Федеральное агентство по образованию

Государственное образовательное учреждение
высшего профессионального образования

«Омский государственный технический университет»

МАТЕМАТИЧЕСКИЕ МЕТОДЫ В ПСИХОЛОГИИ

Конспект лекций
для студентов 2 курса гуманитарных специальностей
дневного, вечернего и заочного отделений

Омск – 2008

Составитель Ананко Алла Александровна, ст. преподаватель

Печатается по решению редакционно-издательского совета Омского государственного технического университета.

ОГЛАВЛЕНИЕ

Лекция 1. Измерения и шкалы
1.1. Типы измерений

Любое эмпирическое научное исследование начинается с того, что исследователь фиксирует выраженность интересующего его свойства, как правило, при помощи чисел. Таким образом, следует различать ***объекты исследования*** (в психологии это чаще всего люди, испытуемые), их ***свойства*** (то, что интересует исследователя, составляет предмет изучения) и ***признаки***, отражающие в числовой шкале выраженность свойств.

Измерение в терминах производимых исследователем операций - это приписывание объекту числа по определенному правилу. Это правило устанавливает соответствие между измеряемым свойством объекта и результатом измерения - признаком.

В обыденном сознании, как правило, нет необходимости разделять свойства вещей и их признаки: такие свойства предметов, как вес и длина, мы отождествляем, соответственно, с количеством граммов и сантиметров. Если нет необходимости в измерении, мы ограничиваемся сравнительными суждениями: этот человек тревожный, а этот - нет, этот более сообразителен, чем другой, и т.д.

В научном исследовании нам исключительно важно отдавать себе отчет в том, что точность, с которой признак отражает измеряемое свойство, зависит от процедуры измерения.

Пример. Мы можем разделить всех наших испытуемых на две группы по сообразительности: сообразительные и не очень. И далее приписать каждому испытуемому символ (например, 1 и 0) в зависимости от его принадлежности к той или другой группе мы можем упорядочить всех испытуемых по степени выраженности сообразительности, приписывая каждому его ранг, от самого сообразительного (1 ранг), самого сообразительного из оставшихся (2 ранг) и т. д. до последнего испытуемого. В каком из этих двух случаев измеренный признак будет точнее отражать различия между испытуемыми по измеряемому свойству, догадаться нетрудно.

В зависимости от того, какая операция лежит в основе измерения признака, выделяют так называемые измерительные шкалы. Они еще называются шкалами С. Стивенса, по имени ученого-психолога, который их предложил. Эти шкалы устанавливают определенные соотношения между свойствами чисел и измеряемым свойством объектов. Шкалы разделяют на метрические (если есть или может быть установлена единица измерения) и неметрические (если единицы измерения не могут быть установлены).

1.2. Измерительные шкалы

Номинативная шкала (неметрическая), или шкала наименований (номинальное измерение). В ее основе лежит процедура, обычно не ассоциируемая с измерением. Пользуясь определенным правилом, объекты группируются по различным классам так, чтобы внутри класса они были идентичны по измеряемому свойству. Каждому классу дается наименование и обозначение, обычно числовое. Затем каждому объекту присваивается соответствующее обозначение.

Примеры. Примеры номинативных признаков: «пол» (1-мужской, 0-женский), «национальность» (1 - русский, 2 - белорус, 3 - украинец), «предпочтение домашних животных» (1 - собаки, 2 - кошки, 3 - крысы, 0 - никакие) и т. д. В последнем случае если одному испытуемому присвоена 1, а другому 2, то это обозначает только то, что у них разные предпочтения: у первого - собаки, у второго - кошки. Из того, что 1<2, нельзя делать вывод, что у второго предпочтение выражено больше, чем у первого, и т. д.

Заметим, что в этом случае мы учитываем только одно свойство чисел - то, это разные символы. Привычные операции с числами - упорядочивание, сложение-вычитание, деление - при измерении в номинативной шкале теряют смысл. При сравнении объектов мы можем делать вывод только о том, принадлежат они к одному или разным классам, тождественны или нет по измеренному свойству. Несмотря на такие ограничения, номинативные шкалы широко используются в психологии, и к ним применимы специальные процедуры обработки и анализа данных.

Ранговая, или **порядковая шкала (неметрическая)** (как результат ранжирования). Как следует из названия, измерение в этой шкале предполагает приписывание объектам чисел в зависимости от степени выраженности измеряемого свойства.

Существует множество способов получения измерения в порядковой шкале. Но суть остается общей: при сравнении испытуемых друг с другом мы можем сказать, больше или меньше выражено свойство, но не можем сказать, насколько больше или насколько меньше оно выражено, а уж тем более — во сколько раз больше или меньше. При измерении в ранговой шкале, таким образом, из всех свойств чисел учитывается то, что они разные, и то, что одно число больше, чем другое.

Пример. Четверым бегунам присвоены ранги в соответствии с тем, кто раньше достиг «финиша» (ранг 1 - самый быстрый):

Бегун	Ранг
A	1
B	2
C	3
D	4

Основываясь только на этих данных, мы можем судить о том, кто раньше прибежал, а кто позже. Но мы не можем судить, насколько каждый из них пробежал быстрее или медленнее другого.

При ранжировании «вручную», а не при помощи компьютера, следует иметь в виду два обстоятельства:

1. Установите для себя и запомните порядок ранжирования. Вы можете ранжировать испытуемых по их «месту в группе»: ранг 1 присваивается тому, у которого наименьшая выраженность признака, и далее - увеличение ранга по мере увеличения уровня признака. Или можно ранг 1 присваивать тому, у которого 1-е место по выраженности данного признака (например, «самый быстрый»). Строгих правил выбора здесь нет, но важно помнить, в каком на правлении производилось ранжирование.

2. Соблюдайте правило ранжирования для связанных рангов, когда двое или более испытуемых имеют одинаковую выраженность измеряемого свойства. В этом случае таким испытуемым присваивается один и тот же, средний ранг. Например, если вы ранжируете испытуемых по «месту в группе» и двое имеют одинаковые самые высокие исходные оценки, то обоим присваивается средний ранг 1,5: (1+2)/2=1,5. Следующему за этой парой испытуемому присваивается ранг 3, и т.д. Это правило основано на соглашении соблюдения одинаковой суммы рангов для связанных и несвязанных рангов. В соответствии с этим правилом сумма всех присвоенных рангов для группы численностью N должна равняться N(N+1)/2, вне зависимости от наличия или отсутствия связей в рангах.

Интервальная шкала (метрическая). Это такое измерение, при котором числа отражают не только различия между объектами в уровне выраженности свойства (характеристика порядковой шкалы), но и то, насколько больше или меньше выражено свойство. Равным разностям между числами в этой шкале соответствуют равные разности в уровне выраженности измеренного свойства. Иначе говоря, измерение в этой шкале предполагает возможность применения единицы измерения (метрики). Объекту присваивается число единиц измерения, пропорциональное выраженности измеряемого свойства. Важная особенность интервальной шкалы - произвольность выбора нулевой точки: ноль вовсе не соответствует полному отсутствию измеряемого свойства. Произвольность выбора нулевой точки отсчета обозначает, что измерение в этой шкале не соответствует абсолютному количеству измеряемого свойства. Следовательно, применяя эту шкалу, мы можем судить, насколько больше или насколько меньше выражено свойство при сравнении объектов, но не можем судить о том, во сколько раз больше или меньше выражено свойство.

Пример. Наиболее типичный пример измерения в интервальной шкале температура по шкале Цельсия (°С). Важная особенность такого измерения заключается в том, что нулевая точка на шкале не соответствует полному отсутствию измеряемого свойства (0°С - это точка замерзания воды, но не отсутствия температуры, тепла). И если сегодня +5°С, а вчера было +10 °С, то можно сказать, что сегодня на 5 градусов холоднее, но неверно утверждать, что сегодня холоднее в два раза.

Интервальные измерения широко используются в психологии. Примером могут являться тестовые шкалы, которые специально вводятся при обосновании равноинтервальности (метричности) тестовой шкалы.

Абсолютная шкала, или шкала отношений (метрическая). Измерение в этой шкале отличается от интервального только тем, что в ней устанавливается нулевая точка, соответствующая полному отсутствию выраженности измеряемого свойства.

Пример. В силу абсолютности нулевой точки, при сравнении объектов мы можем сказать не только о том, насколько больше или меньше выражено свойство, но и о том, во сколько раз (на сколько процентов и т. д.) больше или меньше оно выражено. Измерив время решения задачи парой испытуемых, мы можем сказать не только о том, кто и на сколько секунд (минут) решил задачу быстрее, но и о том, во сколько раз (на сколько процентов) быстрее. Следует отметить, что, несмотря на привычность и обыденность абсолютной шкалы, в психологии она используется не часто.

Перечисленные шкалы полезно характеризовать еще и по признаку их дифференцирующей способности (мощности). В этом отношении шкалы по мере возрастания мощности располагаются следующим образом: номинативная, ранговая, интервальная, абсолютная. Таким образом, неметрические шкалы заведомо менее мощные - они отражают меньше информации о различии объектов (испытуемых) по измеренному свойству, и, напротив, метрические шкалы более мощные, они лучше дифференцируют испытуемых. Поэтому, если у исследователя есть возможность выбора, следует применить более мощную шкалу.

1.3. Как определить, в какой шкале измерено явление

Определение того, в какой шкале измерено явление (представлен признак), - ключевой момент анализа данных: любой последующий шаг, выбор любого метода зависит именно от этого.

Обычно идентификация номинативной шкалы, ее дифференциация от ранговой, а тем более от метрической шкалы, не вызывает особых проблем.

Рассмотрим вопрос анкеты, для ответа на который испытуемые выбирают один из предложенных вариантов:

«Насколько Вы уверены в своих силах...

1). Совершенно уверен

2). Затрудняюсь ответить

3). Совершенно неуверен»

Если исследователя интересует, в какой степени испытуемые уверены или не уверены в своих силах, то логично предполагать, что признак представлен в ранговой шкале. Если же исследователя интересует то, как распределились ответы по вариантам или чем характеризуется каждая из 3 соответствующих групп, то разумнее рассматривать этот признак как номинативный.

Значительно сложнее определить различие между порядковой и метрической шкалами. Проблема связана с тем, что измерения в психологии, как правило, косвенные. Непосредственно мы измеряем некоторые наблюдаемые явления или события: количество ответов на вопросы, или заданий, решенных за введенное время, или время решения набора заданий и т. д. Но при этом выносим суждения о некотором скрытом, латентном свойстве, недоступном прямому наблюдению: об агрессивности, общительности, способности и т. д.

Количество заданий, решенных за отведенное время, - это, конечно, измерение в метрической шкале. Но само по себе это количество нас интересует лишь в той мере, в какой оно отражает некоторую изучаемую нами способность. Соответствуют ли равные разности решенных задач равным разностям выраженности изучаемого свойства (способности)? Если ответ «да» - шкала метрическая (интервальная), если «нет» - шкала порядковая.

Конечно, проще всего в подобных ситуациях согласиться с тем, что признак представлен в порядковой шкале. Но при этом мы существенно ограничиваем себя в выборе методов последующего анализа. Более того, переход к менее мощной шкале обрекает нас на утрату части столь ценной для нас эмпирической информации об индивидуальных различиях испытуемых. Следствием этого может являться падение статистической достоверности результатов исследования. Поэтому исследователь стремится все же найти свидетельства того, что используемая шкала - более мощная, метрическая.

Лекция 2. Дискретный вариационный ряд и его основные показатели

2.1.Вариация признака в совокупности и значение её изучения

Основой обработки данных статистического наблюдения является построение рядов распределения. Цель его - выявление основных свойств и закономерностей исследуемой статистической совокупности. В зависимости от того, является ли признак, взятый за основу группировки, качественным или количественным, различают соответственно два типа рядов распределения - атрибутивные и вариационные. Ряды распределения, построенные по качественным признакам, называют **атрибутивными**. Примером атрибутивных рядов может служить распределение населения по полу, характеру труда, национальности, профессии и т.д. Ряды распределения, построенные по количественному признаку, называют **вариационными.** Величины того или иного количественного признака у отдельных единиц совокупности более или менее различаются между собой. Такое различие в величине признака носит название вариации.

