Министерство общего и профессионального образования РФ
Ярославский государственный педагогический университет
имени К.Д. Ушинского
Институт педагогики и психологии
Кафедра общей и социальной психологии

Н.Г. Рукавишникова, Е.Г Заверткина

Статистический анализ данных и способы представления результатов исследования

Учебно-методическое пособие к лекционно-практическим курсам
«Экспериментальная психология» и «Психодиагностика»

Ярославль

2000

ББК

Печатается по решению
редакционно-издательского
совета ЯГПУ им. К.Д. Ушинского

Статистический анализ данных и способы представления результатов исследования: Учебно-методическое пособие к курсам «Экспериментальная психология» и «Психодиагностика» / Е.Г. Заверткина, Н.Г. Рукавишникова. Ярославль: Изд-во ЯГПУ, 2000. 47с.

Учебно-методическое пособие адресовано студентам педагогических вузов.

Рецензент: Хахунова М.Н., зав. Информационно-методическим отделом ГЦПОПП «Ресурс»

Содержание

Введение

Проблема измерения или оценки тех или иных цифр, событий, явлений является одной из фундаментальных в науке.

Измерения в психологии и других социальных науках не являются такими точными, как в естественных науках (математике, физике). Тем не менее, они могут быть очень полезными для обобщения данных о поведении человека и описания факторов, влияющих на него.

В психологии существуют различные методы сбора данных о личности и группе – наблюдение, оценочные листы и шкалы, личностные опросники, проективные техники, тесты и др. Ценность данных, полученных с помощью этих методов, зависит как от степени валидности и надежности самого измерительного инструмента, так и от грамотной обработки и интерпретации результатов проведенного исследования. С этой целью в психологии широко используются различные методы статистической обработки результатов, которые позволяют правильно организовать, структурировать полученные данные и адекватно их интерпретировать. Статистические процедуры (такие как среднее арифметическое (M), среднее квадратическое отклонение (σ), коэффициент корреляции (r) и др.) играют ведущую роль в процессе оценивания и измерения. Они обеспечивают исследователя правилами и приемами обобщения и описания полученных данных, помогают в интерпретации полученных в ходе исследования тестовых оценок.

Математическая обработка результатов исследования дает психологу возможность ответить на ряд вопросов:

2. чем один человек отличается от другого (или от группы лиц) по исследуемой психологической характеристике?
3. чем отличается уровень развития одной психологической характеристики от другой у данной личности?
4. как различаются две группы лиц по какой-либо психологической характеристике?

Ответы на эти вопросы (нахождение интраиндивидуальных и меж индивидуальных различий) могут быть получены в ходе психодиагностического обследования и зависят от правильного проведения этого обследования и зависят от правильного проведения этого обследования, а также грамотной обработки и интерпретации полученных результатов.

Любое психологическое исследование включает в себя определенные этапы. Подготовительный этап (организационный), предполагающий:

изучение состояния изучаемой проблемы, т.е. изучение актуального и жизненного значимого вопроса, который требует решения и на который невозможно найти правильный ответ без проведения специального исследования;

- выдвижение рабочей гипотезы исследования, т.е. четко, логически непротиворечиво сформулированного предположения о том, как в принципе может быть разрешена проблема;

- разработка методики исследования, т.е. подбор тех средств и приемов, которые предназначены для проверки гипотезы исследования.

Эмпирический этап или этап сбора фактических данных, где используются выбранные методы.

Этап обработки полученных данных, этап сравнения данных, полученных в исследовании с данными, имеющимися до проведения исследования с использованием достоверных и точных методов математической статистики, дифференциация полученных данных по группам, построение графиков,

Этап интерпретации полученных данных и формулирование выводов.

Настоящее пособие посвящено вопросам, связанным с этапом обработки данных, получаемых в ходе эмпирического исследования. В нем рассматриваются наиболее часто употребляемые методы статистической обработки результатов психодиагностического исследования, а также возможные способы наглядного представления полученных в ходе исследования данных.

Пособие предназначено для студентов, обучающихся, прежде всего на психологические специальности и изучающих курс «Психодиагностика».

Рассматриваемые в пособии вопросы статистической обработки и представления данных могут быть полезны студентам при написании курсовых и дипломных работ, связанных с проведением психолого-педагогического исследования.

I. Методы первичной статистической обработки данных

Методами статистической обработки результатов эксперимента называются математические приемы, формулы, способы количественных расчетов, с помощью которых показатели, получаемые в ходе эксперимента, можно обобщать, приводить в систему, выявляя скрытые в них закономерности. Главная цель любого статистического метода – представить количественные данные в систематизированной и сжатой форме с тем, чтобы облегчить их понимание.

Некоторые из методов математико-статистического анализа позволяют вычислять так называемые *элементарные математические статистики*, характеризующие *выборочное распределение данных*, например *выборочное среднее* (или среднее арифметическое), *выборочная дисперсия, мода, медиана* и ряд других. Иные методы математической статистики, например *дисперсионный анализ, регрессионный анализ,* позволяют судить о динамике изменения отдельных статистик выборки. С помощью третьей группы методов, например, *корреляционного анализа, факторного анализа, методов* сравнения выборочных данных, можно достоверно судить о статистических связях, существующих между переменными величинами, которые исследуются в данном эксперименте.

Все методы математико-статистического анализа условно делятся на первичные и вторичные.

Первичными называются методы, с помощью которых можно получить показатели, непосредственно отражающие результаты производимых в эксперименте измерений. Соответственно под первичными статистическими показателями имеются ввиду те, которые применяются в самих психодиагностических методиках и являются итогом начальной статистической обработки результатов диагностики. К первичным методам статистической обработки относят, например, определение среднего арифметического, дисперсии, моды и медианы.

Вторичными называются методы статистической обработки, с помощью которых на базе первичных данных выявляют скрытые в них статистические закономерности. В число вторичных методов обычно включают корреляционный анализ, регрессионный анализ, факторный анализ, методы сравнения первичных данных двух или нескольких выборок.

Цель данного раздела – выяснить смысл некоторых традиционно применяемых элементарных статистических мер. Прежде чем переходить к описанию их сущности остановимся на понятиях: *генеральная совокупность* и *выборка*, знание которых необходимо для дальнейшего понимания статистических процедур и закономерностей

1.1 Генеральная совокупность и выборка

Обследование тех или иных объектов может охватывать всех членов изучаемой совокупности без единого исключения или ограничиваться обследованием лишь некоторой части данной совокупности. В первом случае обследование называют *полным* или *сплошным*, во втором - *частичным* или *выборочным*. Полное обследование совокупности позволяет получать исчерпывающую информацию об изучаемом объекте, однако к нему прибегают редко, так как эта работа сопряжена с большими затратами времени и труда, а также в силу практической невозможности. В подавляющем большинстве случаев вместо сплошного изучению подвергают некоторую часть обследуемой совокупности, по которой и судят о ее состоянии в целом.

Совокупность, из которой отбирают определенную часть ее членов для совместного изучения, называют *генеральной*. Отобранная тем или иным способом часть генеральной совокупности получила название *выборочной совокупности* или *выборки*. Общую сумму членов генеральной совокупности называют ее объемом и обозначают буквой N. Теоретически объем генеральной совокупности ничем не ограничен, т.е. генеральную совокупность представляют как бесконечно большое множество относительно однородных единиц или членов, составляющих ее содержание. Объем выборки, обозначаемый буквой n, может быть и большим, и малым, но он не может содержать менее двух единиц. Выборочный метод - основной при изучении статистических закономерностей. Его преимущество перед полным учетом всех членов генеральной совокупности заключается в том, что он сокращает время и затраты труда, а главное - позволяет получать информацию о таких групповых объектах, сплошное обследование которых практически невозможно или нецелесообразно.

Основное требование, предъявляемое к любой выборке, сводится к получению наиболее полной информации о состоянии генеральной совокупности, из которой выборка взята. Опыт показал, что правильно отобранная часть генеральной совокупности, т.е. выборка, довольно хорошо отображает структуру генеральной совокупности. Однако полного совпадения выборочных показателей с характеристиками генеральной совокупности, как правило, не бывает. Чтобы выборка наиболее полно отображала структуру генеральной совокупности, она должна быть достаточно представительной или *репрезентативной,* т.е. она должна отражать основные параметры той генеральной совокупности, которую она представляет. Выборка должна отражать специфику генеральной совокупности людей как по составу (пол, возраст, социально-экономические, географические, этнические, культурные характеристики), так и по индивидуальным характеристикам включенных в нее людей. Причем эти характеристики должны быть представлены в выборке стандартизации примерно в том же процентном соотношении, что и в генеральной совокупности.

Перейдем к описанию основных процедур статистического анализа первичных результатов исследования.

1.2. Частотное распределение

Как правило, в результате эмпирического исследования бывает довольно много исходных первичных данных, которые подлежат статистической обработке. Например, колонка из 1000 тестовых показателей может производить внушительное, даже ошеломляющее впечатление. Но в таком виде она мало о чем говорит. Чтобы «навести порядок в этом хаосе цифр», нужно прежде всего составить таблицу *частотного распределения* (табл.1). Что это такое?

Приведем пример.

По субтесту «Сходство» из батареи Векслера был получен ряд первичных данных. Выпишем их в первый столбик таблицы в порядке убывания (см. табл. 1). Когда показатели распределены по порядку, подсчитывают количество случаев для каждого показателя (т.е. количество испытуемых, получивших данный первичный результат). Полученное таким способом число и есть *частота* (количество случаев) для соответствующего показателя. Сумма всех частот равняется общему числу тестовых показателей (или объему выборки n).

Во второй столбик впишем частоту встречаемости каждого первичного результата (см. табл.1). Мы получили таблицу частотного распределения данных по тесту.

Таблица 1

Частотное распределение результатов
56 студентов по субтесту «Сходство»

Первичный результат	Частота
26	2
25	2
24	8
23	5
22	6
21	7

20	9
19	5
18	4
17	1
16	2
15	3
14	2

Информация, содержащаяся в частотном распределении, может быть также представлена графически в виде *кривой распределения.*

На рис.1 данные табл.1 изображены с помощью графика. По горизонтальной оси отложены значения тестового показателя, а по вертикальной - частота, или число случаев. График может быть построен двумя способами, каждый из которых достаточно распространен (гистограмма и полигон частот). Гистограмма строится из столбцов, вычерченных над каждым значением, высота которых соответствует частоте встречаемости каждого значения.

На полигоне частот число испытуемых указывается точкой, расположенной на высоте, соответствующей частоте встречаемости данного первичного результата, а затем точки последовательно соединяются прямолинейными от

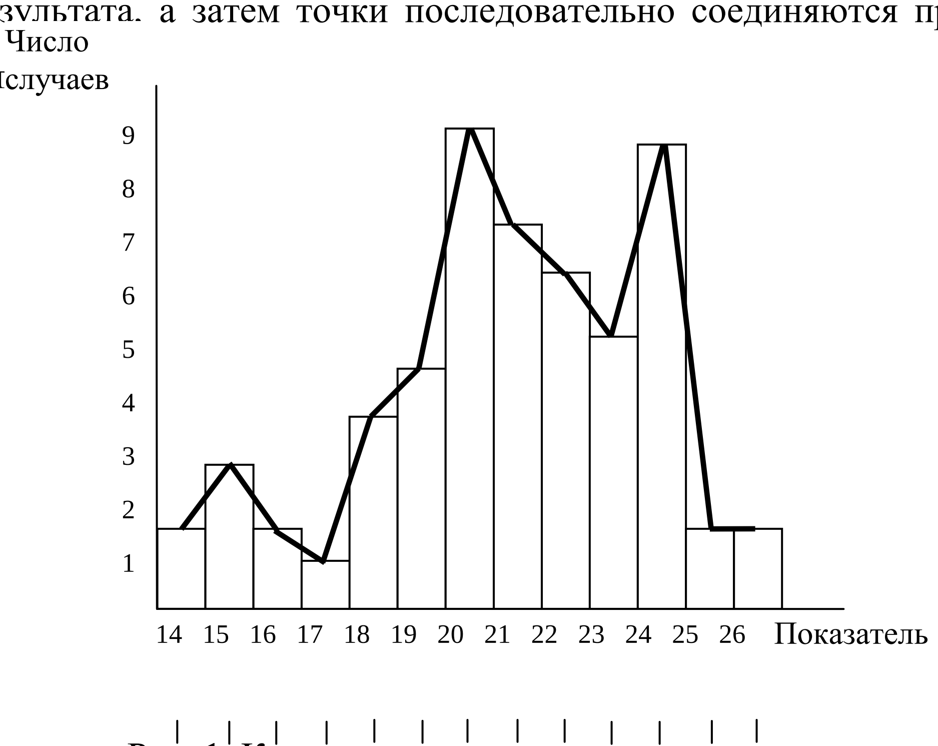


Рис. 1. Кривая полигона частот и гистограмма
(по данным табл. 1)

Для того, чтобы сделать обобщенные данные о характере распределения результатов по тесту и в случае, если получено слишком большое число значений первичного результата, необходимо произвести группировку данных и провести аналогичную процедуру построения частотного распределения. Сгруппируем данные, представленные в табл. 1, объединяя их по 3 единицы в каждой группе (или с интервалом 3 единицы). Группировка данных производится от минимального значения к максимальному. Для того, чтобы интервалы значений были равномерными, добавим еще два значения в нижней части ряда (12 и 13), так как 26 баллов является максимальным для данной методики. Частота встречаемости этих первичных показателей в нашей выборке равна 0. Частотное распределение приобретет следующий вид.

