СТАТИСТИЧЕСКИЕ ЗАДАЧИ ПЛАНИРОВАНИЯ ПРОГОНА ИМИТАЦИОННОЙ МОДЕЛИ

В. Д. Левчук (Гомель, Беларусь)

1. Введение

Общепризнано, что при использовании имитационного моделирования планированию имитационных экспериментов необходимо уделять не меньшее внимание, чем разработке и программированию имитационной модели (ИМ) [1, 2]. Минимальным элементом в плане эксперимента является прогон (реплика). Методика расчета выходных данных в течение прогона модели зависит от наличия или отсутствия «естественного» события, определяющего окончание прогона.

Первый случай соответствует переходному режиму [2]. В нем исследователь задает начальное состояние ИМ и присваивает выходным данным начальные значения (как правило, нулевые). По свершению «естественного» события, соответствующего окончанию работы объекта моделирования, программа ИМ автоматически сохраняет в некотором информационном массиве рассчитанные значения, завершая тем самым прогон. Во втором случае, т. е. в непереходном (стационарном) режиме, исследователя интересует поведение системы в течение длительного периода «нормальной» работы. При этом расчет выходных данных в течение прогона наиболее часто реализуется по следующей схеме:

- как и в предыдущем случае, задается начальное состояние ИМ и присваиваются начальные значения выходным данным;
- по достижению стационарного режима работы фиксируется окончание переходного периода и выходным данным возвращаются начальные значения;
- прогон продолжается, пока не будет получено достаточное количество наблюдений выходных данных.

Таким образом, имеются, по крайней мере, три задачи планирования прогона в стационарном режиме: выбора начального состояния ИМ (А); определения модельного времени окончания переходного периода (Б) и всего прогона (В).

Большинство исследователей в силу простоты программирования предпочитает в качестве начального состояния использовать «пустое», когда в ИМ отсутствуют динамические объекты и свободны статические. Между тем выбор начального состояния, которое близко к стационарному или типовому, позволяет сократить длительность переходного периода.

Универсальных процедур для решения задач Б и В не существует. В каждом конкретном случае исследователь должен отдать предпочтение некоторому методу. Различные рекомендации приводятся в монографиях [2–4]. К сожалению, их анализ приводит к выводу, что более высокая точность оценки искомых параметров достигается за счет сохранения в памяти промежуточных наблюдений в течение всего прогона (и даже нескольких прогонов). Это приводит к неоправданно высоким объемам вычислений и затратам машинного времени.

Поэтому актуальными являются процедуры, предназначенные для решения задач Б и В, исходя из соображений компромисса между эффективностью оценки параметров прогона и объемом хранимых данных. В настоящей статье рассматриваются методы проверки разностей средних уровней и разброса дисперсии, позволяющие оценить время окончания переходного периода в ИМ. Для определения интервала моделирования предлагаются два варианта метода сравнения с контрольным образцом. Данный метод позволяет распространить значение времени окончания прогона, найденное по ограниченной свокупности конфигураций ИМ, на все множество ее рабочих начальных состояний.

2. Методы определения времени окончания переходного периода в ИМ

Для решения поставленной задачи следует выбрать типовое состояние изучаемого процесса в качестве начального и провести эксперимент из 3–5 прогонов, но на большом интервале (в идеале максимально возможном) модельного времени. Математической моделью реализации эксперимента при единственном отклике служит вектор временных рядов, т.е. вектор стохастических процессов в дискретном времени.

Первый этап решения задачи заключается в графическом представлении множества временных рядов в едином модельном времени. Для его выполнения можно порекомендовать как офисный табличный процессор Excel, так и специализированные пакеты статистической обработки данных (в частности, StatGraph, Statistica). В результате анализа графиков исследователь должен сделать три предположения:

- об отсутствии аномальных уровней во временных рядах, не отвечающих потенциальным возможностям изучаемого процесса;
- о значении модельного времени T^*, начиная с которого отсутствует длительная тенденция изменения отклика;
- о выборе наиболее характерного временного ряда для подтверждения гипотезы об отсутствии тренда.

На втором этапе исследователь должен в результате постановки имитационного эксперимента подтвердить или отвергнуть гипотезу об отсутствии тренда после момента T^*. Для этого предлагаются два метода [5].

