# Практическая работа № 5. Эксперимент в анализе систем

## Вопросы

### 1. Охарактеризуйте основные цели экспериментов при получении новых знаний.

Современное понимание эксперимента предполагает, что: 1) измерения кроме количественного могут носить качественный характер; 2) измерение может не снимать неопределенность, если она имеет расплывчатую природу; 3) измерение обычно сопровождается неизбежными погрешностями; 4) интересующая нас величина часто не наблюдаема и поддается лишь косвенным измерениям.

### 2. Что такое измерительные шкалы?

Шкала (измерительная шкала) — это знаковая система, для которой задано отображение (операция измерения), ставящее в соответствие реальным объектам, ситуациям, событиям или процессам тот или иной элемент (значение) шкалы.

### 3. Для чего используются шкалы наименования?

Предположим, что число различимых состояний (математический термин — число классов эквивалентности) конечно. Каждому классу эквивалентности поставим в соответствие обозначение, отличное от обозначений других классов. Теперь измерение будет состоять в том, чтобы, проведя эксперимент над объектом, определить принадлежность результата к тому или иному классу эквивалентности и записать это с помощью символа, обозначающего данный класс. Такое измерение называется измерением в шкале наименований (иногда эту шкалу называют также номинальной или классификационной); указанное множество символов и образует шкалу.

### 4. Охарактеризуйте основные свойства шкал наименования.

Естественнее всего использовать шкалу наименований в тех случаях, когда классифицируются дискретные по своей природе явления (например, различные объекты). Для обозначения классов могут применяться слова естественного языка (географические названия, собственные имена людей и т. д.), произвольные символы (гербы и флаги государств, эмблемы родов войск, всевозможные значки и т. д.), номера (регистрационные номера автомобилей, официальных документов, номера на майках спортсменов), а также их различные комбинации (например, почтовые адреса, экслибрисы личных библиотек, печати и пр.) Все эти обозначения эквивалентны простой нумерации (в некоторых странах человек при рождении получает номер, под которым он фигурирует в государственных информационных системах всю жизнь), но на практике часто предпочтение отдают другим обозначениям (вообразите, что вместо имен и фамилий ваших друзей и знакомых вы должны будете использовать номера).

### 5. Для чего используются порядковые шкалы?

В тех случаях, когда наблюдаемый (измеряемый) признак состояния имеет природу, позволяющую не только отождествлять состояния с одним из классов эквивалентности, но и в каком-то отношении сравнивать разные классы, для измерений можно выбрать более сильную шкалу, чем номинальная. Если же не воспользоваться этим, то мы откажемся от части полезной информации. Однако усиление измерительной шкалы зависит от того, какие именно отношения между классами существуют в действительности. Это и явилось причиной появления измерительных шкал разной силы.

### 6. Охарактеризуйте основные свойства порядковых шкал.

Характерной особенностью порядковых (в строгом смысле) шкал является то, что отношение порядка ничего не говорит о дистанции между сравниваемыми классами. Поэтому порядковые экспериментальные данные, даже если они изображены цифрами, нельзя рассматривать как числа, над ними нельзя выполнять действия, которые приводят к получению разных результатов при преобразовании шкалы, не нарушающем порядка.

При измерениях в порядковых (в строгом смысле) шкалах обработка данных должна основываться только на допустимых для этих шкал операциях – вычислении δij и Ri. С этими числами можно «работать» дальше уже произвольным образом: кроме нахождения частот и мод (как и для порядковой шкалы) появляется возможность определить выборочную медиану (т. е. наблюдение с рангом Ri, ближайшим к числу n/2); можно разбить всю выборку на части в любой пропорции, находя выборочные квантили любого уровня р, 0 < р < 1 (т. е. наблюдения с рангом Ri, ближайшим к величине nр); можно определить коэффициенты ранговой корреляции между двумя сериями порядковых наблюдений; строить с помощью полученных величин другие статистические процедуры.

### 7. Для чего используются модифицированные порядковые шкалы?

По-видимому, опыт работы с сильными числовыми шкалами и желание уменьшить относительность порядковых шкал, придать им хотя бы внешнюю независимость от измеряемых величин побуждают исследователей к различным модификациям, придающим порядковым шкалам некоторое (чаще всего кажущееся) усиление. Другая важная причина попыток усиления шкалы состоит в том, что многие измеряемые в порядковых (принципиально дискретных) шкалах величины имеют действительный или мыслимый непрерывный характер: сила ветра или землетрясения, твердость вещества, глубина и прочность знаний, овладение навыками и т. п. Сама возможность введения между любыми двумя шкальными значениями третьего способствует тому, чтобы попытаться усилить шкалу.

