

СИСТЕМНЫЕ СТАТИСТИЧЕСКИЕ СВОЙСТВА ЭКСПЕРИМЕНТАЛЬНОГО МЕТОДА ИССЛЕДОВАНИЯ

Abstract: It is shown that multifactor statistical mathematical models obtained by results of experimental investigation are the main form of description of complex systems, processes, objects. System statistical properties of the experimental method of investigation have been formulated and studied. The author shows theoretical schemes of results substantiation to which the mentioned system statistical properties lead.

Key words: experimental method of investigation, statistical properties, systems analysis.

Аноація: Показано, що основною формою опису складних систем, процесів, об'єктів є багатофакторні статистичні математичні моделі, які одержують за результатами експериментального дослідження. Сформульовано та досліджено системні статистичні властивості експериментального методу дослідження. Показано, до яких теоретичних схем обґрунтування результатів приводять вказані системні статистичні властивості.

Ключові слова: експериментальний метод дослідження, статистичні властивості, системний аналіз.

Аннотация: Показано, что основной формой описания сложных систем, процессов, объектов являются многофакторные статистические математические модели, получаемые по результатам экспериментального исследования. Сформулированы и исследованы системные статистические свойства экспериментального метода исследования. Показано, к каким теоретическим схемам обоснования результатов приводят указанные системные статистические свойства.

Ключевые слова: экспериментальный метод исследования, статистические свойства, системный анализ.

1. Введение

Многофакторные статистические математические модели получили значительное распространение при исследовании сложных систем, процессов, объектов. Статистические модели позволяют прогнозировать поведение системы, управлять ею, оптимизировать, проводить коррекцию цели и поверхности отклика. Использование моделей позволяет свести к возможному минимуму потребление физических ресурсов: вещественных, энергетических, пространственных, временных. Модели позволяют создать системные ресурсы – функциональные, целевые, эмергентные, оптимизационные, которые существенно изменяют критерии качества системы и делают ее более эффективной. Статистические модели особенно необходимы в тех случаях, когда традиционные конструкторские, технологические и эксплуатационные решения, основанные только на физических принципах, исчерпаны или приводят к невыполнимо большим затратам.

Экспериментальный метод распространен в физических, технических, технологических, агробиологических исследованиях, когда использование теоретического подхода затруднено ввиду сложности, малой изученности объекта исследования, комплексности происходящих разнородных явлений, их многофакторности, многокритериальности. Моделью сложного объекта является модель черного ящика, в которой исследователю доступны входы – управляемые факторы X_1, X_2, \dots, X_k – и выходы – критерии качества системы – y_1, y_2, \dots, y_m . Если оператор f , связывающий входы и выходы $\hat{y} = f(X_1, X_2, \dots, X_k)$, предполагается существующим, то он считается неизвестным и его необходимо определить. Информацию о сложной системе исследователь получает путем проведения с нею экспериментов и фиксирования результатов по выбранным критериям качества.

При проведении эксперимента в лабораторных условиях, на стенде и в других условиях, когда им можно управлять, необходимо использовать методы теории планирования эксперимента.

2. Постановка проблемы

Перечислим системные статистические свойства экспериментального метода исследования.

1. Результаты эксперимента являются следствием суммарного влияния групп управляемых, неуправляемых и неконтролируемых факторов [1, с. 20–21]. Число управляемых факторов может достигать 12...20. Число неуправляемых и неконтролируемых факторов сравнительно большое и исследователю не известно. Физические свойства изучаемой системы, процесса, объекта не могут быть определены полностью, т. е. исчерпывающе, и содержат неопределенности.

2. Результаты экспериментов сопровождаются погрешностями – систематическими, случайными, грубыми (выбросами). Случайные погрешности есть следствие влияния неуправляемых, неконтролируемых факторов и ошибок уровней управляемых факторов. Результаты измерений также являются случайными величинами. Обработка результатов экспериментов сопровождается вычислительными погрешностями.

3. Результаты экспериментов являются численными величинами без указания каких-либо теоретических свойств, а также вида (структуры) связи между условиями эксперимента и его результатами.

4. При проведении экспериментов реализуется индуктивный метод познания – от частного полученного результата к его обобщению на все гипотетические аналогичные результаты.

