Лекция 4

Планирование модельного эксперимента

Автор Джомартова Ш.А.

Цели планирования эксперимента

- сокращение общего объема испытаний при соблюдении требований к достоверности и точности их результатов;
- повышение информативности каждого из экспериментов в отдельности.

Факторное пространство – это множество внешних и внутренних параметров модели, значения которых исследователь может контролировать в ходе подготовки и проведения модельного эксперимента.

Стратегическое планирование

- Цель – получение максимального объема информации об исследуемой системе в каждом эксперименте
- идентификация факторов;
- выбор уровней факторов.

Способы построения стратегического плана

- Эксперимент, в котором реализуются все возможные сочетания уровней факторов называется полным факторным экспериментом (ПФЭ)
- Общее число различных комбинаций уровней в ПФЭ для k факторов можно вычислить так: $N = l1 \cdot l2 \cdot l3 \cdot \ldots \cdot lk$. Здесь li – число уровней i-го фактора.
- Недостаток ПФЭ – большие временные затраты на подготовку и проведение

Способы построения стратегического плана

- Рандомизированный план – предполагает выбор сочетания уровней для каждого прогона случайным образом. При использовании этого метода отправной точкой в формировании плана является число экспериментов, которые считает возможным (или необходимым) провести исследователь.

Способы построения стратегического плана

- Латинский план (или «латинский квадрат») – используется в том случае, когда проводится эксперимент с одним первичным фактором и несколькими вторичными. Суть такого планирования состоит в следующем. Если первичный фактор А имеет l уровней, то для каждого вторичного фактора также выбирается l уровней.

Способы построения стратегического плана

- Дробный факторный эксперимент. Каждый фактор имеет два уровня – нижний и верхний, поэтому общее число вариантов эксперимента $N=2k$, где k – число факторов.

Тактическое планирование эксперимента

- Совокупность методов установления необходимого объема испытаний относят к тактическому планированию экспериментов.
- Поскольку точность оценок наблюдаемой переменной характеризуется ее дисперсией, то основу тактического планирования эксперимента составляют так называемые методы понижения дисперсии.

Формирование простой случайной выборки

- Многократный прогон модели для каждого сочетания уровней факторов, выбранного на этапе стратегического планирования эксперимента называют «формирование простой случайной выборки».
- Если требуемое значение дисперсии Dy до начала эксперимента неизвестно, целесообразно выполнить пробную серию из L прогонов и вычислить на ее основе выборочную дисперсию, значение которой подставляют в выражение и получают предварительную оценку числа прогонов N_T.

Методы понижения дисперсии

- активные (предусматривают формирование выборки специальным образом);
- пассивные (применяются после того, как выборка уже сформирована);
- косвенные (в которых для получения оценок наблюдаемой переменной и используются значения некоторых вспомогательных величин).