Рассмотрим пример:

Таблица 1

Динамика цен 1 кв. м. муниципального жилья в Москве в 1995 г. (млн.руб.)

Местоположение	Май
Митино	5,2
Южное Бутово	4,9
Северное Бутово	5,8
Отрадное	6,5
Веерная улица	7,6
Жулебино	4,9

Числовые значения признака, встречающиеся в данной совокупности, называют вариантами значений. В данном примере встречается пять вариантов значений признака, которые колеблются в пределах от 4,9 до 7,6 млн. руб. за 1 кв. м.

Изучение характера и степени вариации признаков у отдельных единиц, составляющих изучаемую совокупность, является важнейшим вопросом всякого статистического исследования. Управление процессом развития в желаемом направлении требует определения роли не только каждой из выше названных групп факторов в вариации тех или иных признаков, но и роли отдельных факторов соответствующих групп. Для решения такой задачи в статистике применяются специальные методы исследования вариации, основанные на использовании системы показателей, с помощью которой измеряется вариация.

Представленные выше данные о цене 1 кв. м. муниципального жилья по шести районам г. Москвы без какой-либо систематизации образуют так называемый первичный ряд данных. При наличии достаточно большого количества вариантов значений признака первичный ряд становится трудно обозримым и непосредственное рассмотрение его не дает представления о распределении единиц по величине признака в совокупности. Первым шагом в упорядочении первичного ряда является его ранжирование, т.е. расположение всех вариантов ряда в возрастающем (или убывающем) порядке. В мае 1995 г. по районам Москвы мы располагали такой информацией о цене 1 кв. м. муниципального жилья:

Таблица 2

Район Москвы	Южное Бутово	Жулебино	Митино	Северное Бутотово	Отрадное	Веерная улица
Цена 1 кв. м. жилья, млн. руб.	4,9	4,9	5.2	5,8	6,5	7,6

Ранжированный ряд данных позволяет сразу увидеть наименьшее и наибольшее значение признака в совокупности, определить расстояние между крайними значениями признака, а также выделить наиболее часто повторяющиеся значения в обследуемой совокупности. Использование ранжированного ряда также позволяет легко разделить все данные по группам. В нашем примере цена 1 кв. м. жилья варьирует по районам г. Москвы от 4,9 млн. до 7,6 млн. руб., т.е. размах вариации составляет 2,7 млн. руб. за 1 кв. м, можно также видеть, что в двух районах встречается одна и та же цена - 4,9 млн. руб.

По характеру вариации различают дискретные и непрерывные признаки. **Дискретные признаки** отличаются друг от друга на некоторую конечную величину, т.е. даны в виде прерывных чисел. Например, тарифный разряд рабочих, число детей в семье, число рабочих на предприятии и т.д. **Непрерывные признаки** могут отличаться один от другого на сколь угодно малую величину и в определенных границах принимать любые значения. Например, заработная плата рабочих, размер среднедушевого денежного дохода, стоимость основных фондов предприятия и т.д.

Способы построения вариационного ряда для этих видов признаков различны. Для построения дискретного ряда с небольшим числом вариантов достаточно перечислить все встречающиеся варианты значений признака, обозначаемые через x_i, а затем подсчитать частоту повторения каждого варианта f_i (например, распределение рабочих по разрядам, студентов по успеваемости и т.п.).

В тех случаях, когда число вариантов дискретного признака достаточно велико, а также при анализе вариации непрерывного признака, когда значения признака у отдельных единиц могут вообще не повторяться, строятся интервальные ряды распределения. Интервал указывает определенные пределы значений варьирующего признака и обозначается нижней и верхней границами интервала. Такие распределения наиболее распространены в практике статистической работы.

При построении интервальных рядов распределения необходимо прежде всего установить число групп (интервалов), на которые следует разбить все единицы изучаемой совокупности. При группировке внутри однородных совокупностей появляется возможность применения равных интервалов, величина которых зависит от вариации признака в совокупности и от количества обследованных единиц.

Определение величины интервала h для построения вариационного ряда с равными интервалами производится следующим образом:

1) вычисляется разность между максимальным и минимальным значениями признака первичного ряда (определяется размах вариации, R): $R = x_{max} - x_{min}$;

2) размах вариации делится на число групп k, т.е. $h = \frac{R}{k}$.

Число групп приближенно определяется по формуле Стэрджесса: $k = 1 + 3{,}322 \ln n$, где n – общее число изучаемых единиц совокупности.

Указанное выражение почти всегда оказывается дробной величиной, которую округляют до целого числа, поскольку количество групп не может быть дробным.

Для целей анализа и сравнительной характеристики различных рядов распределения применяются обобщающие показатели вариационного ряда. Система таких показателей может быть наглядно представлена при сравнении особенностей нескольких рядов распределения.

1. Выборочное среднее $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$,

где x_i – измеренная величина, n – количество измеренных значений.

Выборочное среднее характеризует среднее значение экспериментального показателя в выборке наблюдений. Этот показатель очень часто используется при сравнении различных выборок наблюдений. Пусть в результате эксперимента получена выборка значений: 3, 2, 1, 0, 3, 2, 3. Рассчитаем ее выборочное среднее

$$\overline{x} = \frac{3+2+1+0+3+2+3}{7} = \frac{14}{7} = 2.$$

2. Выборочная дисперсия $\overline{s}^2 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n-1}$.

Выборочная дисперсия характеризует меру рассеяния случайной величины относительно математического ожидания (выборочного среднего значения). Выборочная дисперсия в условиях предыдущего примера рассчитывается как

$$\overline{s}^2 = \frac{(3-2)^2 + (2-2)^2 + (1-2)^2 + (0-2)^2 + (3-2)^2 + (2-2)^2 + (3-2)^2}{6} = \frac{8}{6} = 1{,}33.$$

3. Выборочная медиана

Выборочная медиана представляет собой то значение в вариационном ряду, которое делит его пополам. Например, в вариационном ряду всего 21 измерение, поэтому значение 11-го измерения и будет значением медианы. Этот показатель очень часто используется в психологических исследованиях.

4. Асимметрия $A = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^3}{n\sigma^3}$.

Если $A < 0$, то эмпирическое распределение несимметрично и сдвинуто вправо. При $A > 0$ распределение имеет сдвиг влево. При $A = 0$ распределение симметрично.

5. Эксцесс Е. Показатель, характеризующий выпуклость или вогнутость эмпирических распределений:

$$E = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^4}{n\sigma^4} - 3.$$

Если E больше или равно нулю, распределение выпукло, в других случаях - вогнуто.

6. Мода M_0, наибольшая частота, в примере $M_0 = 21$.

7. Квантиль

Это значение вариационного ряда, которое соответствует заданному значению вероятности появления признака. При $p = 0{,}25$ это значение называется нижним квартилем, при $p = 0{,}75$ – верхним квартилем. Например, в выборке из 13 значений нижний квартиль равен 13, верхний – 25, а мода 21.

Пример

1	2	3	4	5	6	7	8	9	10	11	12	13
12	13	13	15	16	21	21	21	22	24	25	26	28

Нижний квартиль Мода Верхний квартиль

8. Размах выборки L. $L = x_{max} - x_{min}$.

В примере $x_{max} = 28$; $x_{min} = 12$.
$L = 28 - 12 = 16$.

Лекция 3. Статистический анализ выборочных средних двух выборок

3.1. Выбор метода и общий подход

Очень часто в статистическом анализе необходимо сравнить значения выборочных средних двух выборок измерений. Решение этих задач проводится в тех случаях, когда требуется оценить различия в деятельности двух групп и степень влияния корригирующих воздействий (для психологов), наличие отклонений в динамике изменения показателей (для юристов и экономистов) и т. д. Выбор методов доказательств этих предположений зависит от следующих особенностей выборок случайных величин:

1. Зависимы или независимы между собой исследуемые группы.
2. Подчиняются ли они закону нормального распределения случайных величин.
3. Используются ли в анализе группирующая переменная признаков (пол, факультет, курс, возраст, уровень доходов и другие признаки).

Если две сравниваемые выборки измерений независимы (independed) между собой и подчинены нормальному закону распределения случайных величин, то решение задачи проводится с использованием параметрического критерия ***(t-критерия Стьюдента).*** Анализ вида закона распределения проводится по эмпирическому распределению в виде гистограммы. В том случае, если при выполнении контрольной работы используется пакет Statistika, определение вида закона проводится с использованием статистических критериев Шапиро-Уилкоксена и Колмогорова-Смирнова. Для зависимых выборок измерений (depended) также используется Statistika на основе $t -$***критерия.***

Для непараметрических распределений решение задачи сравнения выборок проводится с использованием знакового критерия **(Siqntest)** для зависимых переменных и критерия ***Манна-Уитни (Mann-Whitney)*** для независимых переменных.

Схема статистического анализа при проверке гипотез о равенстве средних значений в двух выборках для всех возможных случаев приведена на рисунке 1. Зависимыми называются группы, между которыми существует внутренняя связь. Например: родные братья и сестры (генетическая связь), одна и та же группа, исследуемая дважды. Зависимость между двумя группами может быть и по преподавателю, который проводит занятия в двух группах. Во всех случаях зависимость определяется исследователем.

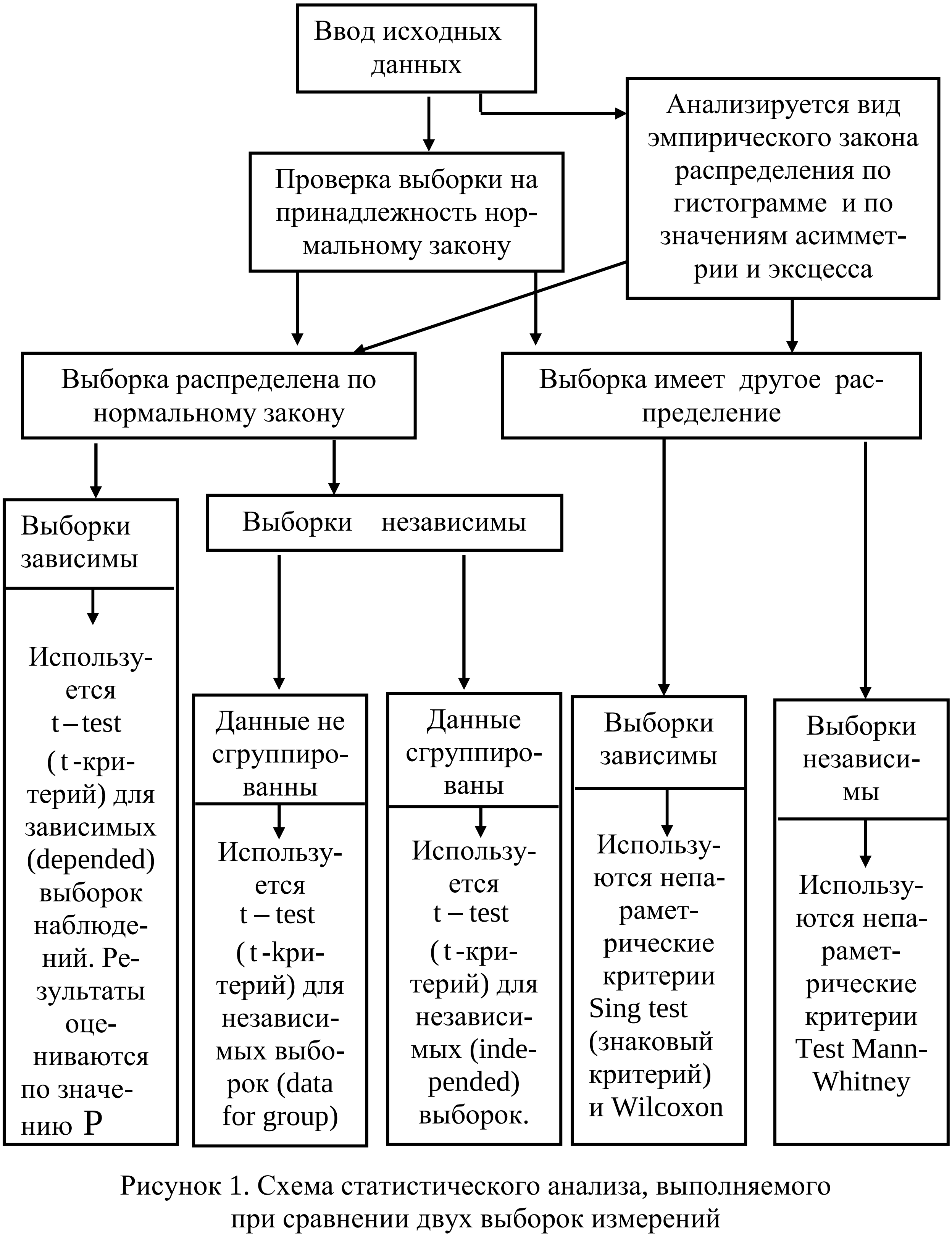

Рисунок 1. Схема статистического анализа, выполняемого при сравнении двух выборок измерений

Процедура проверки статистических гипотез заключается в следующем.