Таблица 2

Интервал значений	Частота
24-26	12
21-23	18
18-20	18
15-17	6
12-14	2

Изобразим его графически.

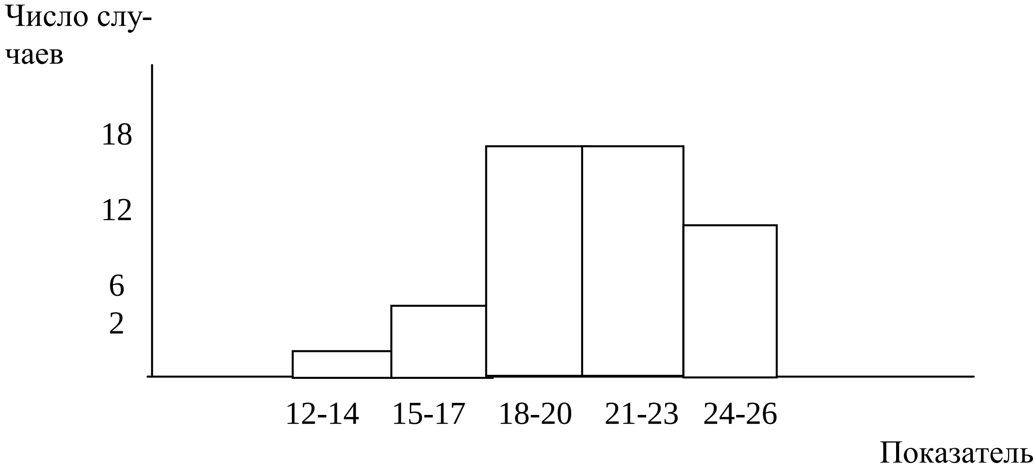


Рис. 2. Кривая полигона частот и гистограмма (по данным табл. 2)

Примечание: на полигоне частот число случаев изображается точкой, расположенной над серединой интервала на высоте, соответствующей его частоте.

Судя по характеру распределения, представленного на рис.2, оно не является нормальным и характеризуется небольшой асимметрией со сдвигом в сторону высоких значений. Это означает, что для данной выборки задания были слишком легкими и большинство испытуемых справились с ними хорошо. Идеальная нормальная кривая изображена на рис. 5, ее свойства описаны в параграфе 1.5.

Этот тип кривой обладает важными математическими свойствами и на ней основаны многие виды статистического анализа. По существу, эта кривая означает, что число случаев максимально в середине распределения и постепенно спадает к ее краям. Кривая симметрична и имеет единственный пик в центре. Большинство распределений численных показателей, от роста и веса до способностей и других характеристик личности, приближаются к нормальной кривой. Можно отметить следующую закономерность: чем больше группа, тем ближе распределение показателей к нормальной кривой.

1.3. Меры центральной тенденции

Рассматривая элементарные методы математической статистики, применяемые для обработки данных тестовых исследований, можно выделить группу методов, которые могут описывать те или иные *меры центральной тенденции*. Такие меры указывают наиболее типичный или репрезентативный результат, характеризующий выполнение теста всей группой. Самой известной из таких мер является **среднеарифметическое значение (М).**

Среднеарифметическое (или выборочное среднее) значение как статистический показатель представляет собой среднюю оценку изучаемого в эксперименте психологического качества. Эта оценка характеризует степень его развития в целом у той группы испытуемых, которая была подвергнута исследованию (*выборки испытуемых*). Сравнивая непосредственно средние значения двух или нескольких групп, мы можем судить об относительной степени развития у людей, составляющих эти группы, оцениваемого качества.

Среднеарифметическое определяется при помощи следующей формулы:

$$M = \frac{\sum_{i=1}^{n} x_i}{n} \quad ,$$

где М – среднеарифметическое значение, n – количество испытуемых, или количество частных показателей; $\sum x_i$ – сумма всех результатов.

Например, допустим, что в результате применения психодиагностической методики для оценки некоторого психологического свойства у десяти испытуемых мы получили следующие частные показатели степени развития данного свойства у отдельных испытуемых: 5, 4, 5, 6, 7, 3, 6, 2, 8, 4.

Следовательно, $M = \frac{5+4+5+6+7+3+6+2+8+4}{10} = \frac{50}{10} = 5$

Для данной выборки среднее арифметическое значение равно 5.

Другой мерой центральной тенденции является **мода** (Мо) или наиболее часто встречающийся результат. В интервальном частотном распределении мода определяется как середина интервала, для которого частота максимальна.

Третья мера центральной тенденции – это **медиана** (Me), т.е. результат, находящийся в середине последовательности показателей, если их расположить в порядке возрастания или убывания. Справа и слева от медианы в упорядоченном ряду остается по одинаковому количеству данных (50% и 50%). Например, для выборки 2,3,4,4,5,6,8,7,9 медианой будет значение 5, т.к. слева и справа от него остается по четыре показателя. Если ряд включает в себя четное число признаков , то медианой будет среднее, взятое как полусумма двух центральных значений ряда. Для следующего ряда 0,1,1,2,3,4,5,5,6,7 медиана будет равна 3,5.

Знание медианы и моды полезно для того, чтобы установить, является ли распределение частных значений изучаемого признака симметричным и приближающимся к нормальному распределению. Среднее арифметическое, медиана и мода для нормального распределения обычно совпадают или очень мало отличаются друг от друга. Для других типов распределений это не характерно. Если выборочное распределение признаков близко к нормальному, то к нему можно применять методы вторичных статистических расчетов, основанные на нормальном распределении данных. В противном случае этого делать нельзя, т.к. в расчеты могут вкрасться серьезные ошибки.

Приведем пример расчета мер центральной тенденции. Для этой цели возьмем следующее распределение первичных результатов теста «Числовые ряды», полученное на выборке из 20 студентов (см. табл. 3).

Таблица 3

Первичный Результат	Частота
20	1
18	1
17	2
16	1
15	3

14	2
13	1
12	2
11	2
10	2
9	1
8	1
5	1

2. Подсчитаем среднее арифметическое:

$$M = \frac{5+8+9+10\times 2+11\times 2+12\times 2+13+14\times 2+15\times 3+16+17\times 2+18+20}{20} = 13{,}1$$

2. Мода (Мо) – наиболее часто встречающаяся оценка. Из таблицы частотного распределения видно, что наибольшей частотой обладает оценка в 15 баллов, следовательно Мо = 15.

3. Медиана (Ме) – точка, которая делит распределение ответов ровно пополам.

В нашем примере серединой распределения является промежуток между оценками 13 и 14, т.к. 10 человек (или 50% выборки) имеет оценки меньше или равные 13 баллам и 10 человек (или 50% выборки) имеют оценки выше или равные 14 баллам.

В этом случае за медиану принимают оценку равную среднему арифметическому 13 и 14 баллов или

$$Me = \frac{13+14}{2} = 13{,}5$$

Таким образом, в нашем примере М=13,1; Мо=15; Ме=13,5 и распределение первичных результатов стремится к нормальному.

Приблизительно судить о том, является или не является полученное распределение близким к нормальному, можно, построив график распределения данных.

Если график оказывается более или менее симметричным, значит, к анализу данных можно применять статистические процедуры, предназначенные для нормального распределения.

Приблизительные картины симметричного и несимметричного распределений признаков показаны на рис.3, где точками Me_1 и Me_2 обозначены те величины признаков, которые соответствуют медианам, M_1 и M_2 - те, которые соответствуют среднеарифметическим значениям, а Mo_1 и Mo_2 – модам. Здесь мы видим, что для симметричных распределений признаков, в том числе для нормального распределения, значения моды, медианы и среднего совпадают. Для других типов распределений, несимметричных, это не характерно.

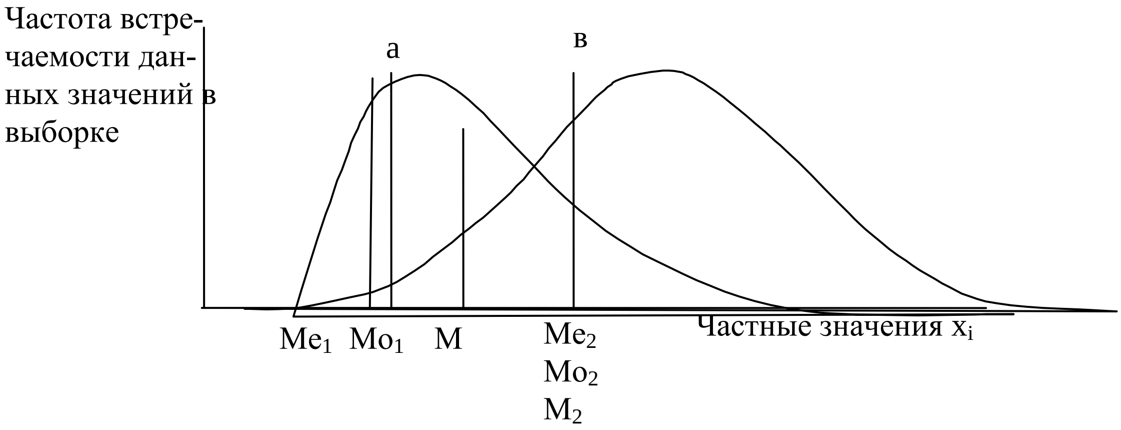


Рис.3. Примеры симметричного (а) и несимметричного (в) распределения первичных результатов

1.4. Меры разброса данных

Для более полного описания результатов эмпирического исследования используются *меры разброса* (или вариативности) данных, характеризующие степень индивидуальных отклонений от центральной тенденции. Наиболее наглядным и известным способом представления разброса является *размах распределения*, т.е. разность между самым высоким и самым низким результатом. Но эта мера крайне неточна и неустойчива, т.к. она определяется только двумя показателями, и единственный необычайно высокий или низкий результат может заметно повлиять на величину размаха. Более точный метод измерения разброса данных основан на *учете разности* между каждым индивидуальным результатом и среднеарифметическим значением по группе. Такой мерой разброса является *дисперсия* или средний квадрат отклонения (σ^2).

Дисперсия как статистическая величина характеризует, насколько частные значения отклоняются от средней величины в данной выборке. Чем больше дисперсия, тем больше отклонение или разброс данных.

Прежде чем представить формулу для расчетов дисперсии, рассмотрим пример. Воспользуемся теми же данными, которые были использованы ранее для вычисления средней величины (5, 4, 5, 6, 7, 3, 6, 2, 8, 4). Мы видим, что все они разные и отличаются друг от друга и от средней величины (M = 5). Меру их общего отличия от средней величины и характеризует дисперсия. Ее определяют для того, чтобы можно было отличать друг от друга выборки, имеющие одинаковую среднюю, но разный разброс. Представим себе другую, отличную от предыдущей выборку первичных значений, например: 5, 4, 5, 6, 5, 4, 5, 5. Легко убедиться в том, что ее средняя величина также равна 5. Но в данной выборке ее отдельные частные значения отличаются от средней гораздо меньше, чем в первой выборке. Можно выразить степень этого отличия при помощи дисперсии, которая определяется по следующей формуле:

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n},$$

где σ^2 - дисперсия; $\sum(x_i - \bar{x})^2$ – выражение, означающее, что для всех значений x от первого до последнего в данной выборке вычисляется разность между частными и средними значениями, эти разности возводятся в квадрат и суммируются; n - количество испытуемых в выборке или первичных значений, по которым вычисляется дисперсия.

