

УДК 621.398

ВЫБОР КРИТИЧЕСКИХ РЕАЛИЗАЦИЙ И ЕГО СОЧЕТАНИЕ С ДРУГИМИ МЕТОДАМИ УСКОРЕНИЯ СТАТИСТИЧЕСКОГО ЭКСПЕРИМЕНТА

Ю. Ф. Подоплекин,

доктор техн. наук, зам. генерального директора, директор по науке

В. Р. Андриевский,

канд. техн. наук, профессор

Федеральное государственное унитарное предприятие «ЦНИИ "Гранит"»

Рассматривается возможность применения метода отбора критических реализаций случайных факторов для решения общей задачи оценки вероятностных показателей качества исследуемой системы. Показывается возможность использовать выбор критических реализаций в сочетании с другими методами сокращения трудоемкости статистического эксперимента: «выделения главной части» и «расслоенной выборки».

The possibility of application of the method of sampling of random factors critical realizations for the solution of the general problem of estimation of probability rates of the investigated system quality is examined. Availability of sampling of critical realizations combined with other methods of labor saving of a statistical experiment such as «leading part separation» and «stratified sample» is discussed.

Метод выбора критических реализаций случайных факторов с исключением области очевидных («неинтересных») исходов при статистическом моделировании был предложен в 1960-х годах сотрудником ЦНИИ «Гранит» Н. П. Яковлевым как средство существенного сокращения трудоемкости оценки вероятности неблагоприятного функционирования системы. Эффективность метода была показана на примере исследования динамики управления крылатой ракетой на участке старта. Идея метода была описана в публикациях 1973–1975 гг. [1, 2]. Метод впоследствии переоткрывался другими исследователями. (Известна ошибочная ссылка в книге [3] на работу В. В. Кузьмина и В. А. Ярошевского в «Ученых записках ЦАГИ», 1984, т. 15, № 2, как на приоритетную в этом вопросе.)

Идея метода заключается в установлении эквивалентности между двумя схемами статистического эксперимента: «обычной», когда по полной программе исследуются все получившиеся в результате «игры со случайным» реализации случайных факторов, и «критической», когда отсеиваются «лишние», не определяющие возможности интересующего исследователя исхода (как правило, неблагоприятного).

Реализация идеи может быть различной. Например, генерация случайных условий эксперимента и

подсчет числа реализаций ведутся обычным образом, а проверка исхода – только в критической области условий.

С учетом того что проверяемая методом статистических испытаний обычно уже довольно отработанная в ходе рабочего проектирования система, поведение которой может вызывать сомнения только при значительном, скажем, не менее чем на две трети допустимого, отклонении и, по крайней мере, двух определяющих факторов одновременно, сокращение числа фактических проверок при сохранении заданной достоверности оценок получается существенным – в десятки раз. В отношении эффективности метод выбора критических реализаций достоин того, чтобы возглавить список известных методов сокращения трудоемкости статистического эксперимента. Прежде всего это относится к математическому и полунатурному моделированию (с ростом размерности пространства контролируемых в эксперименте «случайных» факторов эффективность метода увеличивается).

В пределе идея метода реализуется в единичных проверках по принципу «два ваших – два наших» (варируются по два наиболее определяющих параметра объекта и системы управления). По оценке Н. П. Яковлева, такая проверка получается с запа-

сом и соответствует отходу более чем на четыре среднеквадратических отклонения от номинального поведения системы (при вариации на уровне «2σ» вероятность случайного попадания в критическую область имеет порядок $1/2,56 \cdot 10^6$).

В оригинальной версии метод выбора критических реализаций не претендует на решение задач в общей постановке – оценке математического ожидания некоторой функции или функционала качества системы. Сохранение трудоемкости решения задач такого общего вида сегодня обеспечивается другими известными методами.

Предмет нашего рассмотрения – сочетание идей методов «выделения главной части» и «расслоенной выборки» с принципом выбора критических реализаций для решения задачи в общей постановке.

