Содержание

Введение

Глава 1. Теоретическая часть

1.1 Основные понятия, используемые в математической обработке психологических данных

1.2 Статистический анализ экспериментальных данных

1.3 Вторичные методы обработки материалов психологических исследований

Глава 2. Практическая часть

2.1 Ранговая корреляция

Заключение

Литература

## Введение

Психология получила статус науки благодаря эксперименту и использованию математики при обработке экспериментальных данных и психологических исследований. Математика в психологии служит таким логическим инструментом доказательства, давая возможность научного понимания психологических закономерностей и более глубокого их анализа Математическая статистика - область современной математики, основанная на теории вероятностей и занятая поиском законов изменения и способов измерения случайных величин, обоснованием методов расчетов, производимых с такими величинами.

Математическая статистика возникла (XVII в) и развивалась параллельно с теорией вероятностей. Дальнейшее развитие математической статистики (вторая половина XIX - начало XX в) обязано, в первую очередь, П.Л. Чебышеву, А.А. Маркову, А.М. Ляпунову, а также К. Гауссу, А. Кетле, Ф. Гальтону, К. Пирсону и др.

В XX в. Наиболее существенный вклад в математическую статистику был сделан советскими математиками (В.И. Романовский, Е.Е. Слуцкий, А.Н. Колмогоров, Н.В. Смирнов), а также английскими (Стъюдент, Р. Фишер, Э. Пирсон) и американскими (Ю. Нейман, А. Вальд) учеными.

Еще в середине XIX начале XX века наблюдается, правда, еще не вполне регулярные, но, тем не менее, приносящие обоюдную пользу, - попытки провести аналогии между психологическими и физическими исследованиями, особенно в области построения лабораторного эксперимента, анализа и обработки экспериментальных данных. Почти одновременно в психологию и физику приходят вероятностные и статистические методы, теория дифференциальных уравнений, вариационное исчисление и другие. О том, чтобы математически описать деятельность мозга мечтал И.П. Павлов.

Психология получила статус науки благодаря эксперименту (как естественно-научная дисциплина) и математической статистике. Благодаря проникновению в количественные свойства психических явлений, психология получила множество логических доказательств, которые явились научным обоснованием изучения психики человека. Именно поэтому математика как строгая логическая дисциплина необходима любому специалисту, практикующемуся в области психологии. Современная математическая статистика представляет собой большую и сложную систему знаний. Математическая статистика нужна психологу не только для проведения научных исследований, а постоянно в его повседневной работе. Статистики разработали целый комплекс простых методов, которые совершенно доступны любому квалифицированному специалисту психологу.

## Глава 1. Теоретическая часть

## 1.1 Основные понятия, используемые в математической обработке психологических данных

1) признаки и переменные.

2) шкалы измерения.

3) Статистические гипотезы.

4) Статистические критерии.

1. Признаки и переменные - это измеряемые психологические явления. Такими явлениями могут быть время решения задачи, количество допущенных ошибок, уровень тревожности, показатель интеллектуальной лабильности, интенсивность агрессивных реакций, угол поворота корпуса в беседе, показатель социометрического статуса и множество других переменных. Понятия признака и переменной могут использоваться как взаимозаменяемые. Они являются наиболее общими. Иногда вместо них используются понятия показателя или уровня, например уровень настойчивости, показатель вербального интеллекта и др.

Математическая обработка - это оперирование со значениями признака, полученными у испытуемых в психологическом исследовании. Такие индивидуальные результаты называют также "наблюдениями", "наблюдаемыми значениями", "вариантами", "датами" и др. значение признака определяется при помощи специальных шкал измерения.

2. Шкалы измерения. Измерение - это приписывание числовых форм объектами или событиям в соответствии с определенными правилами.С. Стивенсом предложена классификация из 4 типов шкал измерения:

а) Номинативная, или номинальная, или шкала наименований;

б) Порядковая, или ординальная, шкала;

в) Интервальная, или шкала равных интервалов;

г) Шкала равных отношений.

Шкала наименований**.** К этой шкале относятся материалы, в которых изучаемые объекты отличаются друг от друга по их качеству. При обработке таких материалов нет никакой нужды в том, чтобы располагать эти объекты в каком-то порядке, исходя из их характеристик.

Шкала порядка. Если в шкале наименований порядок следования изучаемых объектов практически не играет никакой роли, то в шкале порядка - это видно из ее названия - именно на эту последовательность переключается все внимание. К этой шкале в статистике относят такие исследовательские материалы, в которых рассмотрению подлежат объекты, принадлежащие к одному или нескольким классам, но отличающиеся при сравнении одного с другим: больше - меньше, выше - ниже и т.п.