Определим дисперсию для двух приведенных выше выборок частных значений.

Мы видим, что дисперсия по второй выборке (0,4) значительно меньше дисперсии по первой выборке (3), хотя среднеарифметические значения у выборок равны. Если бы не было дисперсии, то мы не в состоянии были бы различить данные выборки.

Очень часто вместо дисперсии для выявления разброса частных данных относительно средней используют производную от дисперсии величину, называемую *стандартное (или выборочное) отклонение*. Оно равно квадратному корню, извлекаемому из дисперсии, и обозначается тем же знаком, только без квадрата (σ).Эта величина в ряде случаев оказывается более удобной характеристикой варьирования, чем дисперсия, так как выражается в тех же единицах, что и средняя арифметическая величина.

Пример: Для расчета стандартного отклонения (σ) воспользуемся результатами теста на исследование зрительной памяти на выборке из 10 студентов (см. табл. 4).

Для подсчета σ необходимо найти разности первичных оценок и среднего арифметического по выборке. Представим полученные данные в виде таблицы.

Таблица 4

Первичный результат (x_i)	$x_i – M$	$(x_i - M)^2$
6	- 2,6	6,76
7	- 1,6	2,56
7	- 1,6	2,56
8	- 0,6	0,36
8	- 0,6	0,36
8	- 0,6	0,36
10	1,4	1,96
10	1,4	1,96
11	2,4	5,76
12	3,4	11,56
M = 8, 6		$S (x_i - M)^2 = 34,2$

Далее разности возводятся в квадрат и суммируются. Мы получили S $(x_i - M)^2$, теперь разделим ее на объем выборки (n=10) и извлечем квадратный корень. В нашем примере $\sigma = 1{,}85$.

Стандартное отклонение (σ) как мера разброса результатов теста широко применяется при сравнении разбросов в различных группах. На рис.4, например, представлены два распределения, имеющее одно и тоже среднее значение, но отличающееся разбросом. Распределение, характеризуемое большими индивидуальными различиями, в отличие от распределений с различиями меньшими, имеет большее стандартное отклонение.

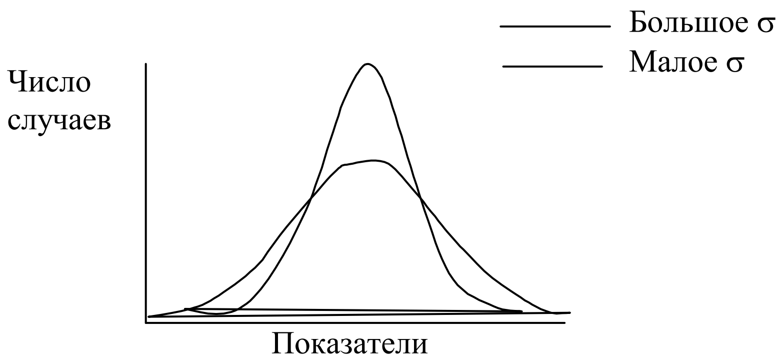


Рис. 4. Примеры распределений, отличающихся разбросом

Особенно четкой оказывается интерпретация σ применительно к нормальной или к приблизительно нормальной кривой распределения, т.к. здесь существует прямое соответствие между σ и относительным количеством случаев. На рис.5 по горизонтальной оси отложены интервалы, соответствующее отклонению в 1σ, 2σ и 3σ вправо и влево от среднего значения (М). Процент случаев, приходящихся на интервал М $+1\sigma$ в нормальном распределении равен 34,13. Поскольку кривая симметрична, 34,13% случаев приходится также на интервалы от М до -1σ, так что диапазон от -1σ до

+1σ охватывает 68,26% случаев. Почти все случаи (99,72%), то есть почти все показатели лежат от -3σ до +3σ относительно среднего значения.

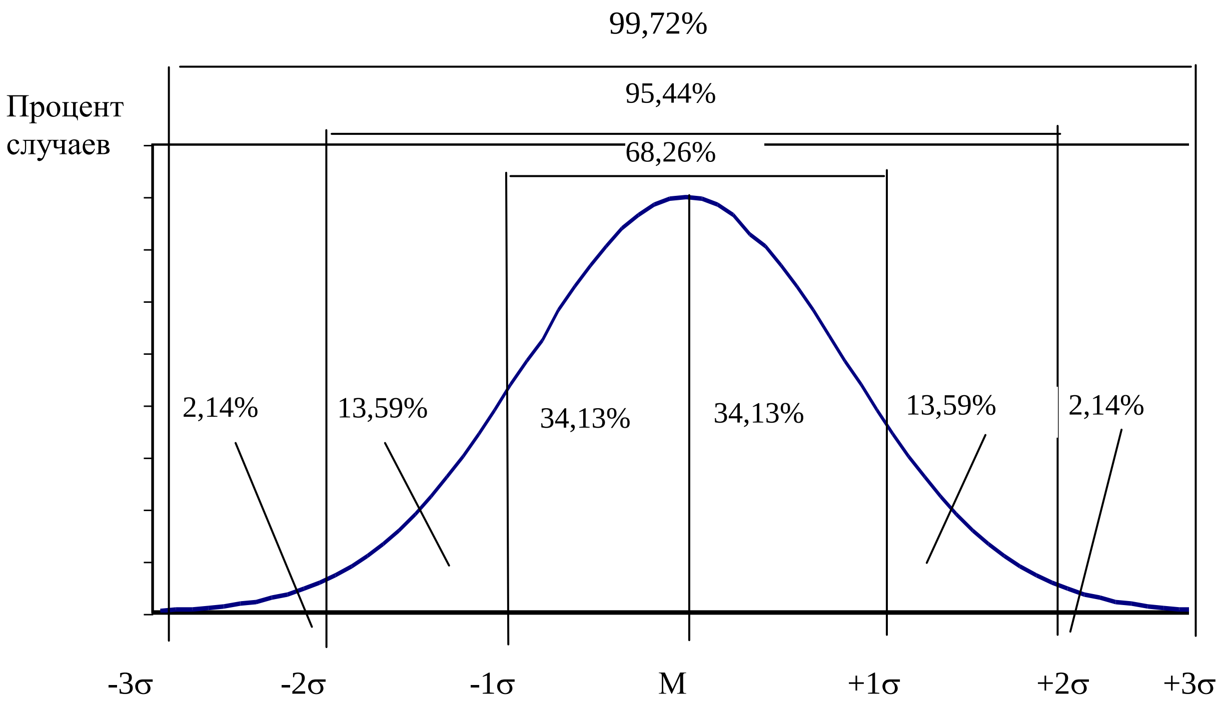


Рис. 5. Процентное соотношение случаев для кривой нормального распределения.

Эта закономерность известна как «закон трех сигм», и является одной из характеристик нормального распределения. Эти соотношения имеют особое значение для стандартизации или нормирования первичных показателей, получаемых в ходе исследования.

1.5. Нормирование (стандартизация) данных

Большинство психологических тестов дает непосредственные числовые сведения об исполнении их человеком, называемые сырой оценкой. Ею может быть количество вопросов, на которые были даны ответы, время, необходимое для выполнения задания, или другие сведения. Взятая сама по себе сырая оценка по психологическому тесту не имеет значения. Ее можно интерпретировать только сопоставив с каким-нибудь стандартом. На практике ее обычно сопоставляют с нормами выполнения данного теста в выборке стандартизации. Соотнесение первичного результата (или сырой оценки) с распределением результатов, полученных в выборке стандартизации показывает, какое место он занимает в этом распределении.

Чтобы определить более точно положение результатов конкретного испытуемого, полученный результат переводится в некую относительную меру (стандартизованную оценку). Таким образом преобразованные оценки называют тестовыми нормами. Тестовые нормы представляют собой эмпирически установленные на базе репрезентативной выборки усредненные количественные данные о результатах выполнения теста, полученные в определенных условиях.

Тестовые нормы позволяют: 1) устанавливать для данного теста место каждого испытуемого по отношению к другим в составе генеральной совокупности; 2) устанавливать статистически значимые сходство или различие между группами (и между испытуемыми) по одному и по различным тестам.

Пользователь теста всегда должен знать способ, которым устанавливались нормы теста. Применительно к психологическим тестам они не абсолютны, не универсальны и не постоянны. Они отражают лишь выполнение теста испытуемыми из выборки стандартизации. Таким образом, нормы устанавливаются эмпирически, соответственно тому, как выполняет задания теста репрезентативная группа испытуемых. После этого сырой балл испытуемого соотносится с нормами, полученными на выборке стандартизации и делается вывод об уровне развития исследуемой психологической переменной (соответствует ли результат среднему выполнению в нормативной группе, ниже среднего или значительно выше среднего).

Большинство распределений первичных результатов исследования ближе к нормальному распределению, чем к какому-либо иному. Нормальный закон распределения во всех естественных науках имеет фундаментальное значение. И в психологии его значение трудно переоценить. Достаточно сказать, что все психологические шкалы основываются на этом законе, поскольку ему следуют распределения большинства человеческих способностей и свойств. Большинство статистических процедур, применяемых в психологии построено на гипотезе нормального распределения данных. Поэтому на практике прежде чем перейти к построению норм проверяется предположение, что анализируемое распределение соответствует нормальному (т.е. близость свойств эмпирического распределения со свойствами нормального распределения). Каковы же **свойства нормального распределения**?

1. В нормальном распределении все меры центральной тенденции равны между собой, т.е. сходятся в одной точке на графике (M = Me = Mo).

2. В нормальном распределении примерно 99% всех значений исследуемой переменной находится в пределе M+3σ. Процентное распределение случаев на кривой нормального распределения в психологии называют “законом трех сигм” (см. рис. 5).

3. Кривая нормального распределения имеет вид колокола, она симметрична (асимметрия отсутствует, As = 0) и не имеет слишком острой или слишком плоской вершины (эксцесс отсутствует, Ex = 0).

На рис. 5 изображена стандартная кривая нормального распределения, которая наглядно отображает все вышеизложенные свойства.

Существуют специальные формулы и компьютерные программы проверки распределения на его близость к нормальному, но для небольших практических исследований чаще всего бывает достаточно проверить свойства полученного эмпирического распределения по этим трем критериям. Если различия незначительны, то делается вывод о том, что распределение близко к нормальному. Это означает, что полученное распределение может рассматриваться как устойчивое - репрезентативное по отношению к генеральной совокупности и, следовательно, на его основе можно построить репрезентативные тестовые нормы. Если проверка не выявила нормальности на требуемом уровне, то это означает, что либо выборка мала и нерепрезентативна по отношению к популяции, либо измеряемые свойства и сам тест в силу тех или иных причин вообще не дают нормального распределения.

1.6. Типы норм или стандартных оценок.

Описание почти всех современных стандартизованных тестов в той или иной форме содержат данные о внутригрупповых нормах. С помощью таких норм индивидуальное выполнение оценивается в соответствии с выполнением наиболее сопоставимой выборки стандартизации. Например, первичный показатель ребенка сравнивается с показателями детей того же возраста и того же года обучения.

Одной из наиболее простых для подсчета и распространенных в психодиагностике форм стандартных оценок является процентиль.

Процентиль - это процентная доля индивидов из выборки стандартизации, первичный результат которых равен данному или ниже его. Например, если 28% людей из выборки стандартизации правильно решают 15 задач в тесте на исследование математических способностей, то первичному показателю 15 соответствует 28-й процентиль. Процентили указывают на относительное положение индивида в выборке стандартизации. Чем ниже процентиль, тем хуже позиция индивида, и наоборот. 50-й процентиль соответствует медиане. Как рассчитываются процентили? Возьмем для примера первичные результаты по тесту “числовые ряды”, представленные в виде частотного распределения (см. Табл. 5).