О возможности сочетания метода выбора критических реализаций с другими методами

Наиболее известны метод выделения главной части и его варианты (корреляционная выборка, комбинированные методы определения вероятностных характеристик) [4, 5, 6]. Основная идея методов этой группы – использование упрощенной (в идеале – аналитической) модели системы, учитывающей влияние основных случайных факторов на ее поведение. Это позволяет свести обработку результатов статистического эксперимента с более точной моделью или с самой системой к оцениванию поправок к вероятностным характеристикам системы, определенным на упрощенной модели. Оценка математического ожидания вектора показателей качества, полученная на упрощенной модели, $m_{y,m} = M[Y_m]$, уточняется по выборке $\{Y\}$ реализаций более точной модели или натурального эксперимента с формированием выборки разностей между показателями качества системы и ее упрощенной модели $\Delta Y = Y - Y_m$.

Далее, $\hat{m}_y = m_{y,m} + M[\Delta Y]$ и используется то, что дисперсии разностей ΔY существенно меньше дисперсий самих показателей качества (наиболее существенное для исследователя в поведении системы, модель, даже упрощенная, должна предсказывать, иначе это – не модель).

Поскольку $D[\Delta Y] \ll D[Y]$, постольку и $D[m_y]$, приближенно равная $\frac{1}{n} D[\Delta Y]$, где n – объем выборки, получается много меньше, чем при прямой оценке m_y по выборке $\{Y\}$. Потребный объем выборки, оцениваемый по заданной относительной погрешности результата, уменьшается во столько раз, во сколько раз $D[\Delta Y]$ меньше $D[Y]$. Как именно выбираются сочетания случайных факторов в каждой реализации, при этом имеет второстепенное значение. Главное – чтобы они совпадали для полной (системы) и упрощенной модели и чтобы адекватность упрощенной модели обеспечивала значение коэффициента корреляции между Y и Y_m более 0,5.

Введение в методы выделения главной части отбора критических реализаций становится возможным при естественном, как представляется, уточнении требований к адекватности упрощенной модели. Если в области, близкой к номинальным характеристикам системы, возможные отклонения ΔY существенно меньше, чем на краях поля допусков, они не повлияют на результат и соответствующие реализации можно просто не проверять, приближенно полагая для них $\Delta Y = 0$.

Чтобы избежать необходимости доказательства малости возможного смещения оценки \hat{m}_y , понятие области критических реализаций также может быть уточнено. К ней целесообразно отнести все, что остается после исключения прилегающей к номинальным характеристикам области адекватности упрощенной модели (уточнение – только на периферии генеральной совокупности). Поскольку основная вероятностная масса обычно сосредоточена именно в центральной области, вокруг точки номинальных характеристик, характерный для выбора критических реализаций эффект существенного сокращения трудоемкости исследования обеспечивается.

Использование приближенной модели является существенным элементом и другой эффективной методики – так называемой «расслоенной выборки» (в работе [7] наряду с изложением основ приведена подробная библиография). Если обозначить X – случайный вектор параметров системы с функцией плотности вероятности $w(x)$ и $Y = f(X)$ – показатель качества системы, рассматриваемая задача сводится к оценке интеграла

$$J = \int_{\Omega} f(x)w(x)dx$$

В методе расслоенной выборки множество возможных значений случайного вектора X разбивается на L непересекающихся «слоев» $\Omega_1, \dots, \Omega_L$, и из каждого слоя Ω_h берется простая случайная выборка объемом n_h . В качестве оценки вероятностной характеристики системы используется взвешенная сумма послыжных оценок

$$\hat{J} = \sum_{h=1}^L \omega_h \hat{J}_h$$

где

$$\hat{J}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} y_i, \quad y_i = f(x_i)$$

при выборе $x_i \in \Omega_h$,

$$\omega_h = \int_{\Omega_h} w(x)dx.$$

Так как слои не пересекаются, то

$$D[\hat{J}] = \sum_{h=1}^L w_h^2 D[\hat{J}_h] = \sum_{h=1}^L \frac{w_h^2}{n_h} D_h.$$

Здесь $D_h = D[f(X)]$ при выборе $X \in W_h$.