5. Предполагается, что комплекс условий проведения эксперимента является типичным для класса изучаемых систем (процессов, объектов), и поэтому полученный результат может быть распространен на другие представители изучаемого класса систем.

3. Концепция решения

Рассмотрим, к каким теоретическим схемам обоснования результатов приводят системные свойства экспериментального метода исследования. Будем придерживаться приведенной последовательности свойств 1...5.

1. Исследование влияния группы управляемых факторов X_1, X_2, \dots, X_k проводится при изменчивости неуправляемых и неконтролируемых факторов. При многофакторной схеме эксперимента необходимо использовать такой план эксперимента, который позволяет представить в математической модели эффекты – главные и взаимодействия – ортогонально друг другу или близко к ортогональным [1, с. 66–78, 146–169]. Эксперимент, в общем случае, проводится в неоднородных условиях. Поэтому необходима рандомизация последовательности опытов, т.е. проведение их в случайном порядке, разбиение плана эксперимента на ортогональные блоки.

Если указанные условия проведения многофакторного эксперимента не выполняются, то решаемая экспериментальная задача будет некорректно поставленной.

Формирование результатов опытов как суммарное влияние групп управляемых, неуправляемых и неконтролируемых факторов привело к решению класса так называемых обратных задач.

Н. Винер, А. Розенблют, М. Планк и другие ученые рассматривали постановку эксперимента как задание определенного вопроса. «Любой эксперимент – всегда некий вопрос. Если вопрос не точен, получить точный ответ на него трудно. Глупые ответы, т. е. противоречивые, расходящиеся друг с другом или не относящиеся к делу irrelevantные (неуместные – P.C.) результаты экспериментов, обычно указывают на то, что сам вопрос был поставлен глупо» [2, с. 171].

Использование методологии теории планирования эксперимента [3; 1, с. 117–119] необходимо при решении многофакторных экспериментальных задач. «Глупая» постановка «вопроса» означает, что экспериментальная задача является некорректно поставленной; это может выражаться в значительной коррелированности эффектов математической модели между собой. Последнее может быть обусловлено нестандартной областью факторного пространства, малым числом проведенных опытов по сравнению с числом определяемых эффектов и другими причинами.

Выбор устойчивого (робастного) последовательного плана проведения многофакторного эксперимента необходим для решения корректно поставленной задачи.

2. Результаты экспериментов рассматриваются как случайные величины. Систематические ошибки должны быть исключены. Если в результатах присутствуют грубые ошибки (выбросы), то их, по возможности, необходимо исключить или использовать для обработки устойчивые (робастные) методы оценивания.

Случайный характер результатов обуславливает использование статистических методов оценивания математической и прикладной статистики: дисперсионного, корреляционного, регрессионного анализов; теории планирования экспериментов и др.

Важнейшим требованием при обработке случайных исходных данных является использование методов устойчивого оценивания [1, с. 14–16]:

- робастный последовательный план эксперимента;
- формализованный выбор устойчивых структур многофакторных статистических моделей, заранее не известных исследователю;
- устойчивое оценивание коэффициентов статистических моделей, в том числе и в условиях исходной мультиколлинеарности факторов.

Подробное рассмотрение решения указанных проблем приведено в [1].

3. При формализации результатов экспериментов в виде математических моделей и выводов используются системы предпосылок относительно некоторых свойств рассматриваемых сложных систем, теоретический аппарат логического вывода и его реализация, формулировка полученных следствий (результатов).

Комплекс условий, реализуемый в эксперименте, не является формально заданным, а следовательно, может уточняться и изменяться.

Реализация выбора предпосылок и начальных условий, осуществляемая в чисто математическом подходе и для решения прикладных задач, может принципиально различаться по поставленным целям и фактически достигнутым результатам [1, с. 56–57]. На важность соответствия принятых предпосылок фактическим условиям обращали внимание французский физик Л. Бриллюэн [4, с. 96, 187], д.ф.-м.н. Е.С. Вентцель (Грекова И.) [5, с. 111], д.э.н. В.В. Швырков [6, с. 7].

При решении прикладных задач необходимо не только сформулировать систему необходимых предпосылок, но и методики их проверок; устойчивость предпосылок и метода получения результатов к сравнимо малым нарушениям принятых условий; систему действий исследователя, если принятые предпосылки не выполняются фактически. Рассмотрение указанных задач приведено в [1, с. 57–65].