1. Выбирается некоторый критерий проверки гипотезы (рис. 1).
2. Вычисляется по определенному алгоритму значение критической об-

ласти критерия из исходной выборки наблюдений Q_p.

3. По исходным данным в специальных таблицах (приложение А) определяются критические значения области критерия Q_k при заданном уровне значимости 0,05 и 0,01.

4. Сравнивают расчетное и критическое значения критерия. Если расчетное значение больше или равно критическому значению, гипотезу H_0 отвергают. Исключение составляют только критерии знаков G, критерий Уилкоксена T и

критерий Манна-Уитни U. Для них устанавливаются обратные соотношения.

Могут быть и другие правила статистической проверки гипотез, в которых используются два критических значения при уровнях значимости 0,05 и 0,01. В том случае, если расчетное значение больше или равно критическому значению при уровне значимости 0,05, гипотеза H_0 отклоняется, а гипотеза H_1 еще не принимается. Гипотеза H_1 принимается только в том случае, если расчетное значение больше или равно критическому на уровне значимости 0,01.

3.2. t-критерий Стьюдента

Назначение критерия. Критерий предназначен для оценки различий между ***двумя*** параметрическими эмпирическими распределениями по среднему значению ***(уровню)*** какого-либо признака, количественно измеренного. Он позволяет выявлять различия между выборками, когда $n > 3$.

Описание критерия. Эмпирическое расчетное значение критерия t_p отражает, насколько велика зона совпадения между рядами. Чем ***больше*** t_p, ***тем более*** вероятно, что различия ***достоверны.***

Гипотезы

H_0: Уровень признака в группе 2 не ниже уровня признака в группе 1.

H_1: Уровень признака в группе 2 ниже уровня признака в группе 1.

Алгоритм расчета t-критерия Стьюдента для независимых выборок измерений

1. Определить расчетное значение t-критерия по формуле

$$t_p = \frac{\left|\bar{x}_1 - \bar{x}_2\right|}{\sqrt{\dfrac{(n_1-1)\ \bar{s}_1^{\,2} + (n_2-1)\ \bar{s}_2^{\,2}}{f} \cdot \dfrac{n_1+n_2}{n_1 n_2}}}, \qquad (1)$$

где $f -$ степень свободы, которая определяется как $f = n_1 + n_2 - 2$.

2. Определить критическое значение t-критерия при заданном уровне значимости и степени свободы.

3. Сравнить расчетное и критическое значения t-критерия. Если расчетное значение больше или равно критическому, то гипотеза (H_0) равенства средних значений в двух выборках измерений отвергается. Во всех других случаях она принимается на заданном уровне значимости.

Пример. Две группы студентов обучались по двум различным методикам. В конце обучения с ними был проведен тест по всему курсу. Необходимо оценить, насколько существенны различия в полученных знаниях. Результаты тестирования представлены в таблице 3.

Таблица 3

25	18	9	13	8	20	25	18	6	12
19	13	12	12	18	9	7	10	18	20

Рассчитаем выборочное среднее, дисперсию и стандартное отклонение:

$x_1 = 15,4;$ $x_2 = 13,8;$ $s_1 = 6,83;$ $s_2 = 4,61;$ $f = 18.$

Определим по формуле (1) значение

$$t_p = 0,62.$$

По таблице находим критическое значение t_k для уровня значимости $\alpha = 0,01.$

$$t_k = 2,88.$$

Вывод*: так как расчетное значение критерия меньше критического* $(0,62 < 2,88)$, *гипотеза* H_0 *подтверждается и существенных различий в методиках обучения нет на уровне значимости 0,01.*

3.4. Алгоритм расчета t-критерия Стьюдента для зависимых выборок измерений

1. Определить расчетное значение t-критерия по формуле

$$t = \frac{\left|\bar{d}\right| \sqrt{n}}{S_d},$$

где $$\bar{d} = \frac{\sum_{i=1}^{n}\left(x_{i1} - x_{i2}\right)}{n}, \qquad S_d = \sqrt{\frac{\sum_{i=1}^{n}\left(d_i - \bar{d}\right)^2}{n-1}}.$$

6. Рассчитать степень свободы $f = n - 2.$

7. Определить критическое значение t-критерия (см. прил., табл. А3).

8. Сравнить расчетное и критическое значение t-критерия. Если расчетное значение больше или равно критическому, то гипотеза Но равенства средних значений в двух выборках изменений отвергается. Во всех других случаях она принимается на заданном уровне значимости.

Лекция 4. Критерии для непараметрических распределений

4.1. U-критерий Манна-Уитни

Назначение критерия. Критерий предназначен для оценки различии между ***двумя*** непараметрическими выборками по ***уровню*** какого-либо признака, количественно измеренного. Он позволяет выявлять различия между ***малыми*** выборками, когда $n < 30$.

Описание критерия

Этот метод определяет, достаточно ли мала зона пересекающихся значений между двумя рядами. Чем меньше эта область, тем более вероятно, что различия достоверны. Эмпирическое значение критерия и отражает то, насколько велика зона совпадения между рядами. Поэтому, ***чем меньше*** U, ***тем более*** вероятно, что различия ***достоверны.***

Гипотезы

Ho: Уровень признака в группе 2 не ниже уровня признака в группе 1.

$H1$: Уровень признака в группе 2 ниже уровня признака в группе 1.

Алгоритм расчета критерия Манна-Уитни U

1. Перенести все данные испытуемых на индивидуальные карточки.
2. Пометить карточки испытуемых выборки 1 одним цветом, скажем, красным, а все карточки из выборки 2 – другим, например синим.
3. Разложить все карточки в единый ряд по степеням нарастания признака, не считаясь с тем, к какой выборке они относятся, как если бы была одна большая выборка.
4. Проранжировать значения на карточках, приписывая меньшему значению меньший ранг.
5. Вновь разложить карточки на две группы, ориентируясь на цветные обозначения: красные карточки в один ряд, синие – в другой.
6. Подсчитать сумму рангов отдельно на красных карточках (выборка 1) и на синих карточках (выборка 2). Проверить, совпадает ли сумма рангов с расчетной.
7. Определить большую из двух ранговых сумм.
8. Определить по формуле значение

$$U = (n_1 \cdot n_2) + \frac{n_x \cdot (n_x + 1)}{2} - T_x ,$$

где n_1 – количество испытуемых в выборке 1; n_2 – количество испытуемых в выборке 2; T_x – большая из двух ранговых сумм; n_x – количество испытуемых в группе с большей суммой рангов.

9. Определить критические значения U. Если $U_{эмп} > U_{кр\ 0,05}$, то гипотеза H_0 принимается. Если $U_{эмп} \le U_{кр}$, то отвергается. Чем меньше значения U, тем достоверность различий выше.

Пример. Сравнить эффективность двух методов обучения в двух группах. Результаты испытаний представлены в таблице 4.

Таблица 4

18	10	7	14	11	13					
15	20	10	8	16	10	19	7	15	14	29

Перенесем все данные в другую таблицу, выделив данные второй группы, подчеркиваем и делаем ранжирование общей выборки (см. алгоритм ранжирования в методических указаниях к заданию).

Значения	7	$\underline{7}$	$\underline{8}$	10	$\underline{10}$	$\underline{10}$	11	13	$\underline{14}$	14	$\underline{15}$	$\underline{15}$	$\underline{16}$	18	$\underline{19}$	$\underline{20}$	$\underline{29}$
Ранги	1,5	1,5	3	5	5	5	7	8	9,5	9,5	11,5	11,5	13	14	15	16	17
Номер	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

Найдем сумму рангов двух выборок и выберем большую из них: $T_x = 108$.

Рассчитаем эмпирическое значение критерия по формуле (3): $U_p = 24$.

Определим критическое значение критерия при уровне значимости $\alpha = 0{,}05 : U_k = 19$ (см. прил. табл. А1)

Вывод: *так как расчетное значение критерия* U *больше критического при уровне значимости* $\alpha = 0{,}05$ *и* $24 > 19$, *гипотеза о равенстве средних принимается, различия в методиках обучения будут несущественны.*

4.2. Критерий знаков G

Назначение критерия. Предназначен для оценки различий между ***двумя зависимыми*** непараметрическими выборками по ***уровню*** какого-либо признака, количественно измеренного. Он позволяет выявлять различия между малыми выборками.

Описание критерия. Эмпирическое значение критерия G отражает то, насколько велика зона совпадения между зависимыми выборками. ***Чем меньше*** G, ***тем более*** вероятно, что различия ***достоверны.***

Гипотезы

H_0: Уровень признака в группе 2 не ниже уровня признака в группе 1.

H_0: Уровень признака в группе 2 ниже уровня признака в группе 1.

Алгоритм расчета критерия знаков G

1. Подсчитать количество нулевых реакций и исключить их из рассмотрения. В результате n уменьшится на количество нулевых реакций.

2. Определить преобладающее направление изменений. Считать сдвиги в преобладающем направлении «типичными».

3. Определить количество «нетипичных» сдвигов. Считать это число эмпирическим значением G.

4. Определить критические значения G для данного n.

5. Сопоставить расчетное и критическое значения критерия G. Если расчетное значение критерия больше критического, то различия между выборками на заданном уровне значимости отсутствуют, во всех других случаях сдвиг в типичную сторону может считаться достоверным.

Пример. В группе изучались два способа решения проблем. Чтобы сравнить, какой из них совершенней, были даны две серии заданий. Определить, существуют ли преимущества у какого-либо способа в решении проблем. Результаты выполненных заданий двумя способами приведены в таблице 3.

Таблица 5

Испытуемые	1	2	3	4	5	6	7	8	9	10	11
1-й способ	5	4	10	6	7	7	3	9	7	9	6
2-й способ	7	4	8	7	5	6	5	7	4	9	4

Определим направления изменения знаков, для чего вычтем значения второй выборки из первой

1-й способ	5	4	10	6	7	7	3	9	7	9	6
2-й способ	7	4	8	7	5	6	5	7	4	9	4
Знак	–	=	+	–	+	+	–	+	+	=	+

Так как два значения были определены со знаком равенства (нулевые), рассчитаем новое значение n; $n = 11 - 2 = 9$.

Типичный сдвиг в сторону положительных значений (+) и таких случаев 6, а нетипичных сдвигов (–) всего 3. Это и будет значение G_p. $G_p = 3$.

$G_t = 1$ при уровне значимости $\alpha = 0,05$ (см. прил. табл. А1)

Вывод: *так как расчетное значение критерия* $G_p > G_t$, *гипотеза* Но *принимается и существенных различий в способах решения проблем нет.*

Лекция 5. Вычисление и анализ коэффициента ранговой корреляции
5.1. Выполнить ранжирование по следующему алгоритму

1. Меньшему значению начисляется меньший ранг. Наименьшему значению начисляется ранг 1. Наибольшему значению начисляется ранг, соответствующий количеству ранжируемых значений. Например, если $n = 7$, то наибольшее значение получит ранг 7, за возможным исключением тех случаев, которые предусмотрены правилом 2.

2. Если несколько значений равны, им начисляется ранг, представляющий собой среднее значение из тех рангов, которые они получили бы, если бы не были равны. Например, три наименьших значения равны десяти. Если бы время измеряли более точно, то эти значения могли бы различаться и составля-

ли бы, скажем, 10,2 с, 10,5 с, 10,7 с. В этом случае они получили бы ранги, 1, 2 и 3 соответственно. Но поскольку полученные значения равны, каждое из них получает средний ранг: $(1+2+3)/3=2$.