Таблица 5

Первичный результат	Частота	Накопленная частота	Процентиль
20	1	20	100
18	1	19	95
17	2	18	90
16	1	16	80

15	3	15	75
14	2	12	60
13	1	10	50
12	2	9	45
11	2	7	35
10	2	5	25
9	1	3	15
8	1	2	10
5	1	1	5

В первом столбике таблицы расположены первичные результаты исследования в порядке убывания, во втором - частота встречаемости каждого результата в выборке объемом 20 человек (n = 20). Третий столбик представляет собой сумму числа испытуемых, имеющих первичный результат равный или ниже данного (суммирование частот производится снизу вверх, от минимального результата к максимальному). И, наконец, в последнем столбике даны процентильные ранги каждого первичного результата, которые рассчитываются следующим образом: накопленная частота умножается на 100 и делится на объем выборки (в нашем примере n = 20).

Процентильные показатели легко рассчитать и понять даже сравнительно неподготовленному человеку. Их применение достаточно универсально, они подходят к любому типу теста, измеряет он способности или свойства личности; их применение целесообразно в том случае, если тип распределения неизвестен.

Большинство стандартных показателей получаются путем линейного и нелинейного преобразования первичных показателей. Если распределение близко к нормальному, используется линейное преобразование, при этом сохраняются соотношения между первичными показателями, поскольку они вычисляются вычитанием из каждого первичного показателя одной и той же величины с последующим делением результата на другую постоянную величину. При этом относительная величина разницы между стандартными показателями в точности соответствует относительной величине различий между первичными показателями. Линейно преобразованный стандартный показатель обычно называют просто z - показатель. Чтобы вычислить z, находят разность между первичным результатом индивида и средним арифметическим для нормативной группы, а затем делят эту разность на среднее квадратическое отклонение (σ) нормативной группы (см. формулу и пример ниже).

Таблица 6

Вычисление значений стандартных показателей

$Z = \dfrac{X - M}{\sigma}$	M = 60	σ = 5
Результат Саши		Результат Толи
$X_1 = 65$		$X_2 = 58$
$Z_1 = \dfrac{65-60}{5} = +1.0$		$Z_2 = \dfrac{58-60}{5} = -0.4$

Всякий первичный показатель равный среднему арифметическиму, имеет z = 0. Отрицательные значения стандартного показателя означают, что выполнение теста индивидом было ниже среднего, положительные - выше среднего. В целом значения z колеблются в пределах 3σ выше и ниже среднего.

При желании в качестве M и σ можно выбрать любые другие удобные значения, например, показатели отдельных субтестов в тесте интеллекта Векслера преобразуются так, что M = 10, а σ = 3.

От стандартной шкалы z легко осуществить переход к любой другой, более удобной в обращении шкале. В практике психологического тестирования часто используется ряд нормализованных шкал:

1. Шкала интеллекта Векслера: $IQ = 100 + 15 \cdot z$;

2. Шкала Т-баллов (MMPI): $T = 50 + 10 \cdot z$;

3. Шкала стенов Кеттелла: $ST = 5,5 + 2 \cdot z$.

Взаимосвязи между разными типами стандартизованных шкал на основе кривой нормального распределения приведены на рис. 6.

Линейно преобразованные стандартные показатели сопоставимы именно потому, что их исходные распределения имеют приблизительно одну и ту же форму (близкую к нормальному распределению). Если распределение отличается от нормального, оно может быть нормализовано искусственно. В результате такого нелинейного преобразования получают нормализованные стандартные показатели. Их значения могут быть найдены с помощью таблиц.

Нормализация может производиться с использованием компьютерных программ, однако преобразование распределения следует проводить при наличии большой и репрезентативной выборки, когда есть основание считать, что отклонение распределения от нормального произошло в силу определенных дефектов теста, а не особенностей выборки или действия других факторов, влияющих на исследуемую функцию.

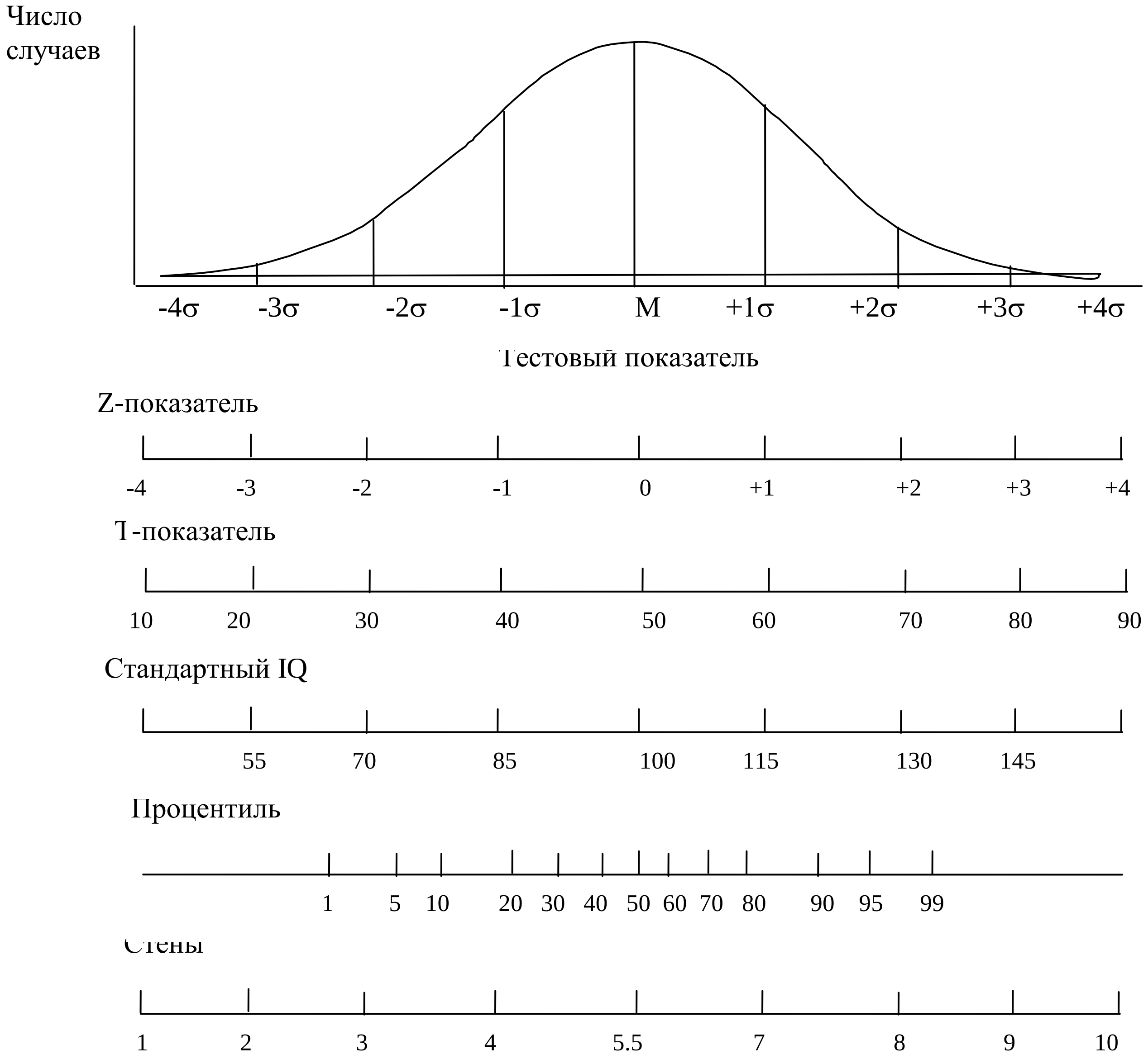


Рис. 6. Соотношения между разными типами стандартизованных оценок

II. Методы вторичной статистической обработки результатов

С помощью вторичных методов статистической обработки данных непосредственно проверяются, доказываются или опровергаются гипотезы, связанные с эмпирическим исследованием. Эти методы, как правило, сложнее, чем методы первичной обработки, и требуют от исследователя хорошей подготовки в области элементарной математической статистики.

Чаще всего в прикладных психолого-педагогических исследованиях применяют следующие методы вторичной статистической обработки результатов:

1. методы сравнения двух или нескольких элементарных статистик (средних, дисперсий и т.п.), относящихся к разным выборкам;

2. методы установления статистических связей между переменными (например, их корреляции друг с другом);

3. методы выявления внутренней статистической структуры эмпирических данных (например, факторный анализ).

Рассмотрим подробнее каждую из перечисленных выше групп методов вторичной статистической обработки на примерах.

2.1. Методы сравнения элементарных статистик

Ни одно психологическое исследование не обходится без сравнений. Сравнивать приходится результаты, полученные двумя разными группами испытуемых или результаты, полученные одной выборкой испытуемых, но в разное время или в разных условиях. О различиях между ними судят обычно по разности между средними, дисперсиями и другими выборочными показателями.

В статистике широкое применение получила так называемая *нулевая гипотеза* (H_o). Сущность ее сводится к предположению, что разница между генеральными параметрами сравниваемых групп равна нулю и что различия, наблюдаемые между выборочными характеристиками носят не систематический, а исключительно случайный характер. Для проверки данной гипотезы используют специальные критерии достоверности. В практических исследованиях по педагогике и психологии чаще всего применяют два вида статистических критериев: t - *критерий* Стьюдента и F - *критерий* Фишера (при условии нормального распределения изучаемой переменной). Первый используют для сравнительной оценки средних величин, второй - для сравнительной оценки дисперсий. Ниже рассмотрен отдельно каждый из этих критериев.

2.1.1. t - критерий Стьюдента

Для сравнения выборочных средних величин, принадлежащим к двум совокупностям данных, и для решения вопроса о том, отличаются ли средние значения статистически достоверно друг от друга, используют t – критерий Стьюдента. Это параметрический метод, используемый для проверки гипотез о достоверности разности средних при анализе количественных данных. Метод Стьюдента различен для независимых и зависимых выборок. Независимые выборки получаются при исследовании двух различных групп испытуемых. В случае независимых выборок для анализа разницы средних применяют формулу:

$$t = \frac{\left|\bar{x}_1 - \bar{x}_2\right|}{\sqrt{\left|m_1^2 + m_2^2\right|}} ,$$

где $\bar{x}_1$ - среднее значение переменной по одной выборке данных, $\bar{x}_2$ - среднее значение переменной по другой выборке данных, m_1 и m_2– интегрированные показатели отклонений частотных значений из двух сравниваемых выборок от соответствующих им средних величин, m_1 и m_2 вычисляются по формулам:

$$m_1^2 = \frac{\bar{S}_1^2}{n_1}; \qquad m_2^2 = \frac{\bar{S}_2^2}{n_2} ,$$

где $\bar{S}_1^2$ и $\bar{S}_2^2$ – выборочные дисперсии переменных по первой и второй выборкам, n_1– число частных значений переменной в первой выборке, n_2– число частных значений переменной по второй выборке.

После того как при помощи формулы вычислен показатель t, по специальной таблице для заданного числа степеней свободы, равного n_1+n_2-2 и избранной вероятности допустимой ошибки (см. таблицу 1 в Приложении), находят нужное табличное значение t и сравнивают с ним вычисленное значение t. Если вычисленное значение t больше или равно табличному, то делают вывод о том, что сравниваемые средние значения из двух выборок действительно статистически достоверно различаются с принятой вероятностью допустимой ошибки, т.е. нулевая гипотеза не верна.

Рассмотрим процедуру вычисления t – критерия Стьюдента и определения на его основе разницы в средних величинах на конкретном примере. Допустим, что имеются две выборки данных:2, 4, 5, 3, 2, 1, 3, 2, 6, 4, (n_1 =10) и 4, 5, 6, 4, 4, 3, 5, 2, 2, 7 (n_2 =10). Это данные изучения слуховой вербальной памяти девушек и юношей 12-14 лет. Результаты показывают количество проб, потребовавшееся испытуемым для запоминания 12 пар слов, не связанных по смыслу. Средние значения по этим выборкам соответственно равны: $\bar{x}_1$ =3,2; $\bar{x}_2$ = 4,2. Кажется, что они существенно друг от друга отличаются. Но так ли это и насколько статистически достоверны эти различия? На данный вопрос может точно ответить статистический анализ с использованием t-критерия Стьюдента.