### 8. Охарактеризуйте основные свойства модифицированных порядковых шкал.

Опыт работы с сильными числовыми шкалами и желание уменьшить относительность порядковых шкал, придать им хотя бы внешнюю независимость от измеряемых величин побуждают исследователей к различным модификациям, придающим порядковым шкалам некоторое (чаще всего кажущееся) усиление. Другая важная причина попыток усиления шкалы состоит в том, что многие измеряемые в порядковых (принципиально дискретных) шкалах величины имеют действительный или мыслимый непрерывный характер: сила ветра или землетрясения, твердость вещества, глубина и прочность знаний, овладение навыками и т. п. Сама возможность введения между любыми двумя шкальными значениями третьего способствует тому, чтобы попытаться усилить шкалу.

### 9. Для чего используются шкалы интервалов?

Если упорядочивание объектов можно выполнить настолько точно, что известны расстояния между любыми двумя из них, то измерение окажется заметно сильнее, чем в шкале порядка. Естественно выражать все расстояния в единицах, хотя и произвольных, но одинаковых по всей длине шкалы. Это означает, что объективно равные интервалы измеряются одинаковыми по длине отрезками шкалы, где бы они на ней ни располагались. Следствием такой равномерности шкал этого класса является независимость отношения двух интервалов от того, в какой из шкал эти интервалы измерены (т. е. какова единица длины интервала и какое значение принято за начало отсчета). В самом деле, если два интервала в одной шкале выражаются числами Δ1x и Δ2x, а при другом выборе нуля и единицы — числами Δ1y и Δ2y, то, поскольку это объективно те же самые интервалы, имеем Δ1x/Δ2x = Δ1y/Δ2y; откуда следует, что введенные шкалы могут иметь произвольные начала отсчета и единицы длины, а связь между показаниями в таких шкалах является линейной: у = ах + b, а > 0, –∞ < b < ∞. Это соотношение можно выразить словами: «шкала интервалов единственна с точностью до линейных преобразований». Построенные таким образом шкалы называются интервальными.

### 10. Охарактеризуйте основные свойства шкал интервалов.

Название «шкала интервалов» подчеркивает, что в этой шкале только интервалы имеют смысл настоящих чисел и только над интервалами следует выполнять арифметические операции: если произвести арифметические операции над самими отсчетами по шкале, забыв об их относительности, то имеется риск получить бессмысленные результаты. Например, если сказать, что температура воды увеличилась в два раза при ее нагреве от 9 до 18° по шкале Цельсия, то для тех, кто привык пользоваться шкалой Фаренгейта, это будет звучать весьма странно, так как в этой шкале температура воды в том же опыте изменится от 37 до 42°.

### 11. Для чего используются шкалы отношений?

Пусть наблюдаемые величины удовлетворяют аксиомам аддитивности. Это существенное усиление шкалы: измерения в такой шкале являются «полноправными» числами, с ними можно выполнять любые арифметические действия, так как вычитание, умножение и деление — лишь частные случаи сложения. Введенная таким образом шкала называется шкалой отношений.

### 12. Охарактеризуйте основные свойства шкал отношений.

Этот класс шкал обладает следующей особенностью: отношение двух наблюдаемых значений измеряемой величины не зависит от того, в какой из таких шкал произведены измерения: x1/x2 = y1/y2. Этому требованию удовлетворяет соотношение вида у = aх (а ≠ 0). Таким образом, величины, измеряемые в шкале отношений, имеют естественный, абсолютный нуль, хотя остается свобода в выборе единиц.

### 13. Для чего используются шкалы разностей?

Повторно применяя сдвиг к y(z = у + b = х + 2b), затем к z и т. д., обнаруживаем, что в такой шкале значение не изменяется при любом числе сдвигов: у = х + nb, n = 0, 1, 2, ... Постоянная b является характерным параметром шкалы и называется ее периодом. Полученную шкалу будем называть шкалой разностей (иногда ее также называют циклической или периодической). В таких шкалах измеряется направление из одной точки (шкала компаса, роза ветров и т. д.), время суток (циферблат часов), фаза колебаний (в градусах или радианах).