Метод проверки разностей средних уровней

Реализация этого метода состоит из трех шагов. На первом шаге исходный временной ряд $y_1 = y(t_1)$, $y_2 = y(t_2)$, …, $y_n = y(t_n)$, из которого удалены значения до момента T^*, соответствующие предполагаемому переходному периоду, разбивается на две примерно равные по числу уровней части: в первой части n_1 первых уровней исходного ряда, во второй – n_2 остальных уровней ($n_1 + n_2 = n$).

Второй шаг заключается в проверке равенства (однородности) дисперсий обеих частей ряда с помощью F-критерия Фишера [6], которая основана на сравнении расчетного значения этого критерия с табличным (критическим) значением критерия Фишера F_α с заданным уровнем значимости (уровнем ошибки) α. Если расчетное значение меньше табличного F_α, то гипотеза о равенстве дисперсий принимается, и переходят к следующему шагу. Если расчетное значение больше или равно F_α, гипотеза о равенстве дисперсий отклоняется, и делается вывод, что данный метод для определения отсутствия тренда ответа не дает.

На третьем шаге проверяется гипотеза об отсутствии тренда с использованием t–критерия Стъюдента [6]. Для этого определяется расчетное значение критерия Стьюдента по формуле:

$$t = \frac{\left|\overline{y_1} - \overline{y_2}\right|}{\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\sqrt{\frac{(n_1 - 1)\sigma_1^2 + (n_2 - 1)\sigma_2^2}{n_1 + n_2 - 2}}}$$

Если расчетное значение t меньше табличного значения статистики Стьюдента t_α, взятого при $n_1 + n_2 - 2$ степеней свободы с заданным уровнем значимости α, гипотеза принимается, т. е. тренда нет, и предположение о значении T^* – времени окончания переходного периода является правильным. В противном случае, тренд есть и можно попробовать повторить процедуру для большего значения модельного времени.

Метод проверки разброса дисперсии

Этот метод обладает большими возможностями и дает более надежные результаты по сравнению с предыдущим. Кроме тренда самого ряда (или тренда в среднем), он позволяет установить наличие тренда дисперсии временного ряда: если тренда дисперсии нет, то разброс уровней ряда постоянен; если дисперсия увеличивается, то ряд «раскачивается» и т. д.

Реализация метода содержит четыре шага. Как и в предыдущем случае из исходного временного ряда отсекается та его часть, которая соответствует переходному периоду. На первом шаге производится сравнение каждого уровня полученного временного ряда, начиная со второго уровня, со всеми предыдущими, при этом определяются две числовые последовательности:

$$k_t = \begin{cases} 1, \text{ если } y_t \text{ больше всех предыдущих уровней;} \\ 0, \text{ в противном случае} \end{cases}$$

$$l_t = \begin{cases} 1, \text{ если } y_t \text{ меньше всех предыдущих уровней;} \\ 0, \text{ в противном случае} \end{cases}$$

$$t = 2,3,...,n$$

На втором шаге вычисляются величины s и d:

$$s = \sum_{t=2}^{n}(k_t + l_t), d = \sum_{t=2}^{n}(k_t - l_t)$$

Величина s характеризует изменение временного ряда и принимает значения от 0 (все уровни ряда равны между собой) до n–1 (ряд монотонный). Величина d характеризует изменение дисперсии уровней временного ряда и изменяется от –(n–1) (ряд монотонно убывает) до (n–1) (ряд монотонно возрастает).

Третий шаг заключается в проверке гипотез: можно ли считать случайными

- отклонение величины s от величины μ – математического ожидания величины s для ряда, в котором уровни расположены случайным образом;
- отклонение величины d от нуля.

Эта проверка проводится с использованием расчетных значений t–критерия Стьюдента для средней и для дисперсии:

$$t_s = \frac{|s-\mu|}{\sigma_1}; \mu = \sqrt{20\ln n - 32}; \sigma_1 = \sqrt{2\ln n - 3.4253};$$

,

$$t_d = \frac{|d|}{\sigma_2}; \sigma_2 = \sqrt{2\ln n - 0.8456}$$

- где μ – математическое ожидание величины s, определенной для ряда, в котором уровни расположены случайным образом;
- σ_1 – среднеквадратическое отклонение для величины s;
- σ_2 – среднеквадратическое отклонение для величины d.