При соответствующем выборе разбиения на слои и чисел n_1, n_2, \dots, n_L дисперсия $D[\hat{J}]$ существенно меньше, чем при простой случайной выборке и том же общем числе испытаний, и тем меньше, чем меньше значение D_h , т. е. чем более однородна выборка в каждом слое. Наибольший эффект дает первое разбиение множества Ω на два-три слоя. При дальнейшем дроблении без пересмотра ранее установленных границ дисперсия только убывает, хотя и с все меньшей скоростью.

Наиболее распространенными являются два метода разбиения выборки по слоям:

- 1) пропорциональная (типическая) выборка $n_h = w_h \cdot n$ (n – общий объем выборки);
- 2) оптимальная выборка

$$n_h = \frac{\omega_h \sqrt{D_h}}{\sum_{h=1}^L (\omega_h \sqrt{D_h})}.$$

Самой сложной задачей при планировании рас-слоенной выборки является определение границ слоев. Именно ее решение требует привлечения приближенной (упрощенной или использующей прошлый опыт) модели. Введем следующие обозначения:

y' – исход упрощенной модели ($y' = f_m(x)$);
 $w_m(y')$ – функция плотности вероятности исходов упрощенной модели в пространстве исходов Ω_h ;
 (y'_{h-1}, y'_h) – границы слоев, $h = 1, \dots, L$;
 D_{mh} – дисперсия исходов упрощенной модели в слое Ω_h ;

Тогда некоторые из рекомендуемых в работе [7] правил разбиения на слои имеют вид:

а)
$$\int_{y'_{h-1}}^{y'_h} \sqrt{w_m(y')} dy' = \frac{1}{L} \int_{\Omega_y} \sqrt{w_m(y')} dy';$$

б) $w_h \sqrt{D_{mh}} \cong \text{const}$

(в оптимальной выборке обеспечивается $n_h \gg \text{const}$);

в) $\omega_h (y'_h - y'_{h-1}) \cong \text{const}$.

Веса слоев также рекомендуется оценивать с помощью приближенных моделей.

Таким образом, необходимость использования приближенных моделей определяется как потребностями предшествующего этапа работы над системой, так и «собственными нуждами» методов исследования на заключительном этапе, и в методах «расслоенной выборки» – в не меньшей степени, чем в методах «выделения главной части»

Если приближенная модель обеспечивает хотя бы в одном слое Ω_h получение достаточно точной оценки \hat{J}_h (на периферии распределения и в номинале требования к точности приближенной модели естественно различаются), то в этом слое исследований с использованием точной модели или натурной системы можно просто не проводить без риска искажения общей оценки вероятностных характеристик системы.

Итак, в сочетании с известными методами уменьшения дисперсии результатов статистического эксперимента правомерно использовать метод выбора критических реализаций для еще более существенного сокращения трудоемкости исследований, в том числе при самой общей постановке задачи.

Литература

1. Яковлев Н. П., Богданов В. С., Куликов Ю. П. Способ уменьшения трудоемкости метода статистических испытаний при исследовании нелинейных систем автоматического управления // Морское приборостроение. Сер. VI. – вып. 1. – 1973. – С. 126–132.
2. Богданов В. С., Яковлев Н. П. Об одном приближенном методе вычисления вероятностей при исследовании нелинейных систем // Изв. АН СССР, Техническая кибернетика. – № 5. – 1975. – С. 97–101.
3. Малышев В. В. и др. Оптимизация наблюдения и управления летательных аппаратов. – М.: Машиностроение, 1989. – 312 с.
4. Ермаков С. М., Михайлов Г. А. Курс статистического моделирования: Учебное пособие. – Изд. 2-е. – М.: Наука, 1982. – 320 с.
5. Гнеденко Б. В., Коваленко И. Н. Введение в теорию массового обслуживания. – Изд. 2-е. – М.: Наука, 1987. – 136 с.
6. Пугачев В. Н. Комбинированные методы определения вероятностных характеристик. – М.: Советское радио, 1973. – 168 с.
7. Васильев Д. В., Сабинин О. Ю. Ускоренное статистическое моделирование систем управления. – Л.: Энергоатомиздат, 1987. – 136 с.