Определяем выборочные дисперсии сравниваемых выборок значений:

$$\bar{S}_1^2 = \frac{1}{10}\sum_{k=1}^{10}\left(x_{k1} - \bar{x}_1\right)^2 = 2.2; \qquad \bar{S}_2^2 = \frac{1}{10}\sum_{k=1}^{10}\left(x_{k2} - \bar{x}_2\right)^2 = 2.4;$$

Вычисляем t по формуле:
$$t = \frac{|3.2 - 4.2|}{\sqrt{\frac{2.2}{10} + \frac{2.4}{10}}} = 1.47.$$

Сравним его значение с табличным для числа степеней свободы: 10+10-2=18. Зададим вероятность допустимой ошибки, равной 0.05, и убедимся, что для данного числа степеней свободы и заданной вероятности допустимой ошибки значение t должно быть не меньше чем 2.10. У нас же этот показатель равен 1.47, т.е. меньше табличного. Следовательно, гипотеза о том, что средние этих выборок (3.2 и 4.2) статистически достоверно отличаются друг от друга, не подтвердилось, хотя на первый взгляд казалось, что такие различия существуют.

К зависимым выборкам относятся, например, результаты одной и той же группы испытуемых до и после воздействия независимой переменной. Данный метод сравнения средних величин применяется тогда, когда необходимо ,например, установить, удался или не удался эксперимент, оказал или не оказал он влияние на уровень развития того психического качества, для изменения которого предназначался. Допустим, вводится экспериментальная методика обучения или развивающая программа, рассчитанная на то, чтобы улучшить знания учащихся, повысить уровень их интеллектуального развития. В этом случае выясняется причинно-следственная связь между независимой переменной- программой или методикой, и зависимой переменной – знаниями или уровнем интеллектуального развития. В исследовании предполагается, что введение новой программы или методики обучения должно будет существенно улучшить знания и повысить уровень интеллектуального развития учащихся.

Оцениваются зависимые переменные в начале и в конце экспериментального исследования. Получив такие оценки и вычислив среднее по всей изученной выборке испытуемых, мы можем воспользоваться критерием Стьюдента для точного установления наличия или отсутствия статистически достоверных различий между средними до и после эксперимента. Если окажется, что они действительно достоверно различаются, то можно будет сделать определенный вывод о том, что эксперимент удался. В противном случае нет убедительных оснований для такого вывода даже в том случае, если сами средние величины в начале и в конце эксперимента по своим абсолютным значениям различны.

Для определения достоверности разницы средних в случае зависимых выборок применяется следующая формула:

$$t = \frac{\sum d}{\sqrt{\frac{n\sum d^2 - \left(\sum d\right)^2}{n-1}}},$$

где d – разность между результатами в каждой паре; Σd – сумма этих частных разностей; Σd^2 – сумма квадратов частных разностей.

Полученные результаты сверяют с таблицей t, отыскивая в ней значения, соответствующие объему выборки (n –1); n- это в данном случае число пар данных).

Разберем конкретный пример. У нас имеются данные по изучению вербального мышления школьников 12-14 лет (субтест «Сходство» из батарии Векслера) до и после применения развивающей методики. Необходимо определить является ли статистически достоверной разность результатов ,полученных в ходе изучения вербального мышления до и после использования развивающей программы. Для этого можно использовать t – критерий Стьюдента. Перед тем как использовать формулу, необходимо вычислить для каждой группы частные разности между результатами во всех парах, квадрат каждой из этих разностей, сумму этих разностей и сумму их квадратов.

Необходимо произвести следующие операции (см. таблицу 7):

Таблица 7.

Испытуемые	До применения развивающей программы	После применения развивающей программы	d	d^2
Д 1	19	21	+2	4
2	10	8	-2	4
3	12	13	+1	1
4	13	11	-2	4
5	17	20	+3	9
6	14	12	-2	4
7	17	15	-2	4
Ю 1	15	17	+2	4
2	14	15	+1	1
3	15	15	0	0
4	17	18	+1	1
5	15	16	+1	1
6	18	15	-3	9
7	19	19	0	0

8	22	25	+3	9
			$\sum d = 3$	$\sum d^2 = 55$

$$t = \frac{+3}{\sqrt{\frac{15 \times 55 - 3^2}{15 - 1}}} = \frac{+3}{\sqrt{\frac{825 - 9}{14}}} = \frac{+3}{\sqrt{58.28}} = 0.39$$

Величина t=0,39 ниже табличного значения, соответствующего уровню уровню значимости 0,05 при объеме выборки 14.Таким образом разница между выборками недостоверна.

2.1.2. F - критерий Фишера

Иногда в психологическом экспериментальном исследовании возникает необходимость сравнить дисперсии двух выборок для того, чтобы решить, различаются ли они между собой. Допустим, что проводится эксперимент, в котором проверяется гипотеза о том, что одна из двух предлагаемых программ обучения обеспечивает одинаково успешное усвоение знаний учащимися с разными способностями, а другая программа этим свойством не обладает. С определенной долей вероятности можно говорить о том, что демонстрацией справедливости такой гипотезы было бы доказательство того, что индивидуальный разброс оценок учащихся по одной программе больше, чем индивидуальный разброс оценок по другой программе.

Подобного рода задачи решаются, в частности, при помощи *критерия Фишера*. Его формула выглядит следующим образом:

$$F(n_1 - 1, n_2 - 1) = \frac{S_1^2}{S_2^2},$$

где n_1 – количество значения признака в первой выборке;

n_2 – количество значений признака во второй выборке;

(n_1-1, n_2-1) – число степеней свободы;

S_1^2 и S_2^2 - дисперсии по первой и второй выборке.

Вычисленное с помощью формулы значение F-критерия сравнивается с табличным (см. таблицу 4 в Приложении), и если оно превосходит табличное для избранной вероятности допустимой ошибки и заданного числа степеней свободы, то делается вывод о том, что гипотеза о различиях в дисперсиях подтверждается.

Например, сравним дисперсии следующих двух рядов цифр с целью определения статистически достоверных различий между ними: 1 ряд: 4, 6, 5, 7, 3, 4, 5, 6 (n_1 = 8); 2 ряд: 2, 7, 3, 6, 1, 8, 4, 5 (n_2 = 8). Соответственно: $\bar{x}_1$=5.0, $\bar{x}_2$=4.5.

$$\sigma_1^2=1.5,\ \sigma_2^2=5.25.$$

Частное от деления большей дисперсии на меньшую равно 3.5. Это и есть критерий Фишера (F). Сравнивая его с табличным значением 3.44, приходим к выводу о том, что дисперсии двух сопоставляемых выборок действительно отличаются друг от друга на уровне значимости более 95% или с вероятностью допустимой ошибки не более 0.05%.

2.2. Корреляционный анализ

Еще Гиппократ в VI в. до н.э. обратил внимание на наличие связи между телосложением и темпераментом людей, между строением тела и предрасположенностью к тем или иным заболеваниям. Определенные виды подобной связи выявлены также в животном и растительном мире. Известно, например, что между ростом и массой тела у человека существует положительная связь: более высокие индивиды имеют обычно и большую массу тела, чем индивиды низкого роста. Однако существуют и исключения из общей закономерности, т.к. каждый биологический признак представляет собой функцию многих переменных: на него влияют и генетические, и средовые факторы, что и обусловливает варьирование признаков. Такого рода зависимость между переменными величинами называется *корреляционной* или *корреляцией*. Переменная - любая величина, которая может быть измерена и чье количественное выражение может варьировать в пределах того или иного континуума. Например, тревожность - переменная, т.к. ее можно измерить, а также потому, что люди различаются по степени выраженности у них тревожности. В корреляционном исследовании переменными могут быть данные тестирования, демографические характеристики, личностные черты, мотивы, ценности, физиологические реакции. Задача корреляционного анализа - дать количественную оценку соответствия значений одной переменной значениям другой переменной. Для решения этой задачи вычисляют статистический индекс, называемый *коэффициентом корреляции.* Коэффициент корреляции (обозначается маленькой буквой r) показывает нам две вещи: 1) степень связи двух переменных и 2) направление этой связи (прямая или обратная связь).

Связь между переменными величинами X и Y можно установить, сопоставляя числовые значения одной из них с соответствующими значениями другой. Если при увеличении одной переменной увеличивается другая, это указывает на *положительную связь* между этими величинами, и, наоборот, когда увеличение одной переменной сопровождается уменьшением значений другой, это указывает на *отрицательную связь*. Мерой связи между двумя переменными и является коэффициент корреляции r.

Коэффициент корреляции - это отвлеченное число, лежащее в пределах от -1 (полностью отрицательная, или обратная связь) через 0 (отсутствие связи) до +1 (полностью положительная, или прямая связь). Коэффициент корреляции близкий по значению к нулю, означает, что связь между двумя переменными полностью отсутствует, $r_{xy} = 0$. В качестве примера можно привести связь между следующими двумя переменными: массой тела и интеллектом.

При *положительной или прямой связи*, когда большим значениям одной переменной соответствуют большие же значения другой, коэффициент корреляции имеет положительный знак и находится в пределах от 0 до +1. Например, существует положительная корреляция между ростом и массой тела людей. Другой пример положительной корреляции - связь между количеством сцен насилия, которые дети видят в телевизионных передачах и их тенденцией вести себя агрессивно. В среднем, чем чаще дети наблюдают насилие по телевизору, тем чаще они демонстрируют агрессивное поведение.

Отрицательная или обратная связь означает, что высоким значениям одной переменной соответствуют низкие значения другой переменной и наоборот. При этом коэффициент корреляции имеет отрицательный знак и находится в пределах от 0 до -1. Примером отрицательной корреляции может служить связь между частотой отсутствия студентов в аудитории и успешностью сдачи ими экзаменов. В целом, студенты, имевшие большее количество пропущенных занятий, проявляют тенденцию к получению более низких оценок на экзаменах. Другой пример - отрицательная корреляция между робостью и напористым поведением.

В целом, чем ближе значения коэффициента корреляции к +1 или -1, тем сильнее связь между двумя изучаемыми переменными. Так, коэффициент корреляции +0.8 отражает наличие более сильной связи между двумя переменными, чем коэффициент корреляции +0.3, а коэффициент корреляции -0.65 отражает более сильную связь между двумя переменными, чем коэффициент корреляции –0.25. Величина корреляции (или сила связи) между переменными зависит только от числового значения коэффициента, в то время как знак (+ или -), обозначает направление связи между ними (положительная или отрицательная связь).

Коэффициент *линейной корреляции* определяется по формуле:

$$r_{xy} = \frac{\sum_{i=1}^{n}\left[(x_i - \bar{x})(y_i - \bar{y})\right]}{n - \sqrt{\bar{S}_x^2 \cdot \bar{S}_y^2}}$$

Таким образом, коэффициент корреляции характеризует как силу, так и направление связи между двумя переменными.

Можно выделить несколько видов корреляционного анализа: линейный, ранговый, парный, множественный.

Линейный корреляционный анализ позволяет установить прямые связи между переменными величинами по их абсолютным значениям. Эти связи графически выражаются прямой линией (отсюда и название “линейный”). *Ранговая* корреляция определяет зависимость не между абсолютными значениями переменных, а между порядковыми местами, или рангами, занимаемыми ими в упорядоченном ряду. *Парный* корреляционный анализ включает изучение корреляционных зависимостей только между парами переменных, а *множественный*, или *многомерный* – между многими переменными одновременно. Распространенной в прикладной статистике формой многомерного корреляционного анализа является *факторный анализ.*

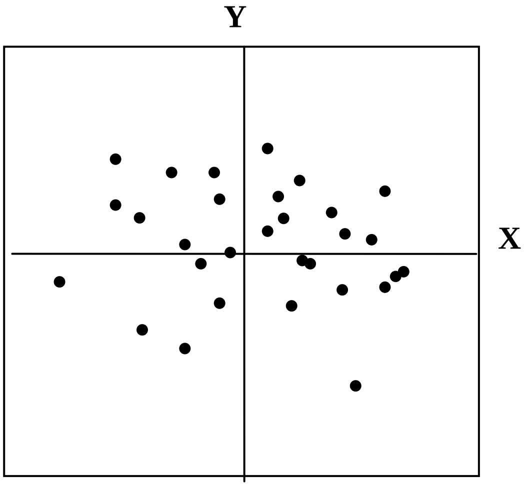
На рис.7 в виде множества точек представлены различные виды связей между двумя переменными x и y (различные виды корреляций между ними).