4. Общая сумма рангов должна совпадать с расчетной, которая определяется по формуле

$$\sum_{i=1}^{n}(R_i)=\frac{N\cdot(N+1)}{2},$$

где N - общее количество ранжируемых наблюдений (значений). Несовпадение реальной и расчетной сумм рангов будет свидетельствовать об ошибке, допущенной при начислении рангов или их суммировании. Прежде чем продолжить работу, необходимо найти ошибку и устранить ее.

5.2. Алгоритм расчета коэффициента ранговой корреляции Спирмена

1. Определить, какие два признака или две иерархии признаков будут участвовать в сопоставлении как переменные A и B.

2. Проранжировать значения переменной A, начисляя ранг 1 наименьшему значению, в соответствии с правилами ранжирования. Занести ранги в таблицу по порядку номеров испытуемых или признаков.

3. Проранжировать значения переменной B в соответствии с теми же правилами. Занести ранги в таблицу по порядку номеров испытуемых или признаков.

4. Подсчитать разности d между рангами A и B по каждой строке и занести в таблицу.

5. Возвести каждую разность в квадрат: d^2. Эти значения также занести в таблицу.

6. Подсчитать сумму квадратов $\sum_{i=1}^{n}=d^2$.

7. При наличии одинаковых рангов рассчитать поправки:

$$T_a=\sum_{i=1}^{k}\left(a^3-a\right)/12, \quad T_в=\sum_{i=1}^{m}\left(b^3-b\right)/12,$$

где a – объем каждой группы одинаковых рангов в ранговом ряду A; b – объем каждой группы одинаковых рангов в ранговом ряду B.

8. Рассчитать коэффициент ранговой корреляции r_B по формуле

а) при отсутствии одинаковых рангов

$$r_в=1-6\cdot\frac{\sum_{i=1}^{n}d^2}{N\cdot\left(N^2-1\right)}.$$

б) при наличии одинаковых рангов

$$r_в = 1 - 6 \cdot \frac{\sum_{i=1}^{n} d^2 + T_a + T_b}{N \cdot \left(N^2 - 1\right)},$$

где $\sum_{i=1}^{n} d^2$ – сумма квадратов разностей между рангами; T_a и T_b – поправки на одинаковые ранги; N – количество испытуемых или признаков, участвовавших в ранжировании.

9. Определить критические значения $r_{вкр.}$ для данного n. Если $r_в$ превышает критическое значение или по крайней мере равен ему, корреляция достоверно отличается от нуля.

Пример. Определение корреляция между двумя признаками [3]. В исследовании, моделирующем деятельность авиадиспетчера, группа испытуемых, студентов психологического факультета ЛГУ, проходила подготовку перед началом работы на тренажере. Испытуемые должны были решать задачи по выбору оптимального типа взлетно-посадочной полосы для заданного типа самолета. Связано ли количество ошибок, допущенных испытуемыми в тренировочной сессии, с показателями вербального и невербального интеллекта, измеренными по методике Д. Векслера? Результаты эксперимента приведены в таблице 6.

Таблица 6

Показатели количества ошибок в тренировочной сессии и показатели уровня вербального и невербалъного интеллекта у студентов-психологов $(N = 10)$

Испытуемый		Количество ошибок	Показатель вербального интеллекта	Показатель невербального интеллекта
1	Т. А.	29	131	106
2	П. А.	54	132	90
3	Ч. И.	13	121	95
4	Ц. А.	8	127	116
5	А. А.	14	136	127
6	К. Е.	26	124	107
7	К. А.	9	134	104
8	Б. Л.	20	136	102
9	И. А.	2	132	111
10	Ф. В.	17	136	99
	Суммы	192	1309	1057
	Средние	19,2	130,9	105,7

Сначала попробуем ответить на вопрос, связаны ли между собой показатели количества ошибок и вербального интеллекта. Сформулируем гипотезы. H_0: Корреляция между показателем количества ошибок в тренировочной сессии и уровнем вербального интеллекта не отличается от нуля. H_1: Корреляция между показателем количества ошибок в тренировочной сессии и уровнем вербального интеллекта статистически значимо отличается от нуля. Далее необходимо проранжировать оба показателя. Приписывая меньшему значению меньший ранг, затем подсчитывают разности между рангами, которые получил каждый испытуемый по двум переменным (признакам), и возводим эти разности в квадрат. Поместим все необходимые расчеты в таблицу 7

Таблица 7

Расчет d^2 для рангового коэффициента корреляции Спирмена при сопоставлении показателей количества ошибок и вербального интеллекта у студентов-психологов $(N = 10)$

Испытуемый		Переменная A: Количество ошибок		Переменная B: Вербальный интеллект		d (ранг A – ранг B)	d^2
		Индивидуальные значения	Ранг	Индивидуальные значения	Ранг		
1	2	3	4	5	6	7	8
1	Т. А.	29	9	131	4	5	25
2	П .А.	54	10	132	5,5	4,5	20,2
3	Ч .И.	13	4	121	1	3	9
4	Ц. А.	8	2	127	3	-1	1
5	А .А.	14	5	136	9	-4	16
6	К. Е.	26	8	124	2	6	36
7	К. А.	9	3	134	7	-4	16
8	Б .Л.	20	7	136	9	-2	4
9	И. А.	2	1	132	5,5	-4,5	20,2
10	Ф. В.	17	6	136	9	-3	9
	Суммы		55		55	0	164,4

В таблице 7 в первой графе представлены значения по показателю количества ошибок; в следующей графе – их ранги. В третьей графе слева представлены значения по показателю вербального интеллекта; в следующей графе – их ранги.

В седьмой представлены разности d между рангом по переменной A (количество ошибок) и переменной B (вербальный интеллект). В последней графе представлены квадраты разностей d^2.

Коэффициент ранговой корреляции Спирмена подсчитывается по формуле

$$r_в = 1 - \frac{6 \cdot \left(\sum_{i=1}^{n} \left(d^2 \right) + T_в \right)}{N \cdot \left(N^2 - 1 \right)}$$

где d - разность между рангами по двум переменным для каждого испытуемого,

$$r_в = 0,039 .$$

N – количество ранжируемых значений, в данном случае количество испытуемых. Полученное эмпирическое значение $r_в$ близко к нулю. И все же определим критические значения $r_в$ при $N = 10$ $r_{вкр} = 0,79$ при уровне значимости $\alpha = 0,01$ ($r_{вкр} = 0,634$, для $\alpha = 0,05$ и $r_{вкр} = 0,437$, для $\alpha = 0,1$).

$$r_{вкр} = t_{кр}(\alpha, \quad n-2) \cdot \sqrt{\left(1 - r_в{}^2\right) \big/ (n-2)},$$

где t_{kp} – критическая точка распределения Стьюдента при уровне значимости α.

Вывод: *гипотеза* H_0 *принимается. Корреляция между показателем количества ошибок в тренировочной сессии и уровнем вербального интеллекта не отличается от нуля.*

Лекция 6. Многомерное шкалирование

6.1. Назначение

Основная цель многомерного шкалирования (МШ) - выявление структуры исследуемого множества объектов - близка к цели факторного и кластерного анализа. Так же, как в факторном анализе, под структурой понимается набор основных факторов (в данном случае - шкал), по которым различаются и могут быть описаны эти объекты. Однако в отличие от факторного, но подобно кластерному анализу, исходной информацией для МШ являются данные о различии или близости объектов.

В психологии чаще всего исходными данными для МШ являются субъективные суждения испытуемых о различии или сходстве стимулов (объектов). Центральное положение МШ заключается в том, что в основе таких суждений лежит ограниченное число субъективных признаков (критериев), определяющих различение стимулов, и человек, вынося свои суждения, явно или неявно учитывает эти критерии. Основываясь на этом положении, решается главная задача МШ - реконструкция психологического пространства, заданного небольшим числом измерений-шкал, и расположение в нем точек-стимулов таким образом, чтобы расстояния между ними наилучшим образом соответствовали исходным субъективным различиям. Таким образом, шкала в МШ интерпретируется как критерий, лежащий в основе различий стимулов.

Геометрические представления МШ основаны на аналогии между понятием различия в психологии и понятием расстояния в пространстве. Чем более субъективно сходны между собой два объекта, тем ближе в реконструируемом пространстве признаков должны находиться соответствующие этим объектам точки. Исходя из такой дистанционной модели, по субъективным данным о различии одного объекта от другого реконструируется их взаимное расположение в пространстве нескольких признаков. Эти признаки трактуются как субъективные шкалы — критерии, которыми пользуется человек при различении объектов. А расстояние между объектами в этом пространстве есть определенная функция от исходных оценок различия.

Общая схема МШ формально может быть представлена следующим образом. На основе суждений экспертов (испытуемых) в отношении интересующих исследователя объектов вначале составляется **симметричная матрица попарных** различий (или матрицы - по одной для каждого эксперта). Допускается и использование **данных о предпочтениях**, содержащих упорядочивание каждым экспертом совокупности объектов по степени их предпочтения. Сравниваемыми объектами могут быть члены коллектива, предметы домашнего обихода, литературные отрывки, цветовые оттенки и т. д. Модель МШ предполагает, что эксперт производит сравнение, осознанно или нет пользуясь одним или несколькими признаками этих объектов. В отношении сотрудников подразделения такими признаками могут быть должностной статус, профессионализм, доброжелательность и т. д.

В процессе МШ определяется, сколько признаков-шкал необходимо и достаточно для построения координатного пространства и размещения в нем точек-объектов. Если δ_{ij} - это оценка экспертом различия между объектами, а число признаков, которыми пользуется эксперт при сравнении, - K, то задача многомерного шкалирования сводится к определению всех x_{ik} и x_{jk} как координат этих объектов в пространстве K признаков. При этом предполагается, что число критериев, которыми пользуется эксперт, значительно меньше числа сравниваемых объектов. Если, например, i *и* j - сотрудники, а признак K - доброжелательность, то x_{ik} и x_{jk} - доброжелательность этих сотрудников. Важно отметить, что исследователю эмпирически даны только **оценки различий** δ_{ij}. Величины значений признаков x_{ik} и x_{jk} непосредственно не даны, но оцениваются в результате МШ в виде матрицы:

$$X = \begin{vmatrix} x_{i1} & \dots & x_{1k} \\ \dots & \dots & \dots \\ x_{p1} & \dots & x_{pk} \end{vmatrix}$$

где P - количество сравниваемых объектов, K - количество шкал.

Элементы x_{ij} указанной матрицы рассматриваются как координаты P объектов в пространстве K признаков. Пространство определено так, что чем больше исходное различие между объектами, тем дальше друг от друга расположены объекты в этом пространстве. Каждая шкала результирующего пространства получает интерпретацию через объекты, находящиеся на противоположных полюсах шкалы.

Следует отметить, что исходными данными для МШ могут являться не только субъективные оценки различий, но и обычные данные типа «объект-признак». Но поскольку МШ предназначено для анализа различий, то для данных типа «объект-признак» необходимо, во-первых, определить, что будет подлежать шкалированию - сами объекты (строки) или признаки (столбцы). Во-вторых, необходимо задать метрику различий - то, как будут определяться различия между всеми парами изучаемых элементов.

Выбирая МШ, исследователь должен отдавать себе отчет в том, что это довольно сложный метод, применение которого к тому же связано с неизбежными потерями исходной информации о различии объектов. Поэтому, если задача исследования ограничивается классификацией объектов и нет оснований полагать, что эта классификация обусловлена небольшим числом независимых причин - критериев различий, то целесообразнее воспользоваться более простым методом - кластерным анализом.

Рассмотрим исходные данные и основные результаты применения МШ на простом примере. Попытаемся, исходя из субъективных оценок расстояний между совокупностью объектов, реконструировать конфигурацию их взаимного расположения. Допустим, субъекту предъявляется 10 объектов, расположенных на плоскости в некоторой произвольной конфигурации, и дана инструкция оценить расстояние между каждым объектом и всеми остальными, присвоив 1 наименьшему расстоянию, 2 - следующему по величине и т. д. Примерно одинаковым расстояниям разрешим присваивать одинаковые числовые значения. В результате выполнения такого задания наблюдатель заполнил нижний треугольник матрицы попарных различий между объектами, в данном случае - расстояний (табл. 8). Можно ли восстановить исходную конфигурацию объектов по такой матрице различий? Оказывается, МШ справляется с подобными задачами. Применение программы неметрического МШ (программа SPSS) дает 2-шкальное решение (табл. 8).

