухстороннего степенного и бета-распределений для значений параметра n, равных соответственно: 2 (математическое ожидание 0,4; параметры бета-распределения 1,2; 1,8); 4 (математическое ожидание 0,32; параметры бета-распределения 1,6; 3,4); 1,5 (математическое ожидание

0,44; параметры бета-распределения 5,5; 7) и 0,5 (математическое ожидание 0,6; параметры бета-распределения 0,9; 0,6).

На графиках видно, что при небольших изменениях оценки математического ожидания существенно меняется вид плотности распределения. В то же время дисперсия двухстороннего степенного распределения:

$$\sigma^2 = (b-a)^2 \frac{n - 2(n-1)\dfrac{c-a}{b-a}\dfrac{b-c}{b-a}}{(n+2)(n+1)^2}.$$

Для рассмотренных параметров распределения дисперсии равны, соответственно, 0,04; 0,02; 0,06 и 0,11. Разброс значений дисперсии позволяет эксперту более наглядно задать вид плотности распределения, используя оценку дисперсии, вместо того, чтобы задавать оценку математического ожидания.

Программно-алгоритмический комплекс для моделирования входных воздействий рассмотрен в [1]. Для того чтобы выбрать модель без использования экспериментальных данных, в него добавлен модуль выбора модели по экспертным оценкам параметров распределения с использованием двухстороннего степенного или бета-распределения.

В случае двухстороннего степенного распределения эксперту необходимо задать интервал, на котором ограничено распределение случайного входного воздействия (минимальное и максимальное значения) и моду. Далее эксперт может изменять значения оценки дисперсии (или параметра n), визуально оценивая изменения плотности распределения. После выбора параметров интервал сохраняется, и можно сгенерировать данные по выбранному распределению с заданными параметрами.

На графиках также видно, что плотность бета-распределения может иметь моду, которая отличается от заданной в экспертной оценке. Поэтому для задания параметров бета-распределения были выбраны экспертные оценки максимального и минимального значений, а также математического ожидания и дисперсии (или параметров распределения напрямую, как и в случае двухстороннего степенного распределения). При выборе параметров эксперт также может визуально оценивать изменения плотности распределения.

Заключение. Рассмотренные в работе ограниченные непрерывные распределения вероятностей — двухстороннее степенное и бета-распределение — можно использовать для моделирования случайной величины по экспертным оценкам интервала распределения, моды, математического ожидания и дисперсии. Двухстороннее степенное распределение может служить в качестве более гибкой альтернативы треугольному распределению при выборе модели случайного входного воздействия. Моделирование случайных входных воздействий с использованием указанных распределений позволяет осуществлять приведенная в работе расширенная функциональность программного обеспечения.

СПИСОК ЛИТЕРАТУРЫ

1. Рудаков И. В., Шляева А. В. Моделирование входных данных для стохастических имитационных моделей систем // Информационные технологии. – 2006. – № 11. – С. 8–12.
2. Кельтон В., Лоу А. Имитационное моделирование: Классика CS. – 3-е изд. – СПб.: Питер; Киев: Издательская группа BHV, 2004. – 847 с.
3. Hill I. D., Hill R., Holder R. L. Fitting Johnson curves by moments // Applied Statistics. – 1976. – Vol. 25. – P. 180–189.
4. Univariate input models for stochastic simulation / M.E. Kuhl, J.S. Ivy, E.K. Lada, N.M. Steiger, M.A. Wagner and J.R. Wilson // Journal of Simulation. – 2010. – Vol. 4. – P. 81–97.
5. Kotz S., van Dorp J. R. A novel method for fitting unimodal continuous distributions on a bounded domain utilizing expert judgment estimates // IIE Transactions. – 2006. – Vol. 38, 5. – P. 421–436.
5. Van Dorp J. R., Kotz S. The standard two sided power distribution and its properties with applications in financial engineering // The American Statistician. – 2002. – 56 (2). – P. 90–99.

Статья поступила в редакцию 10.05.2012