На четвертом шаге расчетные значения t_s и t_d сравниваются с табличным значением t–критерия Стьюдента с заданным уровнем значимости t_α. Если расчетное значение меньше табличного, то гипотеза об отсутствии тренда принимается, а в противном случае тренд есть. Например, если t_s больше табличного значения t_α, а t_d меньше t_α, то для данного временного ряда имеется тренд в среднем, а тренда дисперсии уровней ряда нет. Что, в конечном итоге, требует нового предположения о времени окончания переходного периода в ИМ.

3. Метод определения времени окончания прогона ИМ

Предлагаемый метод состоит в организации серии экспериментов, характеризующихся одинаковым начальным состоянием, количеством прогонов, временем окончания переходного периода, но различными интервалами реализации прогонов в разных экспериментах.

На первом шаге следует задать T_0 – максимально разумный (возможный) интервал реализации эталонного прогона. Желательно, чтобы количество прогонов в эксперименте обеспечивало выборку значений отклика, подтверждающую гипотезу о нормальном законе распределения отклика с заданным уровнем значимости. По результатам постановки первого эксперимента исследователь получает эталонную выборку значений отклика $\{Y_0\}$, которая будет использоваться далее.

На втором шаге следует выбрать «минимально разумное» время интервала моделирования и повторить опыт с тем же начальным состоянием и количеством прогонов. В результате будет получена выборка значений отклика $\{Y_1\}$.

На третьем шаге необходимо проверить с помощью одного или нескольких критериев (параметрических при нормальности распределения отклика и непараметрических в противном случае) гипотезу об однородности выборок $\{Y_0\}$ и $\{Y_1\}$, т.е. принадлежат ли они одной и той же генеральной совокупности [6]. Если гипотеза не подтверждается, то решение о минимальном интервале моделирования оказалось ошибочным. Поэтому следует задать значение $T_2>T_1$ и повторить шаги 2 и 3, где в качестве эталонной по–прежнему должна использоваться выборка $\{Y_0\}$. В результате конечного числа итераций будет найдено минимальное время окончания прогона ИМ.

В идеале можно обойтись единственным экспериментом. В каждом прогоне необходимо сохранять значения отклика в моменты T_1, T_2, … , останавливая прогон в момент T_0. Проверка гипотез будет выполняться после завершения эксперимента. В данном случае всегда можно выполнить любое количество дополнительных прогонов. Проблема заключается в том, что подобные инструменты трассировки прогона не поддерживаются многими средствами автоматизации моделирования.

Любой из вариантов предлагаемого метода следует проверить для нескольких конфигураций изучаемой сложной системы. В подавляющем большинстве случаев время окончания прогона, найденное по тестовой совокупности конфигураций, можно использовать в различных имитационных экспериментах с этой ИМ.

4. Апробация методов

Тестовые результаты эксперимента брались из ИМ с непереходным режимом различной степени сложности. Методы, изложенные в данной статье, были реализованы в среде табличного процессора Excel в виде макросов на языке программирования VBA. Оценки параметров прогона, рассчитанные с помощью данных методов, сравнивались с оценками, полученными с помощью процедуры Велча, приведенной в [2]. Существенных расхождений не было обнаружено. Однако при этом для процедуры Велча требовалось выполнять большее количество прогонов ИМ.

Таким образом, результаты апробации позволяют сделать вывод о рекомендации предложенных методов для решения статистических задач планирования прогона с ИМ сложных систем.

Литература

1. Технология системного моделирования/**Е. Ф. Аврамчук, А. А. Вавилов, С. В. Емельянов и др.** Под общ. ред. С.В. Емельянова, В.В. Калашникова и др. – М.: Машиностроение; Берлин: Техник, 1988. – 520 с.
2. **Лоу А., Кельтон** *В*. Имитационное моделирование. Классика CS. 3-е изд. – СПб.: Питер; Киев: BHV, 2004. – 847 с.
3. **Прицкер А.** Введение в имитационное моделирование и язык СЛАМ II. – М.: Мир, 1987. – 646 с.
4. **Максимей И. В.** Имитационное моделирование на ЭВМ. – М.: Радио и связь, 1988. – 232 с.
5. Экономико–математические методы и прикладные модели: Учеб. пособие для вузов/Под ред. В.В. Федосеева. – М.: ЮНИТИ, 2001. – 391 с.
6. **Шеннон Р.** Имитационное моделирование систем – искусство и наука. – М.: Мир, 1978. – 418 с.