Таблица 8

Субъективные оценки расстояний между 10 объектами

№	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	0									
2	1	0								
3	2	1	0							
4	3	2	1	0						
5	3	2	1	1	0					
6	3	2	2	2	1	0				
7	3	3	3	3	2	1	0			
8	2	2	2	3	2	1	1	0		
9	1	1	2	3	2	2	2	1	0	
10	2	1	1	2	1	1	2	1	1	0

В связи с тем, что типичными исходными данными для МШ в психологии являются все же непосредственные оценки различий, изложение этой главы сопровождается примерами анализа исходной информации именно этого типа. Тем не менее, подчеркнем, что для МШ допустимо применение и любых других исходных данных.

6.2. Многомерные методы и модели

Таблица 9

Результаты МШ субъективных оценок расстояний между 10 объектами (по данным табл. 8)

№ объектов	Шкала 1 (Dim. 1)	Шкала 2 (Dim. 2)
1	0,932	-0,006
2	0,471	-0,302
3	0,019	-0,559
4	-0,471	-0,804
5	-0,497	-0,257
6	-0,493	0,263
7	-0,461	0,810
8	0,026	0,558
9	0,474	0,296
10	0,000	0,000

Каждая строчка таблицы - это координаты соответствующего объекта на плоскости. Графическое изображение всех 10 точек, в соответствии с табл. 9, приведено на рис. 8.

Взаимное расположение объектов в точности соответствует исходной конфигурации, предлагаемой наблюдателю (рис. 9). При этом обращает на себя внимание тот факт, что информация, полученная от наблюдателя, носит неметрический характер, так как расстояния оценивались по шкале порядка. Итоговая же конфигурация воспроизводит метрические соотношения в расположении объектов. Это связано с тем, что информация о различиях, содержащаяся в матрице субъективных оценок, хотя и является по сути порядковой, но обладает избыточностью, которая и позволяет восстановить метрические соотношения.

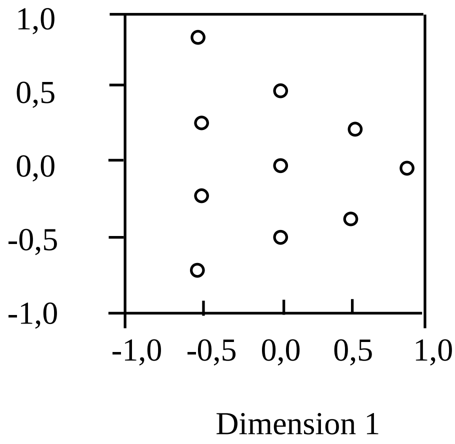

Рис. 2. Субъективное пространство 10 объектов по табл. 9.

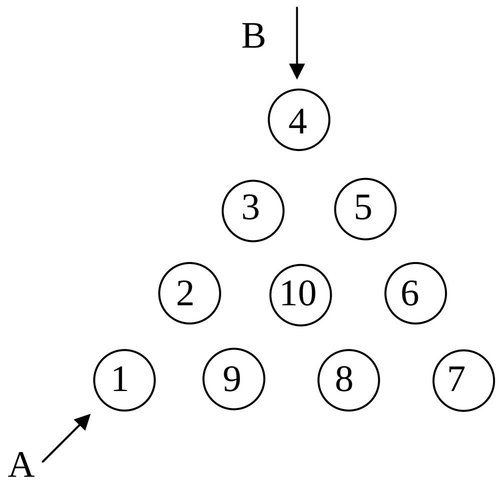

Рис.3

МШ в своих основных трех модификациях позволяет решать три группы задач:

1. Исходные данные - прямые оценки субъектом различий между стимулами или вычисленные расстояния между объектами, характеризующимися совокупностью признаков. Примером второго типа данных могут являться расстояния между ролями (объектами), вычисленные по совокупности конструктов (репертуарные решетки Келли). МШ позволяет реконструировать психологическое пространство субъекта, как конфигурацию стимулов в осях существенных признаков, по которым эти стимулы различаются субъектом.

2. Исходные данные - те же, что и в предыдущем случае субъективные различия между стимулами (оцененные прямо или вычисленные), но полученные не от одного, а от группы субъектов. Взвешенная модель индивидуальных различий позволяет получить групповое психологическое пространство стимулов в осях общих для данной группы существенных признаков. Дополнительно к этому для каждого субъекта - индивидуальные веса признаков как меру учета соответствующих точек зрения при различении стимулов.

3. Исходные данные - результаты упорядочивания каждым из группы субъектов набора стимулов по степени предпочтения. Модель анализа пред почтений позволяет получить групповое психологическое пространство стимулов в осях существенных признаков и размещенные в этом же пространстве идеальные точки для каждого субъекта.

6.3. Неметрическая модель

Это основной вариант многомерного шкалирования, применяемый в настоящее время. Он лежит в основе всех остальных вариантов метода, Исходные данные для этого метода - матрица размерностью $P \times P$, каждый элемент которой - мера (оценка) различия между двумя объектами из Р. Рассмотрим кратко основные математико-статистические идеи метода, необходимые для его использования на практике.

Многомерное шкалирование, как новый шаг в математике, начинается с **метрического шкалирования**, предложенного в 50-х годах У. Торгерсоном. В модели Торгерсона вводится жесткое предположение о том, что оценки различия между объектами равны линейному расстоянию между ними в евклидовом пространстве.

Пусть δ_{ij} - имеющаяся в распоряжении исследователя оценка различия между объектами i и j - координаты этих объектов по оси k, одной из осей искомого пространства размерностью K. Расстояние между объектами в искомом пространстве обозначим как d_{ij}. Тогда основное предположение Торгерсона можно выразить формулой:

$$\delta_{ij} = d_{ij} = \left[\sum_{k=1}^{k} \left(x_{ik} - x_{jk} \right)^2 \right]^{1/2}.$$

Торгерсон показал, что при соблюдении этого условия возможен переход от исходной матрицы различий между стимулами к их координатам в пространстве K признаков. Тогда, по Торгерсону, справедливо выражение:

$$\delta^*_{ij} = \sum_{k=1}^{K} x_{ik} \cdot x_{jk}$$

или в матричной виде: $\Delta^* = X\,X'$

где X — матрица координат стимулов, размерностью $P \cdot K$.

Это уравнение аналогично главному уравнению факторного анализа, и решается оно относительно X методом главных компонент с заданным числом K.

В современных алгоритмах МШ метод Торгерсона используется на этапе предварительной оценки координат объектов по матрице исходных различий. Далее следует **неметрический** этап, соответствующий неметричности исходных данных. На этом этапе исходят из требования **соответствия рангового порядка расстояний между объектами** в результирующем пространстве ранговому порядку исходных различий, то есть, используя принятые обозначения:

$d_{ij} \le d_{lm} \Rightarrow \delta_{ij} \le \delta_{lm}$ для любых $i,\ j,\ l,\ m$ - номеров объектов.

Основной мерой выполнения этого требования является специальный показатель, который называется **стресс - мера отклонения итоговой конфигурации объектов от исходных оценок различия** в смысле указанного требования рангового соответствия. Иногда дополнительно применяют **коэффициент отчуждения** тоже как меру подгонки неметрической модели к данным о различии.

Не рассматривая подробно вычислительные проблемы многомерного шкалирования, укажем, что его алгоритм направлен на нахождение оценок координат объектов, минимизирующих значение стресса. Построен этот алгоритм как **градиентная процедура.** Первый шаг алгоритма - получение стартовой конфигурации, как правило, методом Торгерсона. На каждом последующем шаге, или итерации, координаты стимулов изменяются в сторону уменьшения значения стресса, вычисленного на предыдущем этапе. Итерации повторяются многократно, до выполнения одного из трех заданных изначально условий (в программе SPSS): достижения минимального значения стресса; достижения минимальной разницы между последним и предыдущим значениями стресса; выполнения максимального заданного числа итераций. Каждое из трех условий задано в программе «по умолчанию», но может изменяться пользователем. Уменьшая пороговые величины стресса и его изменения, увеличивая максимальное число итераций, пользователь может добиться повышения точности окончательного решения. Показателем точности является конечная величина стресса. Наиболее приемлемые величины стресса находятся в диапазоне от 0,05 до 0,2.

Одна из основных проблем, возникающих перед исследователем в МШ - это проблема размерности К. Как и при проведении факторного анализа, в М Ш требуется предварительное определение числа шкал. Поэтому от исследователя требуется получить несколько решений в пространствах разной размерности и выбрать из них лучшее. Один из критериев размерности, применяемый для предварительной оценки числа шкал, аналогичен критерию отсеивания Кеттелла в факторном анализе: строится график зависимости стресса от числа шкал по результатам решения в разных размерностях. Истинная размерность соответствует точке перегиба графика после резкого его спада.

Другой критерий числа шкал - абсолютная величина стресса. Если решение одномерно, то приемлемая величина стресса - менее 0,1. Если решение размерностью 2 и выше, то приемлемы значения стресса, меньшие 0,1- 0,15. Однако если уровень ошибок измерения или выборки высок, то можно признать решение и с более высокими значениями стресса. Дополнительно вычисляется величина R2 (RSQ), которая показывает долю дисперсии исходных различий (от единичной), учтенную выделенными шкалами. Чем ближе RSQ к единице, тем полнее данные шкалы воспроизводят исходные различия между объектами.

Окончательный выбор размерности решения определяется на основе критериев интерпретируемости и воспроизводимости, так же, как в факторном анализе. Тем не менее, при размерности 2 и выше, следует избегать решений с величиной стресса выше 0,2. Обычный путь для этого - повышение размерности и исключение объектов.

Результаты применения метода - **таблица координат объектов** в пространстве K шкал-признаков, **величины стресса** и RSQ, **интерпретация** шкал и взаимного расположения объектов по таблице координат.

Исследовалась структура представлений студента о многомерных методах, применяемых в психологии. Студенту было предложено сравнить попарно по степени различия пять методов: множественный регрессионный анализ (МРА), дискриминантный анализ (ДА), кластерный анализ (КА), факторный анализ (ФА) и многомерное шкалирование (МШ). При сравнении было предложено использовать 5-балльную шкалу (1 - очень похожи, 5 - совсем разные). Результаты сравнения приведены в табл. 10.

Таблица 10

Результаты попарного сравнения пяти методов многомерного анализа

Методы	Обозначения	МРА	ДА	КА	ФА	МШ
МРА	MRA	0				
ДА	DA	2	0			
КА	KA	5	2	0		
ФА	FA	2	3	5	0	
МШ	MDS	5	5	3	3	0

Лекция 7. Кластерный анализ

7.1. Назначение

Кластерный анализ решает задачу построения классификации, то есть разделения исходного множества объектов на группы (классы, кластеры). При этом предполагается, что у исследователя нет исходных допущений ни о составе классов, ни об их отличии друг от друга. Приступая к кластерному анализу, исследователь располагает лишь информацией о характеристиках (признаках) для объектов, позволяющей судить о сходстве (различии) объектов, либо только данными об их попарном сходстве (различии). Синонимы кластерного анализа: автоматическая классификация, таксономический анализ, анализ образов.

Первые исследования с использованием этого метода начинают появляться после публикации книги «Начала численной таксономии» биологами Р. Сокэл и П. Снит в 1963 году. Тем не менее, до сих пор в психологии известны лишь единичные случаи удачного применения кластерного анализа, несмотря на его исключительную простоту. Обычно психологи используют для решения простой задачи классификации (объектов, признаков) такой сложный метод, как факторный анализ. Вместе с тем, кластерный анализ не только гораздо проще и нагляднее решает эту задачу, но и имеет несомненное преимущество: результат его применения не связан с потерей даже части исходной информации о различиях объектов или корреляции признаков.