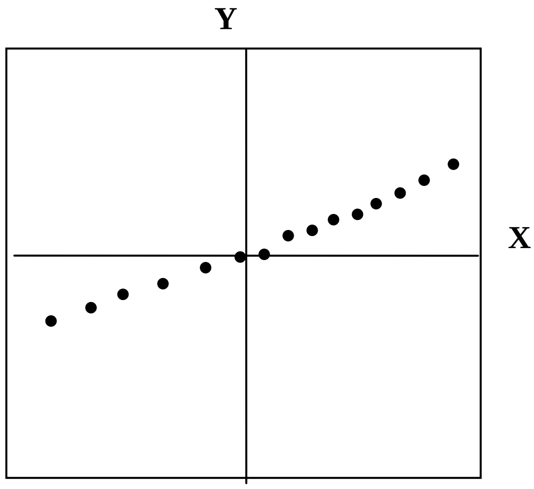
Пример: Для подсчета коэффициента линейной корреляции воспользуемся данными двух психологических тестов на выборке объемом 10 человек. Данные первого (x_i) и второго (y_i) тестов представлены в табл. 8.

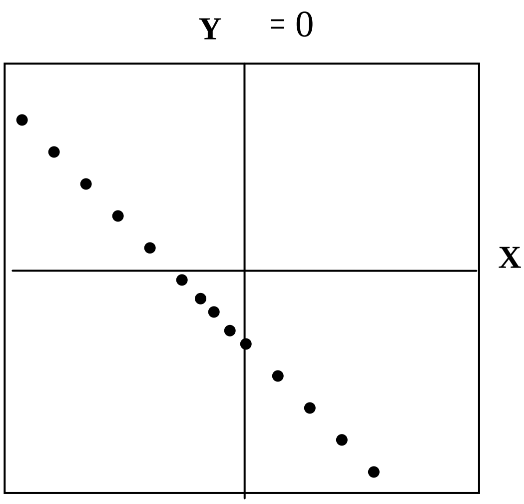
Таблица 8

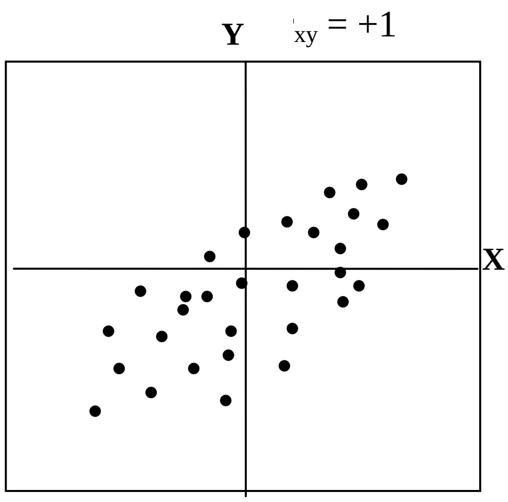
№ испытуемого	Данные по тесту № 1	Данные по тесту № 2	$X = x_i\text{-}M$	$Y = y_i - M$	$(x_i - Mx)$	$(y_i - My)$	X Y
1	65	3,2	11,9	1	141,6	1	11,9
2	32	1,3	-21,1	-0,9	445,2	0,81	18,9
3	45	1,7	-8,1	-0,5	65,6	0,25	4,05
4	56	2,4	2,9	0,2	8,4	0,04	0,58
5	57	2,4	3,9	0,2	15,2	0,04	0,78
6	53	1,5	-0,1	-0,7	0,01	0,49	0,07
7	66	2,5	12,9	0,3	166,4	0,09	3,87
8	51	2,2	-2,1	0	4,41	0	0
9	36	1,9	-17,1	-0,3	292,4	0,09	5,13

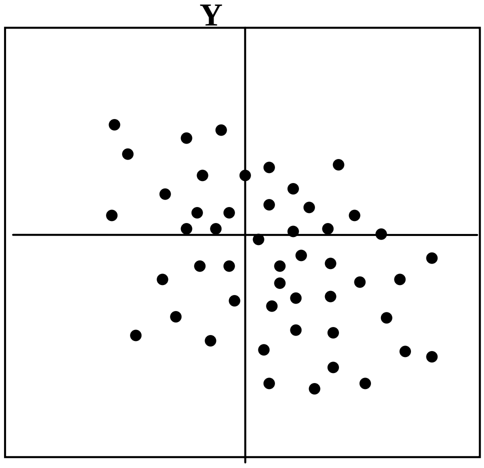
10	70	2,8	16,9	0,6	285,6	0,36	10,14
	Mx= 53,1	My= 2,2			S =424,8	S = 3,17	S = 52,5

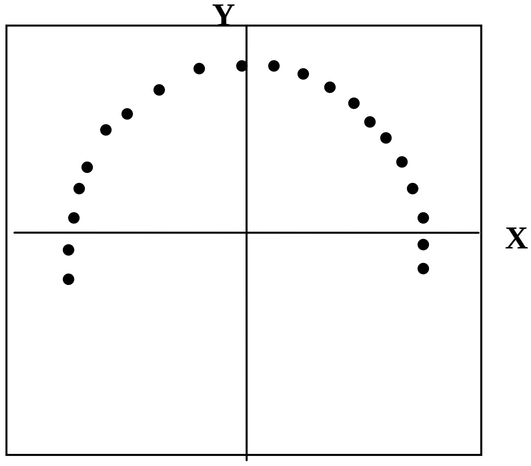












$r_{xy} = -0.3$ $r_{xy} = 0$

Рис. 7. Схематическое представление различных корреляционных зависимостей с соответствующими значениями коэффициента линейной корреляции.

Подсчитав коэффициент корреляции между данными двух тестов (переменными x и y), мы должны ответить на вопрос о вероятности ошибки или уровне значимости полученной связи. Для этого необходимо подставить полученную величину ($r=0.78$) в таблицу и с учетом объема выборки (n) определить уровень значимости коэффициента корреляции. (см. таблицы в книгах Г.Ф.Лакин «Биометрия», М., 1990г., Г.В.Суходольский «Основы математической статистики для психологов», М.,1972г., Немов Р.С. Психология, кн.3, М., 1998).

В нашем случае коэффициент корреляции оказывается достоверным на уровне значимости 0.01, т.е. вероятность ошибки составляет 1% или только в 1 случае из 100 полученный коэффициент корреляции будет недостоверным (см. таблицу 2 в Приложении).

К коэффициенту *ранговой корреляции* в психологических исследованиях обращаются в тех случаях, когда признаки, между которыми устанавливается зависимость, являются качественно различными и не могут быть достаточно точно оценены при помощи так называемой *интервальной* измерительной шкалы. Интервальной называют такую шкалу, которая позволяет оценивать расстояние между ее значениями и судить о том, какие из них больше другого и на сколько.. Например, линейка, с помощью которой оцениваются и сравниваются длины объектов, является интервальной шкалой, т.к. пользуясь ею, мы можем утверждать, что расстояние между 2см и 6см в два раза больше .чем расстояние между 6см и 8см. Если же, пользуясь некоторым измерительным инструментом, мы можем только утверждать, что одни показатели больше других, но не в состоянии сказать на сколько, то такой измерительный инструмент называется не интервальным, а *порядковым*.

Большинство показателей, которые получают в психологических исследованиях, относятся к порядковым, а не к интервальным шкалам (например, оценки типа “да”, “нет”, “скорее нет, чем да” и др., которые можно переводить в баллы). Поэтому коэффициент линейной корреляции к большинству из них не применим. В этом случае используют коэффициент *ранговой корреляции*, формула которого следующая:

$$r_s = 1 - \frac{6\left(\sum d^2\right)}{n^3 - n} \ ,$$

где r_s - коэффициент ранговой корреляции по Спирмену; d – разность между рангами сопряженных значений признаков (независимо от ее знака); n – число пар.

Обычно этот критерий используется в тех случаях, когда нужно сделать какие-то выводы не столько об *интервалах* между данными, сколько об их *рангах,* а также тогда, когда кривые распределения слишком ассиметричны и не позволяют использовать такие параметрические критерии, как коэффициент линейной корреляции.

Рассмотрим в качестве примера данные эксперимента по исследованию зависимости эффективности деятельности и времени реакции испытуемых (см.таблицу 9).

Таблица 9.

Испытуемые	Эффективность деятельности *х*	Время реакции у	Ранги х*	Ранги у*	d	d^2
Д 8	8	17	12	5	7	49
Д 9	20	13	1	2	1	1
Д 10	6	20	15	11.5	3.5	12.25
Д 11	8	18	12	7.5	4.5	20.25
Д 12	17	21	2	13.5	11.5	132.25
Д 13	10	22	8.5	15	6.5	42.25
Д 14	10	19	8.5	9.5	1	1
Ю 9	9	20	10	11.5	1.5	2.25
Ю 10	7	17	14	5	9	81
Ю 11	8	19	12	9.5	2.5	6.25
Ю 12	14	14	4	3	1	1
Ю 13	13	12	5	1	4	16
Ю 14	16	18	3	7.5	4.5	20.25
Ю 15	11	21	7	13.5	6.5	42.25
Ю 16	12	17	6	5	1	1

$$\sum d^2 = 428$$

По формуле подсчитываем значение коэффициента ранговой корреляции:

$$r_S = 1 - \frac{6 \cdot 428}{15^3 - 15} = 1 - \frac{2568}{3360} = 0.24.$$

Сравнивая полученное значение с табличным (см. таблицу 3 в Приложении), получаем, что полученный результат является недостоверным.

2.3. Факторный анализ

Метод множественных корреляций в отличие от метода парных корреляций позволяет выявить общую структуру корреляционных зависимостей, существующих внутри многомерного экспериментального материала, включающего более двух переменных, и представить эти корреляционные зависимости в виде некоторой системы.

Один из наиболее распространенных вариантов этого метода - факторный анализ - позволяет определить совокупность внутренних взаимосвязей, возможных причинно-следственных связей, существующих в экспериментальном материале. В результате факторного анализа обнаруживаются так называемые факторы - причины, объясняющие множество частных (парных) корреляционных зависимостей.

Фактор - математико-статистическое понятие. Будучи переведенным на язык психологии, он становится психологическим понятием. Например, в известном 16-факторном личностном опроснике Р. Кеттелла каждый фактор однозначно связан с определенной чертой личности человека.

Главная цель факторного анализа - сократить число переменных до небольшого числа факторов, лежащих в их основе. Как проводить факторный анализ?

Сама математическая процедура факторного анализа анализа очень сложна. Наша цель - дать общее представление о процедуре проведения факторного анализа и ее использовании в психодиагностике.

Первый шаг - выбор набора исследуемых психологических переменных. Значения переменных могут быть получены на основе одного, двух или более тестов. Далее проводится психодиагностическое исследование этих переменных на репрезентативной выборке испытуемых.

Второй шаг - подсчитывают коэффициенты корреляции между всеми переменными попарно. Коэффициенты корреляции показывают нам степень сходства между переменными. В результате этой процедуры получается матрица интеркорреляций, в каждой клетке которой на пересечении двух переменных стоит коэффициент корреляции между ними.

Третий шаг - выявление небольшого количества факторов на основе матрицы интеркорреляций. Современная психология использует специальные компьютерные программы, с помощью которых возможно провести анализ матрицы интеркорреляций и перейти от нее к гораздо меньшей по объему матрице, состоящей из нескольких независимых факторов. Это и есть базовые факторы, лежащие в основе большого набора сходных между собой переменных.

Четвертый шаг - нахождение факторных весов каждой переменной или того вклада, который она вносит в данный фактор. Переменные, входящие в каждый фактор имеют различные факторные веса. Чем факторный вес больше, тем более сильным оказывается влияние данной переменной на фактор в целом и наоборот. Эта процедура дает возможность определить относительную важность или силу переменных, составляющих данный фактор.