Активные методы

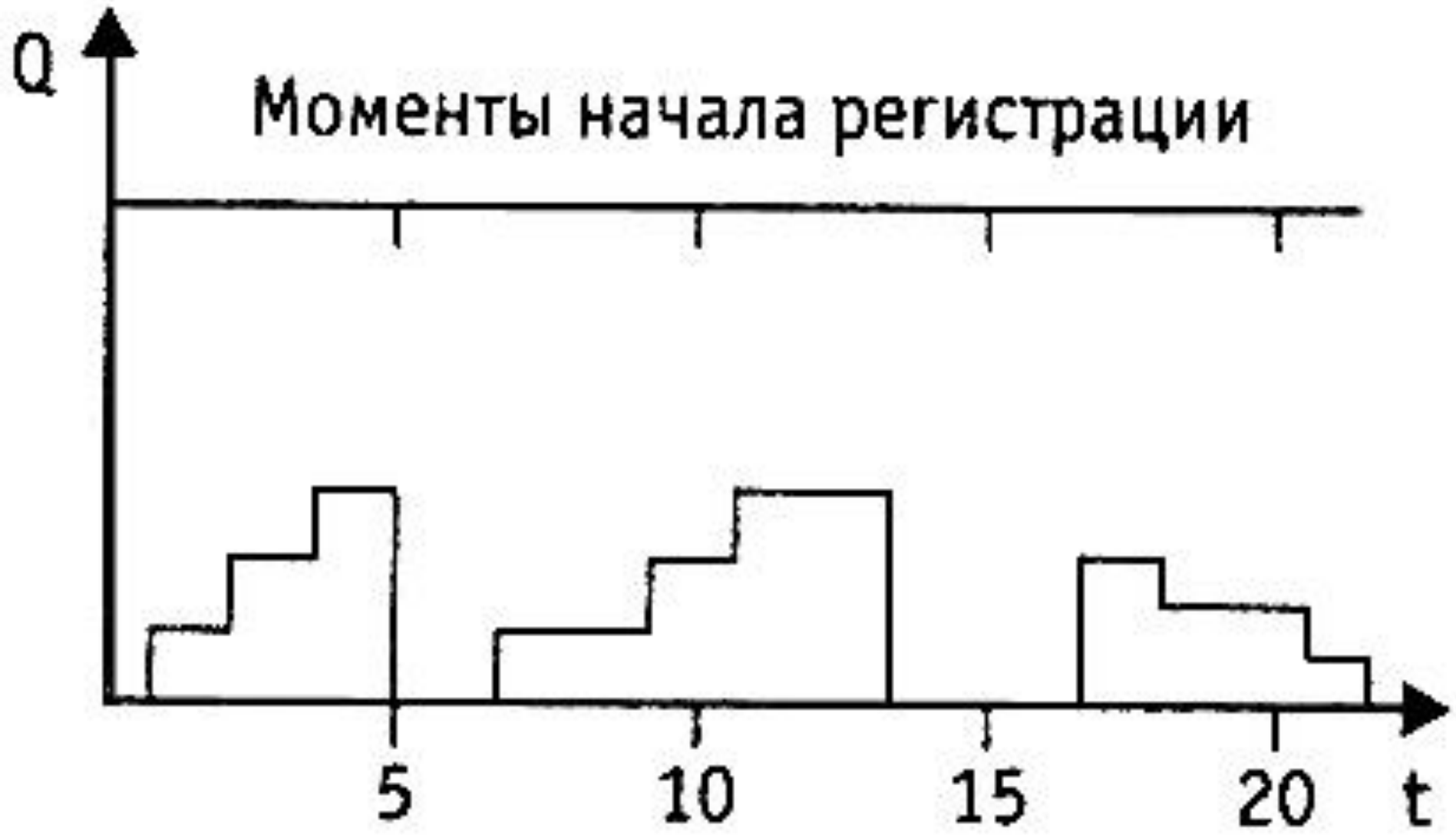
- методом повторения;
- методом подинтервалов;
- методом цикла.

Метод повторений

- При использовании этого метода каждое наблюдение получается при помощи отдельного прогона модели, причем все прогоны начинаются при одних и тех же начальных условиях, но используются различные последовательности случайных чисел.
- Преимуществом метода является статистическая независимость получаемых наблюдений. Недостаток состоит в том, что наблюдения могут оказаться сильно смещенными под влиянием начальных условий.

Метод подинтервалов

- Данный метод основан на разбиении каждого прогона модели на равные промежутки времени. Начало каждого интервала совпадает с началом очередного этапа наблюдений.
- Достоинство метода состоит в том, что влияние переходных условий со временем уменьшается, и наблюдения точнее отражают поведение системы в стационарном режиме.
- Недостаток – значения наблюдаемых переменных, полученные в начале очередного интервала, зависят от конечных условий предыдущего интервала.