При получении многофакторной математической модели необходимо выбрать вид связи (структуру) между условиями эксперимента – факторами X_1, X_2, \dots, X_k – и его результатами – критериями качества $y_1, y_2, \dots, y_w, \dots, y_m$ моделируемой системы.

$$\hat{y}_w = f_w(X_1, X_2, \dots, X_k),$$

где \hat{y}_w – модель w -го критерия качества системы (процесса, объекта); f_w – структура математической модели; k – общее число факторов.

В большинстве решаемых прикладных задач вид структуры статистической модели исследователю заранее не известен.

В [1, с. 106–108] предложена и обоснована формализованная структура многофакторной полиномиальной математической модели, линейной относительно параметров, в виде множества эффектов схемы полного факторного эксперимента:

$$(1 + X_1 + X_1^2 + \dots + X_1^{s_1-1}) \times \dots \times (1 + X_k + X_k^2 + \dots + X_k^{s_k-1}) \rightarrow N_{\Pi}, \quad (1)$$

где 1 – значение фиктивной независимой переменной $x_0 \equiv 1$; X_1, \dots, X_k – факторы искомой математической модели в натуральных значениях; s_1, \dots, s_k – число уровней факторов X_1, \dots, X_k соответственно; N_{Π} – число опытов полного факторного эксперимента, равное числу структурных элементов его схемы $N_{\Pi \text{ СТР}}$.

Выражение (1) используется для создания системы базисных функций. Оно представляет собой совокупность всех многочленов факторов $X_1, X_2, \dots, X_i, \dots, X_k$, каждой степени от 1 до наивысшей $s_i - 1$, и различных взаимодействий этих факторов (точнее степеней этих факторов) по два, три и т.д. до максимально возможных взаимодействий из k различных элементов. Совокупность всех указанных элементов образует вещественное евклидово пространство.

В выражении (1) приведены все возможные эффекты в полиномиальном виде, необходимые и достаточные для адекватной аппроксимации исходных данных схем полных и дробных факторных экспериментов.

При переходе от натуральных значений факторов X_1, \dots, X_k к системе ортогональных полиномов Чебышева (системе ортогональных контрастов) структура математической модели имеет вид

$$(1 + x_1^{(1)} + x_1^{(2)} + \dots + x_1^{(s_1-1)}) \times \dots \times (1 + x_k^{(1)} + x_k^{(2)} + \dots + x_k^{(s_k-1)}) \rightarrow N_{\Pi},$$

где $x_1^{(1)}, \dots, x_1^{(s_1-1)}; \dots; x_k^{(1)}, \dots, x_k^{(s_k-1)}$ – ортогональные контрасты факторов X_1, \dots, X_k ; описание других обозначений приведено выше.

Число структурных элементов $N_{\Pi \text{ СТР}}$ схемы полного факторного эксперимента равно числу опытов полного факторного эксперимента N_{Π} : $N_{\Pi \text{ СТР}} = N_{\Pi}$. Все эффекты для схемы полного факторного эксперимента ортогональны друг к другу. Следовательно, любой полный факторный эксперимент описывается устойчивой математической моделью.

В случае дробного факторного эксперимента число опытов N_D меньше числа структурных элементов $N_{\text{ПСТР}}$ и, следовательно, из $N_{\text{ПСТР}}$ возможных базисных функций необходимо выбрать k' ($k' < N_D$) для получения искомой математической модели. В случае дробного факторного эксперимента некоторые эффекты – взаимодействия – будут коррелированы с другими эффектами, и задача становится, в общем случае, некорректно поставленной.

4. Описание сложных систем с использованием теоретико-аналитического подхода затруднено и для многих прикладных задач невозможно ввиду значительной сложности происходящих в системе комплексных (разнородных) явлений.

Академик А.А. Самарский считает, что «аналитические методы здесь уже, как правило, не работают. Решенные аналитически до конца серьезные нелинейные проблемы насчитываются единицами... Вычислительный эксперимент становится сейчас практически единственным средством проведения теоретических исследований в прикладных задачах» [7, с. 39].

Академик Я.З. Цыпкин также отмечает пригодность аналитических методов для решения относительно простых задач, полученных благодаря далеко идущей идеализации, в результате которой решается не поставленная, а совсем иная задача [8, с. 23].