Приведем пример. Предположим, что психолог - исследователь желает изучить вербальные и математические способности учащихся. Изучив литературу, он приходит к выводу, что для этой цели необходимо провести диагностику следующих переменных:

1. словарь;
2. основные алгебраические понятия;
3. проблема словесных аналогий;
4. основные геометрические понятия;
5. словесные проблемы в алгебре.

Далее проводится исследование, результаты которого обрабатываются на компьютере с использованием стандартной программы факторного анализа. Компьютер прежде всего выдает исследователю матрицу интеркорреляций, состоящую из пяти строк и пяти столбцов (см. табл. 10).

Таблица 10

Переменные	Переменные				
	1	2	3	4	5
1. Словарь	1.00	0.22	0.77	0.20	0.50
2. Алгебра	0.22	1.00	0.21	0.65	0.48
3. Аналогии	0.77	0.21	1.00	0.19	0.52
4. Геометрия	0.20	0.65	0.19	1.00	0.47
5. Алгебра - слова	0.50	0.48	0.52	0.47	1.00

Каждая клетка матрицы - коэффициент корреляции между двумя переменными. Например, корреляция между словарем и алгеброй (переменными 1 и 2) равна 0.22, а между геометрией и словесными аналогиями (переменные 4 и 3) r = 0.19.

Даже первый взгляд на матрицу интеркорреляций позволяет предположить наличие двух факторов: первый связан со словарем и словесными аналогиями, второй - с алгеброй и геометрией. Также видно, что пятая переменная (словесные проблемы в алгебре) связана с обоими факторами.

Эти предположения достаточно очевидны, т.к. исследуемых переменных всего пять. Если переменных больше, сделать эту процедуру вручную невозможно. Компьютерная программа способна объективно выделить факторы, лежащие в основе исходных переменных и преобразовать матрицу следующим образом (см. табл. 11).

Таблица 11

Переменные	Факторы	
	I	II
1. Словарь	0.917	0.101
2. Алгебра	0.113	0.885
3. Аналогии	0.925	0.094
4. Геометрия	0.86	0.891
5. Алгебра - слова	0.594	0.573

Ячейки в таблице 9 являются факторными весами каждой переменной и могут рассматриваться как корреляция между данной переменной и лежащим в ее основе фактором. Например, переменная 1 (словарь) имеет высокую корреляцию с фактором I и низкую корреляцию с фактором II.Р.Кеттелл предположил, что значимыми являются факторные веса более 0.30 - 0.40 (в зависимости от объема выборки). На основе таблицы 11 можно сделать вывод, что переменные 1 и 3 составляют I фактор, переменные 2 и 4 - II фактор, а переменная 5 входит как в I, так и во II фактор.

В последние 20 лет в психологических исследованиях возрос интерес к использованию факторного анализа. Это произошло частично благодаря наличию относительно легких в использовании компьютерных программ.

Факторный анализ используют в целом ряде случаев, например для:

- нахождения факторов, лежащих в основе тестов способностей;
- обнаружения отдельных личностных черт;
- обнаружения клинических синдромов;
- обнаружения факторов удовлетворенности работой и т.д.

Однако нужно помнить, что с помощью факторного анализа невозможно найти значимые общие факторы в хаотичном наборе не связанных друг с другом переменных.

III. Способы наглядного представления результатов исследования

Результаты психолого-педагогического исследования, или психологического тестирования, кроме их текстового описания, можно представить в виде таблиц, схем, графиков, диаграмм и т.д. Такая наглядная форма представления полученных результатов позволяет структурировать, классифицировать, обобщать и сравнивать их между собой.

Таблицы представляют собой упорядоченные по горизонтали и вертикали наборы количественных и качественных данных, заключенных в рамки или без них. Таблицы могут иметь названия, подзаголовки, указывающие на то, какие данные в них содержатся.

Таблицы строятся и оформляются не произвольно, а в соответствии с определенными правилами. Рассмотрим эти правила.

Прежде всего, если в тексте имеется несколько таблиц, они нумеруются. Слово “таблица” обычно пишется справа или в середине вверху над таблицей. Непосредственно под ним располагается название таблицы. Примечания, касающиеся некоторых особенностей материала, содержащегося в таблице, помещаются, как правило, непосредственно под таблицей. Таблица обычно имеет заголовки и подзаголовки, которые указывают на то, что представлено в отдельных столбцах, а также рубрикацию по строкам, где обозначены особенности представляемого материала.

Рассмотрим в качестве примеров способы построения типичных таблиц: таблица без общего заголовка и примечаний (табл. 12); разграфленная таблица с названием и заголовком (табл. 13); разграфленная таблица с названием, заголовками и примечанием (табл. 14).

Таблица 12

	Начальные классы I-IY	Средние классы Y - IX	Старшие классы X - XI
Количество уч-ся	120	150	56
Средний возраст	8.5	12.5	15
Успеваемость (средняя оценка)	3.8	3.5	4.0

В таблице 12 нет общего заголовка, который объединил бы названия всех столбцов, а есть только названия частных подзаголовков, относящихся к отдельным столбцам. Нет также общего названия таблицы, так как содержание представленных в ней данных ясно само по себе. Названия строк таблицы характеризуют собой цифры, имеющиеся в ее строках.

Подобного рода таблицы рекомендуется строить тогда, когда общее количество данных, представленных в столбцах и строках таблицы относительно невелико (не более четырех столбцов и строк). Во всех других случаях рекомендуется строить разграфленные таблицы с названиями, общими и частными подзаголовками (табл. 13).

В тех случаях, когда в таблице необходимо представить очень большое количество данных, которые невозможно полностью описать в подзаголовках столбцов или строк из-за громоздкости самих названий, обращаются к таблицам третьего типа (табл. 14), где названия закодированы, а их расшифровка дается в примечании к таблице.

Таблица 13

Результаты обследования шестилетних и семилетних детей с точки зрения их психологической готовности к обучению в школе (данные представлены в десятибалльной шкале оценок)

Возраст детей. Место их обучения и воспитания до поступления в школу.	Основные показатели психологической готовности детей к обучению в школе									
	Интеллектуальные					Личностные			Межличност-ные	
	Внимание	Воображение	Память	Мышление	Речь	Мотивы учения	Характер	Способности	Общительность	Контактность
Шестилетние дети, посещавшие детский сад	7.2	7.6	7.9	8.0	7.1	6.2	7.2	8.0	8.4	8.4
Шестилетние дети, воспитанные дома	7.6	7.4	7.9	8.3	7.4	7.4	6.9	8.3	7.7	7.6
Семилетние дети, посещавшие детский сад	7.9	8.0	8.1	8.4	8.3	8.2	7.3	8.6	8.9	9.0
Семилетние дети, воспитанные дома	7.8	7.9	8.0	8.6	8.5	8.7	7.0	8.8	8.1	8.3

Таблица 14

Данные комплексного обследования детей 10-х классов средней школы

Условные обозначения детей	Показатели обследования											
	I				II				III			
	1	2	3	4	5	6	7	8	9	10	11	12
А												
Б												

В												
…												
Я												

Примечание. А - Иванов, Б – Петров, В – Сидоров, …, Я – Яковлев; I – социально-демографические данные о детях; II – успеваемость по отдельным предметам; III – данные о психологическом развитии; 1 – возраст, 2 – пол; 3 – социальное происхождение, 4 – место жительства, 5 – математика, 6 – физика, 7 – история, 8 – география, 9 – внимание, 10 – память, 11 – мышление, 12 – речь.

Другой способ наглядного представления эмпирических данных - графический. График на плоскости представляет собой некоторую линию, которая изображает зависимость между двумя переменными, при этом на горизонтальной оси обычно размещают независимую переменную, а на вертикальной оси - зависимую переменную. На рис. 8 представлен график зависимости между математическими способностями и успеваемостью учащихся по математике.

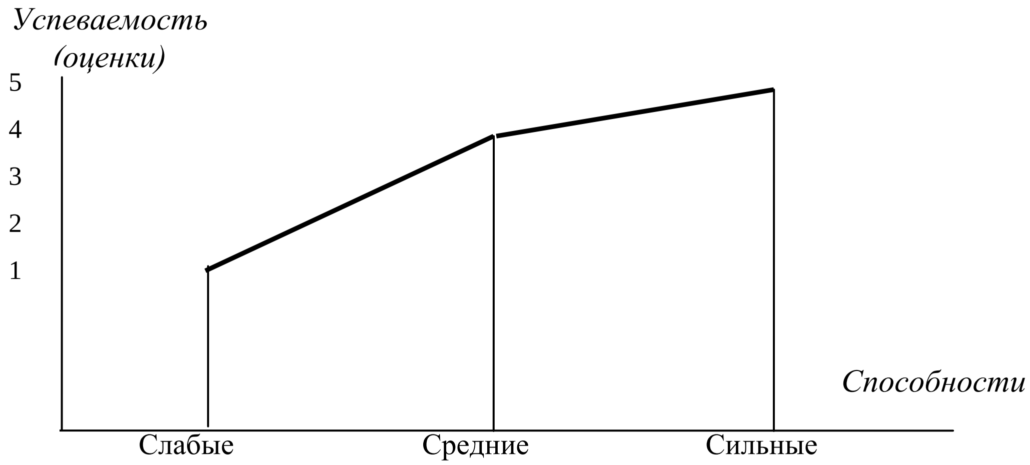


Рис. 8. График зависимости между уровнем способностей и успеваемостью учащихся.

Особую разновидность графических изображений экспериментальных результатов представляют собой гистограммы (см. примеры гистограмм на рис.9).

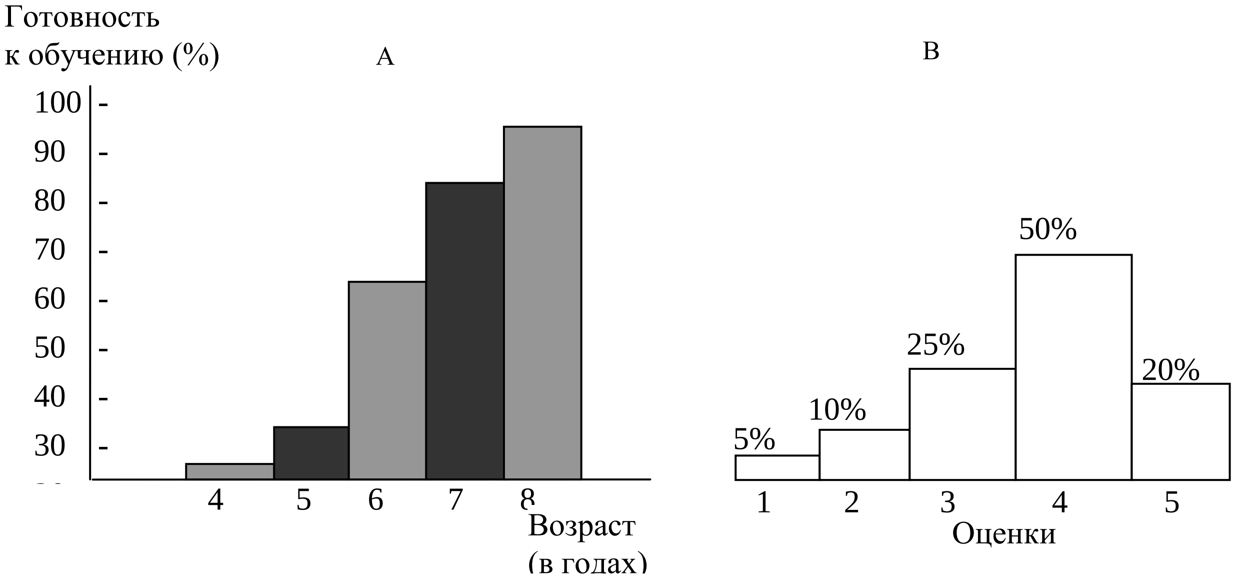


Рис.9. Виды гистограмм на плоскости. А - гистограмма распределения показателей готовности детей разного возраста к обучению в школе; Б- гистограмма распределения оценок в классе.

Это столбчатые диаграммы, состоящие из вертикальных прямоугольников, расположенных основаниями на одной прямой. Их высота отражает степень или уровень развития того или иного качества у испытуемого. Цифры, указывающие на частоту встречаемости качества в выборке испытуемых, размещаются либо над столбцами диаграммы, либо по вертикальной оси графика.