Варианты кластерного анализа - это множество простых вычислительных процедур, используемых для классификации объектов. Классификация объектов - это группирование их в классы так, чтобы объекты в каждом классе были более похожи друг на друга, чем на объекты из других классов. Более точно, кластерный анализ - это процедура упорядочивания объектов в сравнительно однородные классы на основе попарного сравнения этих объектов по предварительно определенным и измеренным критериям.

Существует множество вариантов кластерного анализа, но наиболее широко используются методы, объединенные общим названием иерархический кластерный анализ (Hierarchical Cluster Analysis). В дальнейшем под кластерным анализом мы будем подразумевать именно эту группу методов. Рассмотрим основной принцип иерархического кластерного анализа на примере.

Пример. Предположим, 10 студентам предложили оценить проведенное с ними занятие по двум критериям: увлекательность и полезность. Для оценки использовалась 10-балльная шкала. Полученные данные (2 переменные для 10 студентов) графически представлены в виде графика двумерного рассеивания (рис. 19.1). Конечно, классификация объектов по результатам измерения всего двух переменных не требует применения кластерного анализа: группировки и так можно выделить путем визуального анализа. Так, в данном случае наблюдаются четыре группировки: 9, 2, 3 - занятие полезное, но не увлекательное; 1, 10, 8 - занятие увлекательное, но бесполезное; 5,7 - занятие и полезное и увлекательное; 4, 6 - занятие умеренно увлекательное и умеренно полезное. Даже для трех переменных можно обойтись и без кластерного анализа, так как компьютерные программы позволяют строить трехмерные графики. Но для 4 и более переменных визуальный анализ данных практически невозможен. Тем не менее, общий принцип классификации объектов при помощи кластерного анализа не зависит от количества измеренных признаков, так как непосредственной информацией для этого метода являются различия между классифицируемыми объектами.

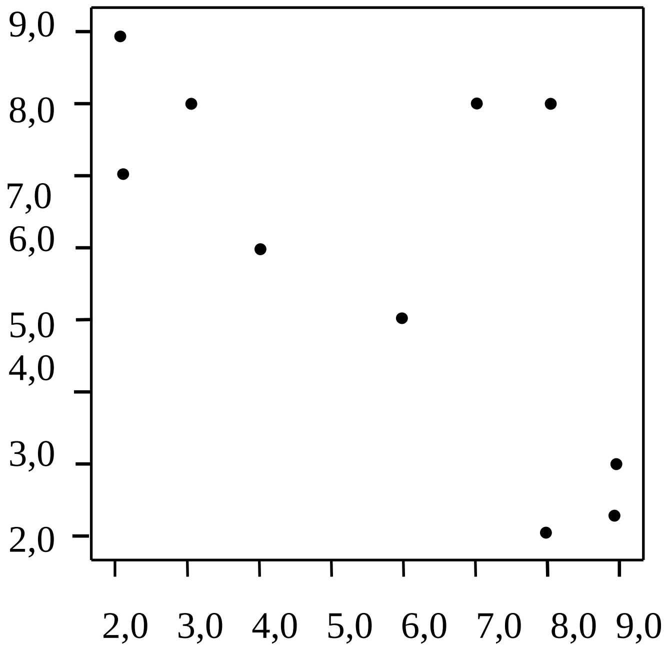

Рис. 4. График двумерного рассеивания переменных «увлекательность» и «польза» для 10 студентов

Кластерный анализ объектов, для которых заданы значения количественных признаков начинается с расчета различий для всех пар объектов. Пользователь может выбрать по своему усмотрению меру различия, обзор которых приведен в соответствующем разделе. В качестве меры различия выбирается расстояние между объектами в P - мерном пространстве признаков, чаще всего - евклидово расстояние или его квадрат. В данном случае $P=2$ и евклидово расстояние между объектами i и j определяется формулой:

$$d_{ij}=\sqrt{\left(x_i-x_j\right)^2+\left(y_i-y_j\right)^2},$$

где x - это значения одного, а y - другого признака.

На первом шаге кластерного анализа путем перебора всех пар объектов определяется пара (или пары) наиболее близких объектов, которые объединяются в первичные кластеры. Далее на каждом шаге к каждому первичному кластеру присоединяется объект (кластер), который к нему ближе. Этот процесс повторяется до тех пор, пока все объекты не будут объединены в один кластер. Критерий объединения объектов (кластеров) может быть разным и определяется методом кластерного анализа. Основным результатом применения иерархического кластерного анализа является **дендрограмма** - графическое изображение последовательности объединения объектов в кластеры. Для данного примера дендрограмма приведена на рис. 5.

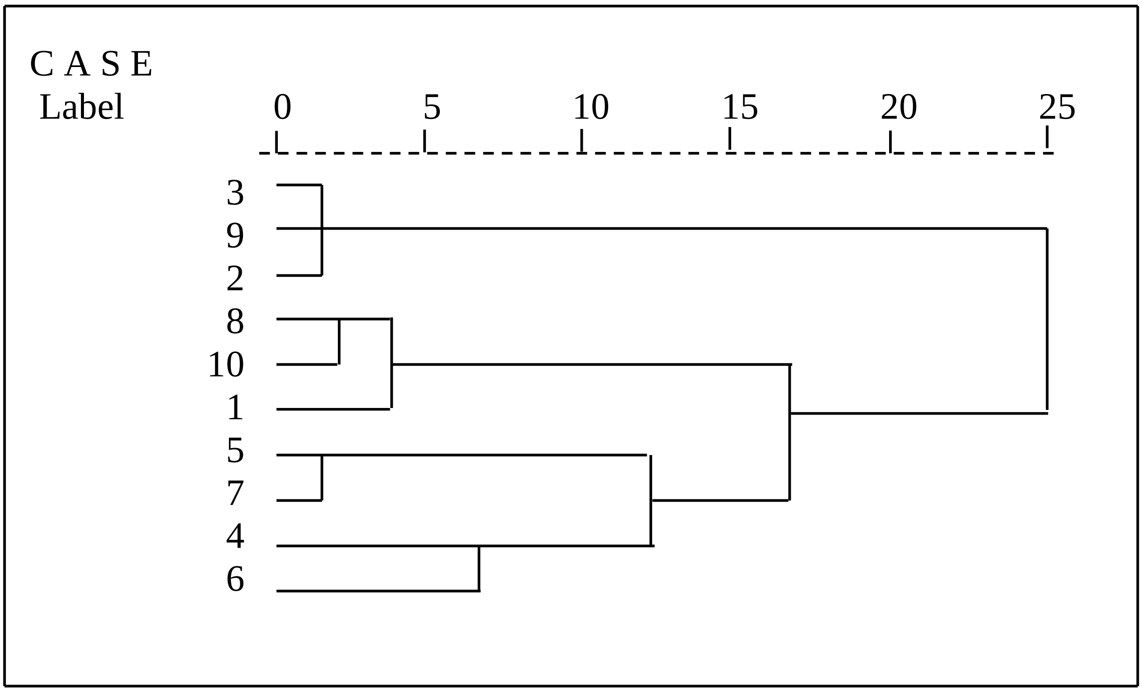

Рис. 5. Дендрограмма для 10 студентов (метод средней связи)

На дендрограмме номера объектов следуют по вертикали. По горизонтали отмечены расстояния (в условных единицах), на которых происходит объединение объектов в кластеры. На первых шагах происходит образование кластеров: (3,9,2) и (5,7). Далее образуется кластер (8,10,1) - расстояния между этими объектами больше, чем между теми, которые были объединены на предыдущих шагах. Следующий кластер - (4, 6). Далее в один кластер объединяются кластеры (5, 7) и (4, 6), и т. д. Процесс заканчивается объединением всех объектов в один кластер. Количество кластеров определяет по дендрограмме сам исследователь. Так, судя по дендрограмме, в данном случае можно выделить три или четыре кластера.

Как видно из примера, кластерный анализ - это комбинаторная процедура, имеющая простой и наглядный результат. Широта возможного применения кластерного анализа очевидна; можно указать ряд задач, при решении которых кластерный анализ является более эффективным, чем другие многомерные методы:

- разбиение совокупности испытуемых на группы по измеренным признакам с целью дальнейшей проверки причин межгрупповых различий по внешним критериям, например, проверка гипотез о том, проявляются ли типологические различия между испытуемыми по измеренным признакам;
- применение кластерного анализа как значительно более простого и наглядного аналога факторного анализа, когда ставится только задача группировки признаков на основе их корреляции;
- классификация объектов на основе непосредственных оценок различий между ними (например, исследование социальной структуры коллектива по данным социометрии - по выявленным межличностным предпочтениям).

Несмотря на различие целей проведения кластерного анализа, можно выделить общую его последовательность как ряд относительно самостоятельных шагов, играющих существенную роль в прикладном исследовании:

1. Отбор объектов для кластеризации. Объектами могут быть, в зависимости от цели исследования: а) испытуемые; б) объекты, которые оцениваются испытуемыми; в) признаки, измеренные на выборке испытуемых.

2. Определение множества переменных, по которым будут различаться объекты кластеризации. Для испытуемых - это набор измеренных признаков, для оцениваемых объектов - субъекты оценки, для признаков испытуемые. Если в качестве исходных данных предполагается использовать результаты попарного сравнения объектов, необходимо четко определить критерии этого сравнения испытуемыми (экспертами).

3.**Определение меры различия** между объектами кластеризации. Это первая проблема, которая является специфичной для методов анализа различий: многомерного шкалирования и кластерного анализа.

4. Выбор и применение метода классификации для создания групп сходных объектов. Это вторая и центральная проблема кластерного анализа. Ее весомость связана с тем, что разные методы кластеризации порождают разные группировки для одних и тех же данных. Хотя анализ и заключается в обнаружении структуры, на деле в процессе кластеризации структура привносится в данные, и эта привнесенная структура может не со впадать с реальной.

5. Проверка достоверности разбиения на классы. Последний этап не всегда необходим, например, при выявлении социальной структуры группы. Тем не менее, следует помнить, что кластерный анализ **всегда** разобьет совокупность объектов на классы, независимо от того, существуют ли они на самом деле. Поэтому бесполезно доказывать существенность разбиения на классы, например, на основании достоверности различий между классами по признакам, включенным в анализ. Обычно проверяют **устойчивость группировки** -

на повторной идентичной выборке объектов. **Значимость разбиения** проверяют по внешним критериям - признакам, не вошедшим в анализ.

7.2. Методы кластерного анализа

Непосредственными данными для применения любого метода кластеризации является **матрица различий** между всеми парами объектов. Определение или задание меры различия является первым и необходимым шагом кластерного анализа.

Из всего множества методов кластеризации наиболее распространены так называемые **иерархические агломеративные методы.** Название указывает на то, что классификация осуществляется путем последовательного объединения (агломерации) объектов в группы, оказывающиеся в результате иерархически организованными. Эти методы — очень простые комбинаторные процедуры, отличающиеся критерием объединения объектов в кластеры.

Критерий объединения многократно применяется ко всей матрице попарных расстояний между объектами. На первых шагах объединяются наиболее близкие объекты, находящиеся на одном уровне сходства. Затем поочередно присоединяются остальные объекты, пока все они не объединятся в один большой кластер. Результат работы метода представляется графически в виде дендрограммы - ветвистого древовидного графика.

Существуют различные методы иерархического кластерного анализа, в частности, в программе SPSS предлагается 7 методов. Каждый метод дает свои результаты кластеризации, но три из них являются наиболее типичными. Поэтому рассмотрим результаты применения этих методов к одним и тем же данным из примера рис.4.