Академик В.С. Пугачев считает, что «для построения моделей сложных систем приходится прибегать к экспериментальному исследованию самих систем или входящих в них подсистем и строить соответствующие модели путем статистической обработки полученных данных» [9, с. 351].

В различных сложных системах многокритериальная компромиссная оптимизация проводится с использованием поисковых (экспериментальных) методов оптимизации. Получение математических моделей осуществляется путем проведения многофакторных экспериментов и аппроксимации их результатов с использованием метода наименьших квадратов. Предполагается, что полученная математическая модель будет соответствовать и другим системам данного класса. Для этого необходимо, чтобы выбранная для эксперимента система по своим физическим, размерно-геометрическим и другим характеристикам была однородна, репрезентативна и случайно выбрана относительно своего класса систем.

5. Использование экспериментального подхода в описании сложных систем приводит к индуктивному методу познания. В отличие от теоретико-аналитического подхода, где используются абстрагирование, идеализация описываемой системы и она становится «унифицированной» относительно своего класса систем, выбранным для проведения эксперимента образцу системы и комплексу условий экспериментирования свойственны только эти условия. Описание в экспериментальном подходе проводится в условиях «индивидуального», «единичного», «конкретного», не отрывая объект изучения от определенных рамок времени и места. Такое изучение получило название «идиографическое» и рассматривалось как методология статистического описания в начале прошлого века чл.-кор. РАН, заведующим кафедрой статистики Петербургского политехнического института проф. А.А. Чупровым [10, с. 39-51]. В противоположность идиографическому описанию номографическое описание объясняет основные, наиболее общие законы, упрощая действительность, обобщая, отвлекаясь от деталей и индивидуальных подробностей.

Номографический подход, порождая «вечные законы», не объясняет конкретные явления реальной действительности в технических, технологических, метрологических и др. сложных сис-

темах (процессах, объектах). Их конструирование, изготовление, эксплуатация требуют информации для принятия решения не только на физическом или химическом уровне, но, главное, на инженерном [1, с. 356–369].

Выборный для эксперимента экземпляр системы и комплекс условий его проведения должны быть эквивалентны системам и условиям эксплуатации класса систем, к которому относится выбранный экземпляр системы. Условия эквивалентности могут быть сформулированы в терминах бинарных отношений порядка отношения на множестве: рефлексивности, симметричности, транзитивности [11, с. 83–84].

Бинарное отношение – совокупность упорядоченных пар элементов данного множества; об элементах a и b , входящих в одну из заданных пар, говорят, что они «находятся в отношении R между собой» и обозначают это формулой aRb . Под a подразумевается система, с которой проводится экспериментальное исследование, под b – любая система класса, к которому относится исследуемая система. Рефлексивное отношение – такое бинарное отношение R в множестве A , что для всех $a \in A$ верно aRa . Под множеством A понимаются системы определенного класса, номинально одинаковые между собой. Симметричное отношение – такое бинарное отношение R в множестве A , что для любых $a, b \in A$ справедливо $aRb \Rightarrow bRa$. Транзитивное отношение – такое бинарное отношение R в множестве A , что для любых $a, b, c \in A$ справедливо $aRb \wedge bRc \Rightarrow aRc$.

Бинарное отношение, которое рефлексивно, симметрично и транзитивно, будет отношением эквивалентности. В технических и технологических экспериментальных исследованиях все системы определенного класса в номинальном смысле эквивалентны. Фактически они различаются в пределах, допускаемых техническими условиями на параметры системы. Практика проведения технических, технологических и метрологических исследований с системами, которые находятся в эксплуатации, показывает, что воспроизводимость результатов экспериментов может нарушаться. Полученные в исследованиях результаты могут существенно отличаться от результатов, полученных в номинально одинаковых условиях с другими экземплярами этих же систем.

В рассматриваемом случае необходимо уменьшить неоднородность характеристик используемой в экспериментах системы, а также неоднородности условий проведения экспериментов и применять методы устойчивого оценивания математических моделей [1, с. 14–17]. Примеры решения многофакторных и многокритериальных экспериментальных задач приведены в [12, 13].