Этап статистической обработки данных, полученных в ходе психологического исследования, и представления их в графической форме завершается качественной интерпретацией. В известной степени интерпретация зависит от интуиции и профессионального опыта психолога и редко протекает по жесткой схеме.

Заключение.

Итак, мы рассмотрели различные статистические методы, используемые в психологии. Главная задача данной работы заключалась в том, чтобы читатель понял, что статистика не так страшна, как кажется. Статистика – это прежде всего способ мышления, и для ее применения нужно лишь иметь немного здравого смысла и знать основы математики. В нашей повседневной жизни мы, сами того не подозревая, занимаемся статистикой. Хотим ли мы спланировать бюджет, рассчитать потребление бензина автомашиной, оценить усилия, которые потребуются для усвоения какого-либо курса, с учетом полученных до сих пор отметок, предусмотреть вероятность хорошей и плохой погоды по метеорологической сводке, или же вообще оценить, как повлияет то или иное событие на наше личное или совместное будущее, - нам постоянно приходится отбирать, классифицировать и упорядочивать информацию, связывать ее с другими данными так, чтобы можно было сделать выводы, позволяющие принять верное решение.

Все эти виды деятельности мало отличаются от тех операций, которые лежат в основе научного исследования. В основе их лежит синтез данных, полученных на различных группах объектов в том или ином эксперименте, сравнение данных с целью выявить различия между ними, сопоставление с целью выявить показатели, изменяющиеся в одном направлении, и, наконец, прогноз событий на основании тех выводов, к которым приводят полученные результаты. Именно в этом заключается цель применения статистики в науках вообще, и в особенности - в гуманитарных. Поскольку именно в этих науках нет ничего абсолютно достоверного, без применения статистических методов выводы в большинстве случаев были бы чисто интуитивными и не могли бы служить основанием для интерпретации данных, получаемых в научных исследованиях.

ПРИЛОЖЕНИЕ

Словарь терминов

Выборочное среднее (М) - см. среднее арифметическое.

Гистограмма - столбчатая диаграмма, состоящая из вертикальных прямоугольников, расположенных основаниями на одной прямой. Их высота отражает степень или уровень развития того или иного качества у испытуемого.

Дисперсия (s^2) - мера разброса данных, которая характеризует насколько индивидуальные результаты отклоняются от среднего арифметического в данной выборке. Чем больше дисперсия, тем больше разброс данных.

Закон трех сигм - закон, характеризующий распределение результатов на кривой нормального распределения. При нормальном распределении почти все случаи (99,72%) лежат в интервале от $-3s$ до $+3s$ относительно среднего арифметического.

Корреляционный анализ - комплекс методов статистического исследования взаимосвязи между переменными. Корреляционный анализ позволяет установить силу и направление связи между двумя и более переменными.

Корреляции коэффициент (r) - количественный показатель от -1 до +1, который обозначает силу и направление связи между двумя переменными. Чем он ближе к +1 или -1 тем связь сильнее. Положительная корреляция обозначает, что большие значения одной переменной соответствуют большим значениям другой переменной. Отрицательная корреляция обозначает, что большие значения одной переменной соответствуют малым значениям другой переменной.

Критерии достоверности различий - статистические методы установления достоверности различий между элементарными статистиками двух выборок (например, средними, дисперсиями).

Медиана (Ме) - первичный результат, находящийся в середине ряда показателей, упорядоченных в порядке возрастания или убывания. Справа и слева от медианы в упорядоченном ряду находится одинаковой количество данных.

Меры центральной тенденции - характеристики совокупности переменных, указывающие на наиболее типичный, репрезентативный для данной выборки результат. Наиболее распространенными в психодиагностике мерами центральной тенденции являются: среднее арифметическое (М), мода (Мо) и медиана (Ме).

Меры разброса (вариативности) - характеризуют степень индивидуальных отклонений от центральной тенденции. Чаще всего в психодиагностике используются такие меры разброса как дисперсия (s^2) и стандартное отклонение (s).

Мода (Мо) - наиболее часто встречающийся первичный результат, либо середина интервала значений, для которого частота максимальна.

Нормальное распределение - вид теоретического распределения переменных. Нормальное распределение наблюдается при изменении переменной под влиянием множества относительно независимых факторов. График нормального распределения представляет собой симметричную колоколообразную кривую, обладающую целым рядом свойств (свойства нормального распределения см. в тексте).

Объем выборки (n) - количество членов выборочной совокупности.

Оценки первичные (сырые баллы) - оценки, получаемые испытуемым на начальном этапе обработки результатов тестовой методики. Обычно это сведения о количестве правильно решенных задач, числе попыток при их решении, реже - о времени выполнения заданий.

Оценки стандартные (шкальные) - способ оценки результата теста путем установления его места на специальной шкале. Шкала содержит данные о внутригрупповых нормах выполнения данной методики в выборке стандартизации. Так, индивидуальное выполнение заданий (первичные оценки испытуемых) сравнивается с их выполнением в нормативной группе (репрезентативной выборке стандартизации).

Полигон частот - графическое изображение частотного распределения, в котором по горизонтальной оси откладываются первичные результаты (либо интервалы значений), а по вертикальной оси - соответствующая им частота. Точки пересечения первичных результатов и их частот соединяются прямолинейными отрезками.

Процентиль - процентная доля испытуемых из выборки стандартизации, первичный результат которых равен данному или ниже его.

Размах распределения - разность между самым высоким и самым низким первичным результатом в выборке.

Репрезентативность - свойство выборочной совокупности представлять характеристики генеральной совокупности. Репрезентативность означает, что с некоторой погрешностью можно считать, что представленное в выборочной совокупности распределение изучаемых переменных соответствует их реальному распределению в генеральной совокупности.

Совокупность выборочная (выборка) - часть элементов генеральной совокупности, подвергающаяся психодиагностическому исследованию. Главное требование к выборочной совокупности - ее репрезентативность.

Совокупность генеральная - множество элементов, объединенных общей характеристикой, указывающей на их принадлежность к определенной системе.

Среднее арифметическое (М) - средняя оценка изучаемой в исследовании переменной, которая вычисляется как сумма всех первичных результатов, деленная на объем выборки.

Стандартизация - приведение к единым нормативам процедуры и оценок теста. Благодаря стандартизации методики достигается сопоставимость полученных результатов у разных испытуемых, появляется возможность выражения тестовых оценок в стандартных показателях, сопоставления таких оценок в разных тестовых методиках.

Стандартное отклонение (s) - мера разброса данных, равная квадратному корню из дисперсии.

Уровень значимости (a)- степень вероятности ошибочного вывода относительно статистической гипотезы, проверяемой на основе выборочных данных. В психологических исследованиях за достаточный уровень значимости обычно принимается a = 0,05, что означает, что вероятность допустимой ошибки равна 5%. Для достаточно больших выборок уровень значимости может быть принят a =0,1 и a = 0,01.

Факторный анализ - математический метод, основанный на многомерном корреляционном анализе эмпирических данных и позволяющий выявить факторы, лежащие в основе ряда сходных переменных.

Частота (количество случаев) - число, показывающее сколько раз данный первичный результат встречается в данной выборке.

Частотное распределение - таблица, состоящая из двух столбцов, в левом столбце которой расположены первичные результаты теста, упорядоченные по убыванию, а в правом - соответствующая каждому результату частота.

Таблицы

Таблица 1

Критические значения t-критерия Стьюдента для различных степеней свободы $\eta = n_1 + n_2 - 2$ и уровней значимости.

η	0.05	0.01
4	2.78	5.60
5	2.58	4.03
6	2.45	3.71
7	2.37	3.50
8	2.31	3.36
9	2.26	3.25
10	2.23	3.17
11	2.20	3.11
12	2.18	3.05
13	2.16	3.01
14	2.14	2.98
15	2.13	2.96
16	2.12	2.92
17	2.11	2.90
18	2.10	2.88
19	2.09	2.86
20	2.09	2.85
21	2.08	2.83
22	2.07	2.82
23	2.07	2.81
24	2.06	2.80
25	2.06	2.79
26	2.06	2.78
27	2.05	2.77
28	2.05	2.76
29	2.05	2.76
30	2.04	2.75
40	2.02	2.70
50	2.01	2.68
60	2.00	2.66
80	1.99	2.64
100	1.98	2.63

Таблица 2

Критические значения коэффициента линейной корреляции r для различных
степеней свободы η= n - 2 и уровней значимости

η	0.05	0.01
2	0.9500	0.9900
3	0.8783	0.9587
4	0.8114	0.9172
5	0.7545	0.8745
6	0.7067	0.8343
7	0.6664	0.7977
8	0.6319	0.7646
9	0.6021	0.7348
10	0.5760	0.7079
11	0.5529	0.6833
12	0.5324	0.6614
13	0.5139	0.6411
14	0.4973	0.6226
15	0.4821	0.6055
16	0.4683	0.5897
17	0.4555	0.5751
18	0.4438	0.5614
19	0.4329	0.5487
20	0.4227	0.5368
21	0.4132	0.5256
22	0.4044	0.5151
23	0.3961	0.5052
24	0.3882	0.4958
25	0.3809	0.4869
26	0.3739	0.4785
27	0.3673	0.4705
28	0.3610	0.4629
29	0.3550	0.4556
30	0.3494	0.4487
31	0.3440	0.4421
32	0.3388	0.4357
33	0.3338	0.4297
34	0.3291	0.4238
35	0.3246	0.4182
36	0.3202	0.4128
37	0.3160	0.4076

38	0.3120	0.4026
39	0.3081	0.3978
40	0.3044	0.3932

Таблица 3.

Критические значения коэффициента ранговой корреляции r_s для различных степеней свободы η= n - 2 и уровня значимости 0.05

η	0.05
2	1.000
3	0.900
4	0.829
5	0.714
6	0.643
7	0.600
8	0.564
10	0.506
12	0.456
14	0.425
16	0.399
18	0.377
20	0.359
22	0.343
24	0.329
26	0.317
28	0.306

Таблица 4

Критические значения F-критерия Фишера для уровня значимости 0.05 и числа степеней свободы n_1 и n_2

n_1 n_2	3	4	5	6	8	12	16	24	50
3	9.28	9.91	9.01	8.94	8.84	8.74	8.69	8.64	8.58
4	6.59	6.39	6.26	6.16	6.04	5.91	5.84	5.77	5.70
5	5.41	5.19	5.05	4.95	4.82	4.68	4.60	4.58	4.44
6	4.76	4.53	4.39	4.28	4.15	4.00	3.92	3.84	3.75
8	4.07	3.84	3.69	3.58	3.44	3.28	3.20	3.12	3.03
12	3.49	3.26	3.11	3.00	2.85	2.69	2.60	2.50	2.40
16	3.24	3.00	2.85	2.74	2.59	2.42	2.33	2.24	2.13
24	3.01	2.78	2.62	2.51	2.36	2.18	2.09	1.98	1.86
50	2.79	2.56	2.40	2.29	2.13	1.95	1.85	1.74	1.60

Примечание: Для больших выборок или иных уровней значимости следует обратиться к таблицам в пособиях по статистике.

ЛИТЕРАТУРА

1. Анастази А. Психологическое тестирование. Том I. М., 1982.
2. Бурлачук Л.Ф., Морозов С.М. Словарь-справочник по психологической диагностике. Киев, 1989.
3. Готтсданкер Р. Основы психологического эксперимента. М., 1982.
4. Закс Л. Статистическое оценивание. М., 1976.
5. Кулагин Б.В. Основы профессиональной психодиагностики. Л., 1984.
6. Лакин Г.Ф. Биометрия. М., 1990.
7. Немов Р.С. Психология. Кн. 3. Психодиагностика. М., 1998.
8. Общая психодиагностика / под ред. А.А. Бодалева, В.В. Столина. М., 1987.
9. Суходольский Г.В. Основы математической статистики для психологов. М., 1972.
10. Фресс П., Пиаже Ж. Экспериментальная психология. Вып. I и II. М., 1966.
11. Практикум по общей психологии / под ред. А.И. Щербакова. М., 1990.
12. Психодиагностические методы (в комплексном лонгитюдном исследовании студентов) / под ред. А.А. Бодалева, М.Д. Дворяшиной, И.М. Палея. Л., 1976.