Dendrogram using Single Linkage

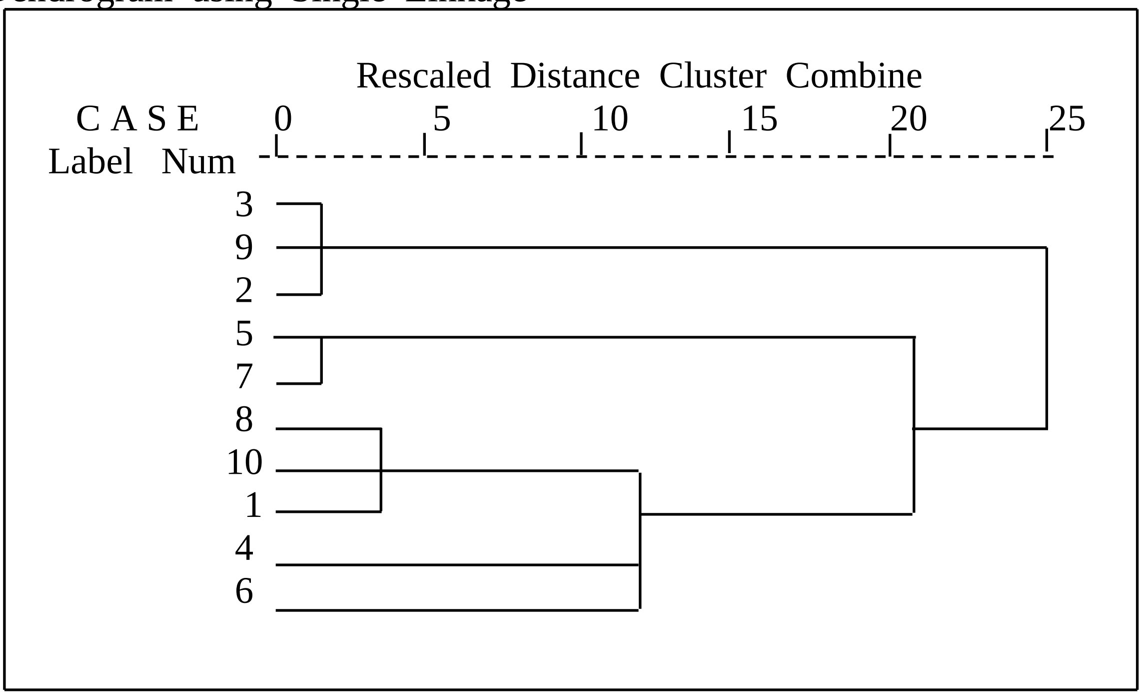

Рис. 6. Дендрограмма для 10 студентов (метод одиночной связи)

Метод одиночной связи (Single Linkage) - наиболее понятный метод, который часто называют методом «**ближайшего соседа»** (Nearest Neighbor). Алгоритм начинается с поиска двух наиболее близких объектов, пара которых образует первичный кластер. Каждый последующий объект присоединяется к тому кластеру, **к одному из объектов которого он ближе.**

На рис. 6 приведен результат применения метода. Сопоставляя эту дендрограмму с рис. 4, можно заметить, что объект 4 присоединяется к кластеру (8, 10, 1) и на том же расстоянии - к объекту 6 в связи с тем, что расстояние от объекта 4 до объекта 6 такое же, что и до объекта 1. Из рисунка видно, что метод имеет тенденцию к образованию длинных кластеров «цепочного» вида. Таким образом, метод имеет тенденцию образовывать **небольшое число крупных кластеров.** К особенностям метода можно отнести и то, что результаты его применения часто не дают возможности определить, как много кластеров находится в данных.

Метод полной связи (Complete Linkage) часто называют методом «дальнего соседа» (Furthest Neighbor). Правило объединения этого метода подразумевает, что новый объект присоединяется к тому кластеру, самый далекий элемент которого находится ближе к новому объекту, чем самые далекие элементы других кластеров. Это правило является противоположным предыдущему и более жестким. Поэтому здесь наблюдается тенденция к выделению большего числа компактных кластеров, состоящих из наиболее похожих элементов.

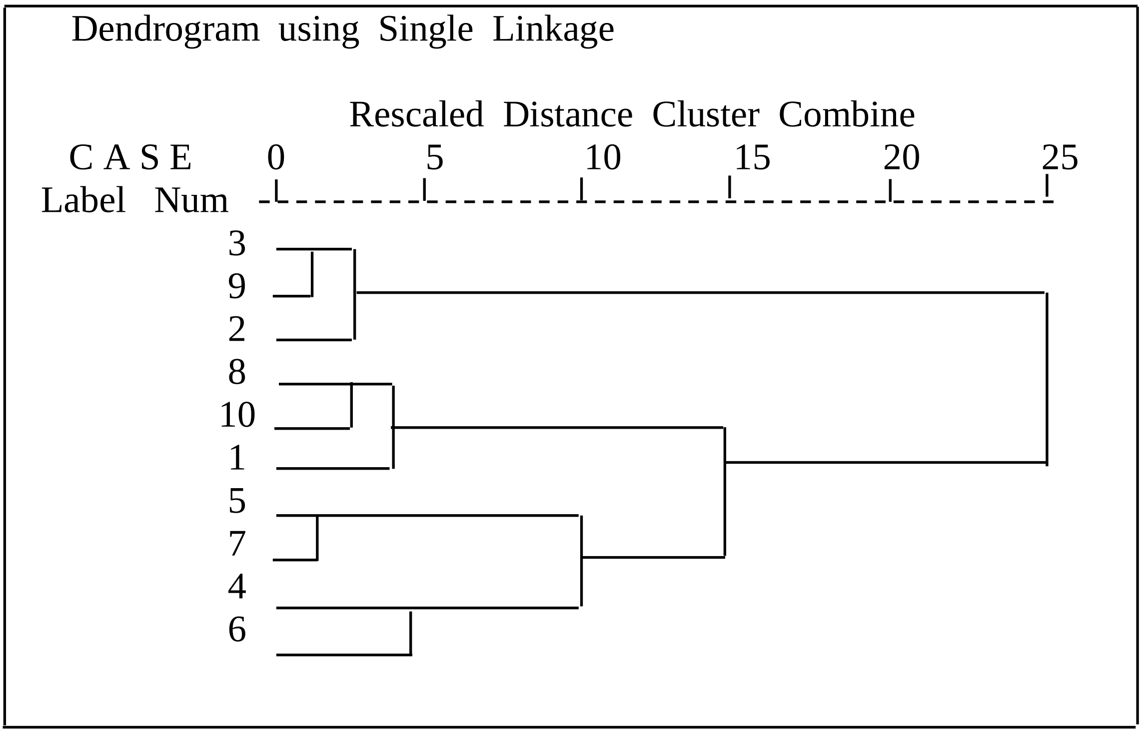

Рис.7. Дендрограмма для 10 студентов (метод полной связи)

Сравним результат применения метода полной связи (рис. 7), метода одиночной связи (рис. 6) и фактическую конфигурацию объектов (рис. 5). Различия в работе методов проявляются прежде всего в отношении объектов 4 и 6. Метод полной связи объединяет их в отдельный кластер и соединяет с кластером (5, 7) раньше, чем с кластером (8, 10, 1) - в отличие от метода одиночной связи. Объект 4 присоединяется сначала к объекту 6, потому что этот последний к нему **ближе, чем самый дальний объект кластера (**8, 10, 1). На этом же основании кластер (4, 6) присоединяется к кластеру (5,7), потому что самый дальний объект 6 кластера (4, 6) ближе к самому дальнему объекту 7 кластера (5, 7), чем к самому дальнему объекту 8 кластера (8, 10, 1).

Метод средней связи (Average Linkage) или межгрупповой связи (Between Groups Linkage) занимает промежуточное положение относительно крайностей методов одиночной и полной связи. На каждом шаге вычисляется среднее арифметическое расстояние между каждым объектом из одного кластера и каждым объектом из другого кластера. Объект присоединяется к данному кластеру, если это среднее расстояние меньше, чем среднее расстояние до любого другого кластера. По своему принципу этот метод должен давать более точные результаты классификации, чем остальные методы. То, что объединение кластеров в методе средней связи происходит при расстоянии большем, чем в методе одиночной связи, но меньшем, чем в методе полной связи, и объясняет промежуточное положение этого метода. Результат применения метода изображен на рис. 5. Поскольку объектов в нашем примере немного, результаты применения методов полной и средней связи различаются незначительно.

В реальных исследованиях обычно имеются десятки классифицируемых объектов, и применение каждого из указанных методов дает существенно разные результаты для одних и тех же данных. Опыт и литературные данные свидетельствуют, что наиболее близкий к реальной группировке результат позволяет получить метод средней связи. Но это не означает бесполезность применения двух других методов. Метод одиночной связи «сжимает» пространство, образуя минимально возможное число больших кластеров. Метод полной связи «расширяет» пространство, образуя максимально возможное число компактных кластеров. Каждый из трех методов привносит в реальное соотношение объектов свою структуру и представляет собой как бы свою точку зрения на реальность. Исследователь, в зависимости от стоящей перед ним задачи, вправе выбрать тот метод, который ему больше подходит.

Численность классов является отдельной проблемой в кластерном анализе. Сложность заключается в том, что не существует формальных критериев, позволяющих определить оптимальное число классов. В конечном итоге это определяется самим исследователем исходя из содержательных соображений. Однако для предварительного определения числа классов исследователь может **обратиться к таблице последовательности агломерации (Agglomeration schedule).** Эта таблица позволяет проследить динамику увеличения различий по шагам кластеризации и определить шаг, на котором отмечается резкое возрастание различий. Оптимальному числу классов соответствует разность между числом объектов и порядкового номера шага, на котором обнаружен перепад различий.

Лекция 8. Уравнение линейной регрессии

8.1. Анализ статистической взаимосвязи между двумя рядами

Данный этап служит для оценки степени тесноты и направления связи между Y и X по их рядам наблюдений и включает следующее:

- вычисление линейного коэффициента парной корреляции;
- заключение о значимости коэффициента корреляции по t-критерию Стьюдента;
- выдачу рекомендации о тесноте и характере связи между Y и X.

Вычисление линейного коэффициента парной корреляции осуществляется по следующей формуле:

$$r = \frac{\sum_{i=1}^{n}\left(x_i - \overline{X}\right)\cdot\left(y_i - \overline{Y}\right)}{\sqrt{\sum_{i=1}^{n}\left(x_i - \overline{X}\right)^2 \cdot \sum_{i=1}^{n}\left(y_i - \overline{Y}\right)^2}} = \frac{\sum_{i=1}^{n}\left(x_i - \overline{X}\right)\cdot\left(y_i - \overline{Y}\right)}{n \cdot \sigma_x \cdot \sigma_y},$$

где $\overline{X}$ и $\overline{Y}$ – среднее, вычисленное по рядам наблюдений соответственно X и Y.

По величине коэффициента корреляции можно количественно оценить тесноту связи $\left(r \in [-1;\ 1]\right)$.

1. Если $r \in [-1;\ 1]$ или $r \in [-1;\ -0,7]$, то между рядами наблюдений существует сильная связь (связь значима).

2. Если $r \in [-0,2;\ 0,2]$, то между рядами наблюдений связь практически отсутствует (связь незначима).

3. Если $r \in [-0,7;\ -0,2]$ или $r \in [0,2;\ 0,7]$, то заключение о тесноте связи необходимо осуществить по t-критерию Стьюдента.

По направлению выделяют связь прямую и обратную. При прямой связи с увеличением (уменьшением) значений фактора X происходит увеличение (уменьшение) значений отклика Y. При обратной связи значения отклика Y изменяются под воздействием фактора X, но в противоположном направлении по сравнению с изменением значений фактора X. Характер направления оценивается по знаку коэффициента корреляции. Если $r > 0$, то связь прямая; если $r < 0$, то связь обратная.

Заключение о тесноте связи по t-критерию Стьюдента осуществляется следующим образом. Сначала вычисляется значение t-критерия Стьюдента по следующей формуле:

$$t_p = \frac{|r| \cdot \sqrt{K}}{\sqrt{1 - r^2}},$$

где $K = (n-2)$ – число степеней свободы; n – количество наблюдений.

Затем на основании заданных уровня значимости α (обычно, $\alpha = 0{,}01$ или $\alpha = 0{,}05$) и числа степеней свободы K определяется табличное значение распределения Стьюдента $t(\alpha, \ K)$ из специальных таблиц. В зависимости от полученных значений t_p и $t(\alpha, \ K)$ формулируется заключение о характере связи между Y и X в следующем виде:

- если $t_p > t \ (\alpha, \ K)$, то при уровне значимости α – связь значима. Есть основание для связи между X и Y.

- если $t_p \le t \ (\alpha, \ K)$, то при уровне значимости α – связь незначима. Нет оснований для связи между X и Y.