### 14. Охарактеризуйте основные свойства шкал разностей.

Циклические шкалы являются частным случаем интервальных шкал. Однако соглашение о хотя и произвольном, но едином для нас начале отсчета шкалы позволяет использовать показания в этой шкале как числа, применять к ним арифметические действия (до тех пор, пока кто-нибудь не забудет об условности нуля, например, при переходе на летнее время или обратно).

### 15. Для чего используется абсолютная шкала?

Рассмотрим такую шкалу, которая имеет и абсолютный нуль, и абсолютную единицу. Эта шкала не единственна с точностью до какого-либо преобразования, а просто уникальна. Именно такими качествами обладает числовая ось, которую естественно назвать абсолютной шкалой.

### 16. Охарактеризуйте основные свойства абсолютной шкалы.

Важной особенностью абсолютной шкалы по сравнению со всеми остальными является отвлеченность (безразмерность) и абсолютность ее единицы. Указанная особенность позволяет производить над показаниями абсолютной шкалы такие операции, которые недопустимы для показаний других шкал: употреблять эти показания в качестве показателя степени и аргумента логарифма. Числовая ось используется как измерительная шкала в явной форме при счете предметов, а как вспомогательное средство присутствует во всех остальных шкалах. Внутренние свойства числовой оси при всей кажущейся ее простоте оказываются чрезвычайно разнообразными, и теория чисел до сих пор не исчерпала их до конца. А некоторые безразмерные числовые отношения, обнаруживаемые в природе, вызывают восхищение и изумление.

### 17. Что такое расплывчатое описание ситуаций?

Расплывчатость — это такое свойство явлений, при котором не выполняется отношение эквивалентности: явление одновременно может принадлежать данному классу и не принадлежать ему. Неопределенность такого типа описывается с помощью функции принадлежности; значение этой функции выражает степень уверенности, с которой мы относим данный объект к указанному классу. Сам класс в итоге становится не определяемым однозначно и называется расплывчатым множеством.

### 18. Что такое вероятностное описание ситуаций?

Самая полная информация, которой можно располагать о случайном объекте, содержится в распределении вероятностей по возможным состояниям этого объекта. Важно, что само конкретное распределение уже есть закономерность: неоднозначная (но и не вполне произвольная) для того, какое именно состояние реализуется, и вполне однозначная для многих важных характеристик, выражаемых функционалами от распределения.

### 19. Охарактеризуйте основные особенности статистических измерений.

Статистический, вероятностный подход относится к неопределенности, описываемой распределениями вероятностей. На то, что методы статистики надо применять осторожно, что многие экспериментальные ситуации могут быть хотя и хаотическими, но не иметь вероятностного характера, обращали внимание многие исследователи. Еще один важный момент состоит в том, чтобы по возможности ослабить или хотя бы учесть влияние измерений на наблюдаемый объект. Особенно это существенно при социальных исследованиях, наблюдениях за людьми: сам факт осознания, что они являются объектом внимания, заметно меняет их поведение. Воздействие измерительного устройства на измеряемый объект должно также учитываться при физических и химических экспериментах.

### 20. Охарактеризуйте основные особенности классификационных моделей.

Классификационные модели являются основополагающими, первичными, исходными формами знания. Узнавание окружающих предметов — типичный пример классификационных процессов в мыслительной деятельности человека (и животных). И в науке познание начинается с соотнесения изучаемого объекта с другими, выявления сходства и различия между ними. Поэтому протокол наблюдений на классификационном уровне эксперимента содержит результаты измерения ряда признаков X для подмножества А объектов, выбранных из множества Г. Признак характеризует конкретное свойство объекта; поэтому и тогда такой протокол называют таблицей «объект — свойство». Способ обработки протокола зависит от цели обработки. Часто оказывается, что задача может быть сформулирована как определение по наблюдавшимся значениям признаков х = (x1, ..., xn) значений ненаблюдаемого («целевого») признака х0. Как правило, целевыми признаками являются те параметры модели, которые требуется уточнить по экспериментальным данным.

### 21. Охарактеризуйте основные особенности числовых моделей.

Числовые модели отличаются от классификационных тем, что: 1) целевые признаки х0 измеряются в числовых шкалах; 2) числа х0 представляют собой функционалы или функции признаковых переменных (которые не обязательно все являются числовыми); 3) в этих моделях гораздо чаще учитываются связи переменных во времени (в классификационных задачах время иногда даже называют «забытой» переменной). В связи с этим и протоколы наблюдений могут не обязательно относиться к множеству объектов: модель можно уточнять и по экспериментам с одним объектом в разные моменты времени. Отметим, что числовые модели могут задавать связь между переменными как в параметризованной форме (т. е. в виде функции с конечным числом параметров), так и в непараметризованной форме (в виде функционала).