4. Выводы

1. Определение эффектов управляемых факторов по экспериментальным данным, представляющих собой результаты суммарного влияния управляемых, неуправляемых и неконтролируемых факторов, в общем случае является решением некорректно поставленной задачи. Многофакторный эксперимент необходимо планировать для исключения неопределенностей в исходной постановке задачи и, следовательно, для решения корректно поставленной задачи.
2. Случайный характер получаемых результатов в экспериментальных исследованиях требует использования методов устойчивого оценивания: устойчивый план эксперимента, структура модели, оценивание коэффициентов модели.
3. Используемая система предпосылок и начальных условий должна быть устойчива к сравнительно малым нарушениям принятых условий. Если принятые предпосылки фактически не выполняют

ся, то исследователь должен изменить метод обработки или модернизировать его с тем, чтобы получаемое решение было устойчивым. Предложенная формализованная структура многофакторной полиномиальной статистической модели, линейной относительно параметров в виде множества эффектов схемы полного факторного эксперимента, необходима и достаточна для адекватной аппроксимации результатов экспериментов при получении многофакторных уравнений регрессии. При этом предполагается, что используется устойчивый (робастный) план эксперимента.

4. Используемый в экспериментальных исследованиях индуктивный метод относится к эмпирическим обобщениям, или к эмпирическим законам. Обобщение проводится на основании результатов проведенных экспериментов, которые всегда ограничены по используемым условиям и числу опытов. Такие обобщения относятся к неполной индукции, и она не является логически обоснованным рассуждением. Полученные экспериментальные результаты следует рассматривать как вероятные гипотезы, которые могут проверяться и подтверждаться или опровергаться.

5. Комплекс условий – параметров проведения эксперимента для экспериментальной установки, используемых экспериментальных образцов, применяемых инструментов, средств измерений и других компонентов определяется группами управляемых, неуправляемых и неконтролируемых факторов. Необходимо проводить эксперимент только в относительно контролируемых и регулируемых условиях. Фактические значения параметров могут изменяться в пределах, установленных техническими условиями на проведение исследований. Обобщенной характеристикой постоянства комплекса условий проведения эксперимента является дисперсия воспроизводимости результатов эксперимента или его среднеквадратическая ошибка.

СПИСОК ЛИТЕРАТУРЫ

1. Радченко С.Г. Устойчивые методы оценивания статистических моделей: Монография. – К.: ПП «Санспарель», 2005. – 504 с.
2. Неуймин Я.Г. Модели в науке и технике. История, теория, практика / Под ред. Н.С. Соломенко. – Л.: Наука, Ленинград. отд., 1984. – 190 с.
3. Налимов В.В. Теория эксперимента. – М.: Наука, 1971. – 208 с.
4. Бриллюэн Л. Научная неопределенность и информация / Пер. с англ. Т.А. Кузнецовой / Под ред. и с послесл. И.В. Кузнецова. – М.: Мир, 1966. – 272 с.
5. Грекова И. Методологические особенности прикладной математики на современном этапе ее развития // Вопросы философии. – 1976. – № 6. – С. 104–114.
6. Швырков В.В. Тайна традиционной статистики Запада. – М.: Финансы и статистика, 1998. – 144 с.
7. Попов Ю.П., Самарский А.А. Вычислительный эксперимент // Компьютеры, модели, вычислительный эксперимент: Введение в информатику с позиций математического моделирования / Авт. пред. А.А. Самарского. – М., 1988. – С. 16–78.
8. Цыпкин Я.З. Адаптация и обучение в автоматических системах. – М.: Наука, 1968. – 399 с.
9. Пугачев В.С. Теория вероятностей и математическая статистика: Учеб. пособие. – 2-е изд., испр. и доп. – М.: Физматлит, 2002. – 496 с.
10. Чупров А.А. Очерки по теории статистики. – М.: Госстатиздат, 1959. – 320 с.
11. Микиша А.М., Орлов В.Б. Толковый математический словарь: Основные термины: около 2500 терминов. – М.: Рус. яз., 1989. – 244 с.
12. Радченко С.Г. Багатофакторне математичне моделювання та компромісна оптимізація технологічного процесу електроерозійного прощиття отворів // Математичні машини і системи. – 2003. – № 3, 4. – С. 186–200.
13. Радченко С.Г. Анализ экспериментальных данных на основе использования многофакторных статистических математических моделей // Математичні машини і системи. – 2005. – № 3. – С. 102–115.