8.2. Построение модели парной регрессии

Построение уравнения регрессии $y = a + b \cdot x$ (1) сводится к оценке (расчету) ее параметров (коэффициентов) a и b. Для оценки параметров уравнения регрессии используют ***метод наименьших квадратов*** (МНК). Сущность МНК заключается в нахождении параметров модели, при которых минимизируется сумма квадратов отклонений эмпирических (фактических) значений отклика от теоретических, вычисленных по модели. Сказанное можно записать в следующем виде:

$$S = \sum_{i=1}^{n} \left(y_i - y(x_i)\right)^2 \to \min.$$

Нахождение коэффициентов уравнения регрессии по МНК производится из системы нормальных уравнений. Данную систему получают путем приравнивания нулю всех частных производных функции S. В качестве аргументов функции S рассматриваются коэффициенты уравнения регрессии.

Для уравнения формула регрессии примет следующий вид :

$$S = \sum_{i=1}^{n} \left(y_i - a - b \cdot x_i\right)^2 \to \min.$$

Дифференцируя функцию S по параметрам a и b, получаем

$$\begin{cases} \dfrac{\partial S}{\partial a} = \sum\limits_{i=1}^{n} 2\left(a + bx_i - y_i\right) = 0; \\ \dfrac{\partial S}{\partial в} = \sum\limits_{i=1}^{n} 2\left(a + bx_i - y_i\right) \cdot x_i = 0; \end{cases}$$

Отсюда система нормальных уравнений для нахождения параметров линейной парной регрессии методом наименьших квадратов имеет следующий вид:

$$\begin{cases} na + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i; \\ a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i. \end{cases}$$

Решая систему относительно a и b, получаем следующие формулы:

$$a = \overline{y} - b \cdot \overline{x}, \quad b = \frac{\overline{y \cdot x} - \overline{y} \cdot \overline{x}}{\overline{x^2} - \overline{x}^2},$$

где $\overline{x}$, $\overline{y}$ – соответствующие средние.

8.3. Анализ качества модели парной регрессии

Под анализом понимается оценка аппроксимации (качества) или проверка адекватности построенной модели. Для оценки аппроксимации используются такие характеристики, как средняя ошибка аппроксимации $\overline{A}$ и F критерий (критерий Фишера) F_p.

Средняя ошибка аппроксимации $\overline{A}$ - среднее отклонение расчетных значений отклика от его фактических – вычисляется по формуле:

$$\overline{A} = \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \cdot 100,$$

где y_i – наблюдаемые значения; $\hat{y}_i = a + bx_i$.

Допустимый предел значений $\overline{A}$ - не более 8-10 %.

Вычисление характеристики F_p осуществляется с использованием следующих формул:

$$\sigma_{фак}^2 = \frac{1}{m}\sum_{i=1}^{n}\left(\hat{y}_i - \overline{Y}\right)^2;$$

$$\sigma_{ост}^2 = \frac{1}{n-m-1}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2;$$

$$F_p = \frac{\sigma_{фак}^2}{\sigma_{ост}^2} = \frac{r^2}{1-r^2}(n-2),$$

где y_i – эмпирические (наблюдаемые) значения Y; $\hat{y}_i$ – расчетные значения y; $\overline{Y}$ – среднее значение, вычисленное по ряду Y; m – количество коэффициентов при переменной x в уравнении регрессии (для уравнения вида (1) – $m=1$).

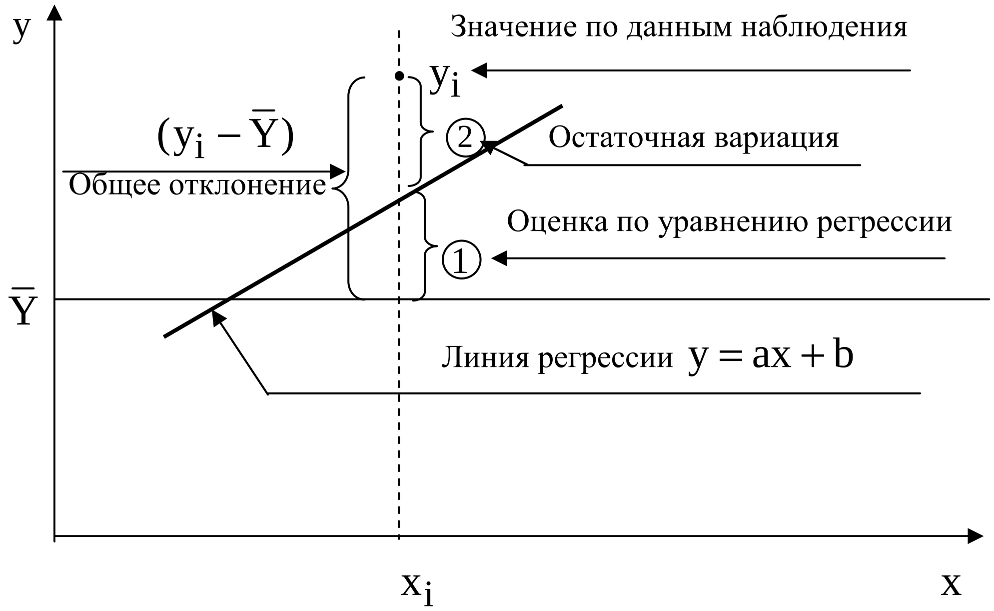

Рис 8. Соотношение между общей вариацией признака, его составляющей, объясняемой влиянием фактора x ①- $(\hat{y} - \bar{Y})$ и остаточной вариацией ②- $(y_i - \hat{y})$

Оценка аппроксимации (адекватность модели) уравнения регрессии осуществляется с помощью F-критерия следующим образом. На основе заданных уровня значимости α (обычно, $\alpha = 0,01$ или $\alpha = 0,05$), чисел степеней свободы $k_1 = m$ и $k_2 = n - m - 1$ определяется табличное значение $F_\alpha(k_1,\ k_2)$ из специальных таблиц, в частности. Если $F_p > F_\alpha(k_1,\ k_2)$, то модель считается адекватной. В противном случае модель неадекватная, т. е. зависимость между случайными величинами X и Y не является линейной.

Критические значения критерия U-Манна-Уитни
(для проверки ненаправленных альтернатив) $P = 0,05$

N_2	N_1													
	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3	1	2	2	3	3	4	4	5	5	6	6	7	7	8
4	3	4	4	5	6	7	8	9	10	11	11	12	13	13
5	5	6	7	8	9	11	12	13	14	15	17	18	19	20
6	6	8	10	11	13	14	16	17	19	21	22	24	25	27
7	8	10	12	14	16	18	20	22	24	26	28	30	32	34
8	10	13	15	17	19	22	24	26	29	31	34	36	38	41
9	12	15	17	20	23	26	28	31	34	37	39	42	45	48
10	14	17	20	23	26	29	33	36	39	42	45	48	52	55
11	16	19	23	26	30	33	37	40	44	47	51	55	58	62
12	18	22	26	29	33	37	41	45	49	53	57	61	65	69
13	20	24	28	33	37	41	45	50	54	59	63	67	72	76
14	22	26	31	36	40	45	50	55	59	64	67	74	78	83
15	24	29	34	39	44	49	54	59	64	70	75	80	85	90
16	26	31	37	42	47	53	59	64	70	75	81	86	92	98
17	28	34	39	45	51	57	63	67	75	81	87	93	99	105
18	30	36	42	48	55	61	67	74	80	86	93	99	106	112
19	32	38	45	52	58	65	72	78	85	92	99	106	113	119
20	34	41	48	55	62	69	76	83	90	98	105	112	119	127

$P = 0,01$

N_2	N_1													
	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3			0	0	0	1	1	1	2	2	2	2	3	3
4	0	1	1	2	2	3	3	4	5	5	6	6	7	8
5	1	2	3	4	4	6	7	7	8	9	10	11	12	13
6	3	4	5	6	6	9	10	11	12	13	15	16	17	18
7	4	6	7	9	9	12	13	15	16	18	19	21	22	24
8	6	7	9	11	11	15	17	18	20	22	24	26	28	30
9	7	9	11	13	13	18	20	22	24	27	29	31	33	36
10	9	11	13	16	16	21	24	26	29	31	34	37	39	42
11	10	13	16	18	18	24	27	30	33	36	39	42	45	48
12	12	15	18	21	21	27	31	34	37	41	44	47	51	54
13	13	17	20	24	24	31	34	38	42	45	49	53	56	60
14	15	18	22	26	26	34	38	42	46	50	54	58	63	67
15	16	20	24	29	29	37	42	46	51	55	60	64	69	73
16	18	22	27	31	31	41	45	50	55	60	65	70	74:	79
17	19	24	29	34	34	44	49	54	60	65	70	75	81	86
18	21	26	31	37	37	47	53	58	64	70	75	81	87	92
19	22	28	33	39	39	51	56	63	69	74	81	87	93	99
20	24	30	36	42	42	54	60	67	73	79	86	92	99	105

Критические значения критерия G знаков
(для проверки ненаправленных альтернатив)

n	P		n	P		n	P		n	P	
	0,05	0,01		0,05	0,01		0,05	0,01		0,05	0,01
5	0	—	27	8	7	49	18	15	92	37	34
6	0	—	28	8	7	50	18	16	94	38	35
7	0	0	29	9	7	52	19	17	96	39	36
8	1	0	30	10	8	54	20	18	98	40	37
9	1	0	31	10	8	56	21	18	100	41	37
10	1	0	32	10	8	58	22	19	110	45	42
11	2	1	33	11	9	60	23	20	120	50	46
12	2	1	34	11	9	62	24	21	130	55	51
13	3	1	35	12	10	64	24	22	140	59	55
14	3	2	36	12	10	66	25	23	150	64	60
15	3	2	37	13	10	68	26	23	160	69	64
16	4	2	38	13	11	70	27	24	170	73	69
17	4	3	39	13	11	72	28	25	180	78	73
18	5	3	40	14	12	74	29	26	190	83	78
19	5	4	41	14	12	76	30	27	200	87	83
20	5	4	42	15	13	78	31	28	220	97	92
21	6	4	43	15	13	80	32	29	240	106	101
22	6	5	44	16	13	82	-33	30	260	116	110
23	7	5	45	16	14	84	33	30	280	125	120
24	7	5	46	16	14	86	34	31	300	135	129
25	7	6	47	17	15	88	35	32			
26	8	6	48	17	15	90	36	33			

Библиографический список

1. Акулич И. Л. Математические программирование в примерах и задачах. М.: Высш. школа, 1993.

2. Вагнер Г. Основы исследования операций. М.: Мир, 1972.

3. Вентцель Е. С. Исследование операций. М.: Сов. радио, 1972.

4. Вилкас Э. И., Майминес Е. З. Решения: теория, информация, модели. М.: Радио и связь, 1981.

5. Воронов М. В., Мещеряков Г. П. Математика для студентов гуманитарных факультетов. Ростов н/Дону: Феникс, 2002.

6. Исследование операций: Метод. указания, Омск: Изд-во ОмГТУ, 1996.

7. Исследование операций /Под ред. Дж. Моудера, С. Элмаграби. М.: Мир, 1981.

8. Калихман И. Л. Сборник задач по математическому программированию. М.: Высш. школа, 1985.

9. Канторович Л. В., Горстко А. Б. Оптимальные решения в экономике. М.: Наука, 1972.

10. Коршунов Ю. М. Математические основы кибернетики. М.: Энергия, 1990.

11. Кендалл М., Стюарт А. Многомерный статистический анализ и временные ряды. М.: Наука, 1976.

12. Шефе Г. Дисперсионный анализ. М., 1980.

13. Замков О. О., Толстопятенко А. В., Черемных Ю. Н. Математические методы в экономике. М.: Изд-во МГУ им. Ломоносова, ДИС, 1998.

14. Наследов А.Д. Математические методы психологического исследования. СПБ, 2007.

15. Ефимова М.Р., Петрова Е.В. Общая теория статистики. М.: Инфра-М, 2004.

Редактор
ИД____________ от ____________
Подписано в печать . Бумага офсетная. Формат 60x84 1/16
Отпечатано на дупликаторе. Усл. печ. л. . Уч.-изд. л. .
Тираж 50 экз. Заказ .

Издательство ОмГТУ. 644050, г. Омск, пр-т Мира, 11
Типография ОмГТУ