Q
Моменты начала регистрации
5
10
15
20
t

Метод циклов

- При использовании метода циклов влияние автокорреляции уменьшается за счет выбора интервалов таким образом, чтобы в их начальных точках условия были одинаковыми.
- Недостатком метода является меньшее по сравнению с методом подинтервалов число получаемых наблюдений.


Q
Моменты начала регистрации
5
10
15
20
t

Пассивные методы

Метод стратифицированной выборки

$$y = \frac{1}{N} \sum_{i=1}^{k} N_i \cdot y_i$$

$$D_y = \frac{1}{N} \sum_{i=1}^{k} N_i \cdot D_{yi}$$

Обработка и анализ результатов моделирования

- полученные результаты обладают требуемой точностью и достоверностью;
- исследователь способен правильно интерпретировать полученные результаты и знает, каким образом они могут быть использованы

Оценка качества имитационной модели

Цели:

- проверить соответствие модели ее предназначению (целям исследования);
- оценить достоверность и статистические характеристики результатов, получаемых при проведении модельных экспериментов.

При аналитическом моделировании

- корректный выбор математического аппарата, используемого для описания исследуемой системы;
- методическая ошибка, присущая данному математическому методу.

При имитационном моделировании

- моделирование случайных факторов, основанное на использовании датчиков случайных чисел, которые могут вносить «искажения» в поведение модели;
- наличие нестационарного режима работы модели;
- использование нескольких разнотипных математических методов в рамках одной модели;
- зависимость результатов моделирования от плана эксперимента;
- необходимость синхронизации работы отдельных компонент модели;
- наличие модели рабочей нагрузки, качество которой зависит, в свою очередь, от тех же факторов.

Целевые свойства имитационной модели

- адекватность;
- устойчивость;
- чувствительность.

Оценка адекватности

Способы оценки адекватности:

- по средним значениям откликов модели и системы;
- по дисперсиям отклонений откликов модели от среднего значения откликов системы;
- по максимальному значению относительных отклонений откликов системы.

Оценка устойчивости

- Устойчивость модели – это ее способность сохранять адекватность при исследовании эффективности системы на всем возможном диапазоне рабочей нагрузки, а также при внесении изменений в конфигурацию системы.

Как оценить устойчивость модели?

- Универсальной процедуры проверки устойчивости модели не существует. Разработчик вынужден прибегать к методам «для данного случая», частичным тестам и здравому смыслу. Часто полезна апостериорная проверка. Она состоит в сравнении результатов моделирования и результатов измерений на системе после внесения в нее изменений. Если результаты моделирования приемлемы, уверенность в устойчивости модели возрастает.

Оценка чувствительности

1. Вычисляется величина относительного среднего приращения параметра X:

$$\Delta X = \frac{(X_{max} - X_{min}) \cdot 2}{X_{max} + X_{min}} \cdot 100\%$$

Оценка чувствительности

2. Проводится пара модельных экспериментов при значениях X=Xmax и X=Xmin и средних фиксированных значениях остальных параметров. Определяются значения отклика модели Y1=f(Xmax) и Y2=f(Xmin).