### 22. Охарактеризуйте основные особенности протоколов наблюдений.

Большая размерность. Во многих исследованиях число объектов N и число признаков n велики, так что произведение n × N достигает нескольких десятичных порядков. Учет времени приводит к еще большему увеличению размерности блока данных.  
Разнотипность данных. Разные признаки могут измеряться в различных шкалах.  
Пропущенные значения. Незаполненная ячейка таблицы данных — не такой уж редкий случай, особенно если эксперимент производится не в лабораторных, а в естественных условиях. Исключить из таблицы строку и столбец, на пересечении которых находится пустая ячейка, — выход далеко не всегда приемлемый.  
Зашумленность. Довольно часто измерение, занесенное в протокол, на самом деле отличается от измеряемого значения на некоторую случайную величину. Статистические свойства этой добавочной помехи могут не зависеть от измеряемой величины, и тогда мы говорим об аддитивном шуме. В противном случае имеет место неаддитивная, или зависимая, помеха.  
Искажения, отклонения от предположений. Приступая к обработке протокола наблюдений, мы всегда исходим из определенных предположений о природе величин, занесенных в протокол. Любой способ обработки дает результаты ожидаемого качества только в том случае, если данные отвечают определенным предположениям. Далеко не всегда в ходе обработки данных обращают внимание на то, действительно ли данные отвечают предположениям, заложенным в алгоритм обработки.

## Терминологический словарь

Шкала (измерительная шкала) — это знаковая система, для которой задано отображение (операция измерения), ставящее в соответствие реальным объектам, ситуациям, событиям или процессам тот или иной элемент (значение) шкалы.

1. Эксперимент — процедура, выполняемая для поддержки, опровержения или подтверждения гипотезы или теории.
2. Неопределённость — отсутствие или недостаток определения или информации о чём-либо.
3. Неопределённость измерения — параметр, относящийся к результату измерения и характеризующий разброс значений, которые могли бы быть обоснованно приписаны измеряемой величине.
4. Косвенные измерения — измерения, при которых значение величины находится на основании известной зависимости между этой величиной и величинами, подвергаемыми прямым измерениям.
5. Система — множество элементов, находящихся в отношениях и связях друг с другом, которое образует определённую целостность, единство.
6. Измерение — совокупность действий для определения отношения одной (измеряемой) величины к другой однородной величине, принятой всеми участниками за единицу, хранящуюся в техническом средстве (средстве измерений).
7. Модель — система, исследование которой служит средством для получения информации о другой системе; представление некоторого реального процесса, устройства или концепции.
8. Статическое измерение — измерение физической величины, принимаемой в соответствии с конкретной измерительной задачей за неизменную на протяжении времени измерения.
9. Отношение эквивалентности — бинарное отношение между элементами данного множества, свойства которого сходны со свойствами отношения равенства.
10. Шкала наименований (номинальная, классификационная) — это шкала, которая используется для измерения значений качественных признаков.
11. Дискретность — свойство, противопоставляемое непрерывности, прерывистость.
12. Порядковая шкала (или ранговая) — это шкала, которая включает отношения тождества и порядка, объекты в данной шкале ранжированы.
13. Интервал — множество всех чисел, удовлетворяющих строгому неравенству a < x < b.
14. Абсолютная шкала (шкала отношений) — это интервальная шкала, в которой присутствует дополнительное свойство — естественное и однозначное присутствие нулевой точки.
15. Расплывчатость — это такое свойство явлений, при котором не выполняется отношение эквивалентности: явление одновременно может принадлежать данному классу и не принадлежать ему.
16. Вероятность — степень (относительная мера, количественная оценка) возможности наступления некоторого события.
17. Случайность — это результат маловероятного или непредсказуемого события; проявление неотъемлемого дополнения к законам необходимости.
18. Классификация — понятие в науке, обозначающее разновидность деления объёма понятия по определённому основанию (признаку, критерию), при котором объём родового понятия (класс, множество) делится на виды (подклассы, подмножества), а виды, в свою очередь делятся на подвиды и т. д.
19. Кластерный анализ — многомерная статистическая процедура, выполняющая сбор данных, содержащих информацию о выборке объектов, и затем упорядочивающая объекты в сравнительно однородные группы.