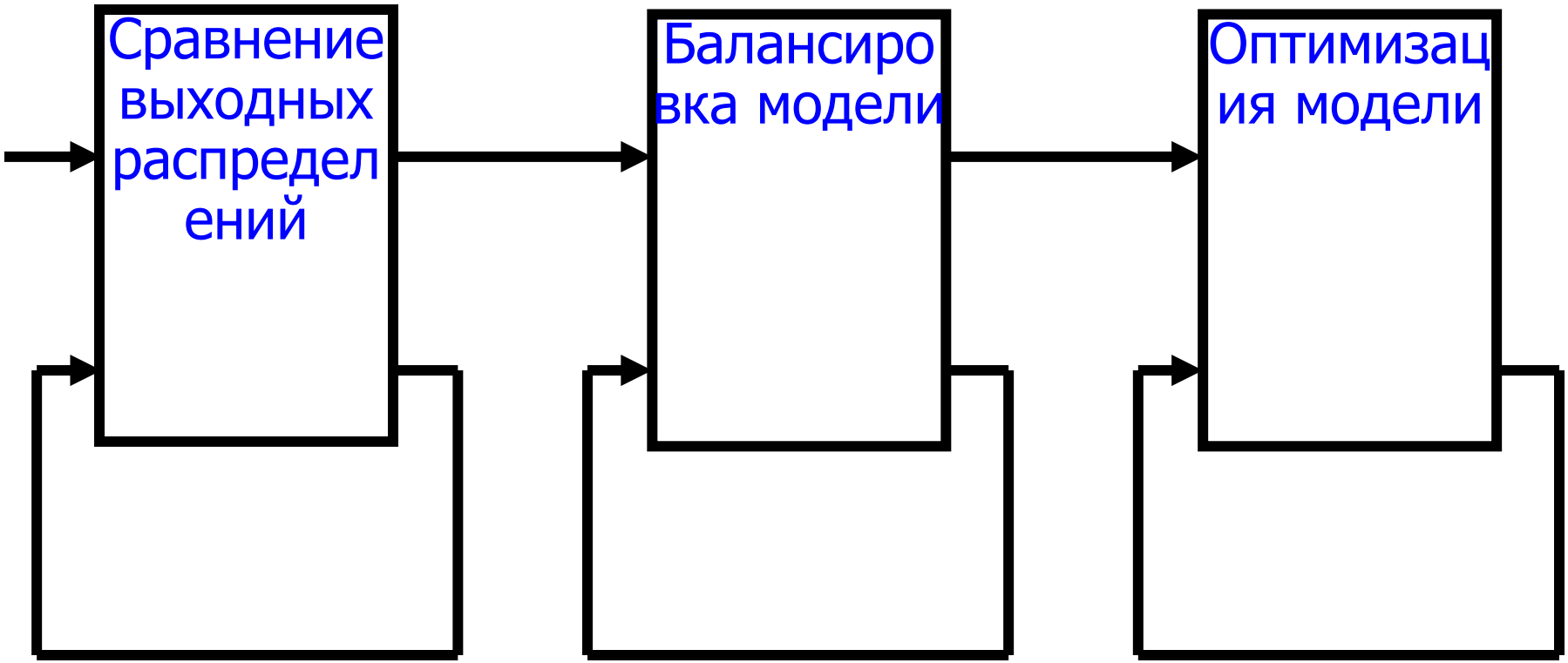
Оценка чувствительности

3. Вычисляется относительное приращение наблюдаемой переменной Y:

$$\Delta Y = \frac{|Y_1 - Y_2| \cdot 2}{Y_1 + Y_2} \cdot 100\%$$

В результате для k-го параметра модели имеют пару значений (ΔXk, ΔYk), характеризующую чувствительность модели по этому параметру.

Калибровка модели



Сравнение выходных распределений

- Цель – оценка адекватности ИМ. Устранение различий на этом шаге основано на внесении глобальных изменений.
- Глобальные изменения модели (например, введение новых процессов, изменение типов событий и т. д.)

Балансировка модели

- Основная задача – оценка устойчивости и чувствительности модели. По его результатам, как правило, производятся локальные изменения (но возможны и глобальные).
- Локальные изменения (в частности, изменение некоторых законов распределения моделируемых случайных величин)

Оптимизация модели

- Цель этого этапа – обеспечение требуемой точности результатов. Здесь возможны три основных направления работ:
  - дополнительная проверка качества датчиков случайных чисел;
  - снижение влияния переходного режима;
  - применение специальных методов понижения дисперсии.

Подбор параметров распределения

- Статическая гипотеза – это утверждение относительно значений одного или более параметров распределения некоторой величины или о самой форме распределения.
- Выбирают H0 – основную гипотезу и H1 – альтернативную.

- Статистическая проверка гипотезы – это процедура выяснения, следует ли принять основную гипотезу H0 или отвергнуть ее.
- Если H0 отвергается, то ошибка первого рода (характеризуется более тяжелыми последствиями); если H0 принимается при истинности H1 – ошибка второго рода.
- Вероятности ошибок 1 и 2 рода зависят от критерия, на основании которого будет выбираться одна из гипотез.

Проверка гипотезы в следующем:

- Производится выборка (эксперимент)
- Вычисляется z – контрольная величина (критерий), представляющая собой некоторую функцию от выборки (результат эксперимента)
- Если z из области B

$$Ð(Z \in B | H0\ \hat{a}\mathring{a}\eth\acute{i}\grave{a}) \leq \alpha$$

то от гипотезы H0 отказываются.

t–критерий

$$T = \frac{\bar{x} - \bar{y}}{\sqrt{(n_1 - 1)D_x + (n_2 - 1)D_y}} \sqrt{\frac{n_1 - n_2 - (n_1 + n_2 - 2)}{n_1 + n_2}}$$

$$\grave{O} \geq t_{kp}$$

F – критерий

$$F = \frac{D_x}{D_y}$$

$$m_1 = n_1 - 1; m_2 = n_2 - 1$$

$$a = \frac{\alpha}{2}$$

$$F \geq F_{a,m_1,m_2}$$

Критерий согласия

- Область значений разбивается на K непересекающихся классов
- Из n опытов формируется выборка
- Вычисляют

$$\chi^2 = \left(\sum_{i=1}^{K} \frac{M_i^2}{np_i} \right) - n$$

$$\chi^2 \geq \chi_a^2$$

Оценка влияния и взаимосвязи факторов

- Необходима количественная оценка влияния фактора на наблюдаемую переменную
- Однофакторный дисперсионный анализ
- Многофакторный дисперсионный анализ
- Корреляционны и регрессионный анализы

Однофакторный дисперсионный анализ

- Пусть x факторов, l –уровней

$$y_i(1), y_i(2), \ldots, y_i(l), i = \overline{1, n}$$

- Проверяется гипотеза о независимости значений y от значений x
- Уравнение однофакторного дисперсионного анализа

$$y_{ij} = m + a_i + e_{ij}$$

Многофакторный дисперсионный анализ

- Пусть есть 1 первичный и 2 вторичных фактора, n –уровней
- Обозначим y_{ijk} результат эксперимента при условии, что фактор а находится на уровне i, фактор в – на уровне j, фактор с – на уровне k
- Уравнение многофакторного дисперсионного анализа

$$y_{ijk} = m + a_i + b_j + g_k + e_{ijk}$$

Многофакторный дисперсионный анализ

- Проверяется гипотеза о независимости результатов от факторов a, в и c

$$H_a : a_i = 0, i = \overline{1, n}$$

$$H_b : b_j = 0, j = \overline{1, n}$$

$$H_c : g_k = 0, k = \overline{1, n}$$

$$SS = \alpha(y_{ijk} - m + a_i + b_j + g_k + e_{ijk})^2$$

Корреляционный анализ

- **Результаты корреляции позволяют делать выводы о степени зависимости**

Оценка величины: $r_{xy} = \frac{\sum D_x D_y}{\sqrt{\sum D_x^2 \sum D_y^2}}$

- Малая(слабая) $0 < |r_{xy}| < 0{,}29$
- Средняя(умеренная) $0{,}3 < |r_{xy}| < 0{,}69$
- Большая(сильная) $0{,}7 < |r_{xy}| < 1$

Регрессионный анализ

Позволяет уточнить:

- Установить наличие возможной причинной связи между переменными
- Предсказывать значение переменных по значениям независимых переменных

Регрессионный анализ

Простая линейная регрессия:

$$y_i = b_0 + b_1 x_i + e_i$$

Спасибо за внимание