Шкала интервалов. К ней относятся такие материалы, в которых дана количественная оценка изучаемого объекта в фиксированных единицах. Например, в опытах учитывалось, сколько точек могут поставить, работая с максимально доступной скоростью, испытуемые. Оценочными единицами в опытах служило число точек. Подсчитав их, исследователь получил то абсолютное число точек, которое оказалось возможным поставить за отведенное время каждому участнику опытов. Главная трудность при отнесении материалов к шкале интервалов состоит в том, что нужно располагать такой единицей, которая была бы при всех повторных изменениях тождественной самой себе, т.е. одинаковой и неизменной.

Шкала отношений. К этой шкале относятся материалы, в которых учитываются не только число фиксированных единиц, как в шкале интервалов, но и отношения полученных суммарных итогов между собой. Чтобы работать с такими отношениями, нужно иметь некую абсолютную точку, от которой ведется отчет. При изучении психологических объектов эта шкала практически неприменима.

3. Статистические гипотезы. Формулирование гипотез систематизирует предположения исследователя и представляет их в четком и лаконичном виде. Благодаря гипотезам исследователь не теряет путеводной нити в процессе расчетов и ему легко понять после их окончания, что, собственно, он обнаружил. Статистические гипотезы подразделяются на нулевые и альтернативные, направленные и ненаправленные.

Нулевая гипотеза - это гипотеза об отсутствий различий. Она обозначается как Н0 и называется нулевой потому, что содержит число 0: Х1 - Х2 = 0, где Х1, Х2 - сопоставляемые значения признаков. Нулевая гипотеза - это то, что мы хотим опровергнуть, если перед нами стоит задача доказать значимость различий.

Альтернативная гипотеза - это гипотеза о значимости различий. Она обозначается как Н1. альтернативная гипотеза - это то, что мы хотим доказать, поэтому иногда ее называют экспериментальной гипотезой.

Нулевая и альтернативная гипотезы могут быть направленными и ненаправленными.

Статистические критерии.

Статистический критерий - это правило, обеспечивающее надежное поведение, то есть принятие истинной и отклонение ложной гипотезы с высокой вероятностью. Статистический критерий обозначает метод расчета определенного числа и само это число.

Параметрические критерии - это критерии, включающие в формулу расчета параметры распределения, то есть средние и дисперсии (t-критерий Стъюдента, критерий F и др.) Непараметрические критерии - это критерии, не включающие в формулу расчета параметров распределения и основанные на оперировании частотами или рангами (критерий-Q Розенбаума, критерий-Т Вилкоксона и др.) Параметрические критерии и непараметрические критерии имеют свои преимущества и недостатки.

Параметрические критерии:

1. Позволяют прямо оценить различия в средних, полученных в двух выборках (t - критерий Стъюдента).

2. Позволяют прямо оценить различия в дисперсиях (критерий Фишера) 3. Позволяют выявить тенденции изменения признака при переходе от условия к условию (дисперсионный однофакторный анализ), но лишь при условии нормального распределения признака.

4. Позволяют оценить взаимодействие двух и более факторов в их влиянии на изменения признака (двухфакторный дисперсионный анализ).

5. Экспериментальные данные должны отвечать двум, а иногда трем, условиям:

а) значения признака измерены по интервальной шкале;

б) распределение признака является нормальным;

в) в дисперсионном анализе должно соблюдаться требование равенства дисперсий в ячейках комплекса.

6. Математические расчеты довольно сложны.

7. Если условия, перечисленные в п.5, выполняются, параметрические критерии оказываются несколько более мощными, чем непараметрические.

Непараметрические критерии.

1. Позволяют оценить лишь средние тенденции, например, ответить на вопрос, чаще ли в выборке А встречаются более высокие, а в выборке Б - более низкие значения признака (критерии Q, U, и др.).

2. Позволяют оценить лишь различия в диапазонах вариативности признака (критерий).

3. Позволяют выявить тенденции изменения признака при переходе от условия к условию при любом распределении признака (критерии L и S).

4. Эта возможность отсутствует.

5. Экспериментальные данные могут не отвечать ни одному из этих условий:

а) значения признака могут быть представлены в любой шкале, начиная от шкалы наименований;

б) распределение признака может быть любым и совпадение его с каким-либо теоретическим законом распределения необязательно и не нуждается в проверке;

в) требования равенства дисперсий отсутствует.

6. Математические расчеты по большей части просты и занимают мало времени (за исключением лишь некоторых критериев).

7. Если условия, перечисленные в п.5, не выполняются, непараметрические критерии оказываются более мощными, чем параметрические, так как они менее чувствительны к "засорениям".

## 1.2 Статистический анализ экспериментальных данных

Методы первичной статистической обработки результатов эксперимента Статистические методы применяются при обработке материалов психологических исследований для того, чтобы извлечь из тех количественных данных, которые получены в экспериментах, при опросе и наблюдениях, возможно больше полезной информации. В частности, в обработке данных, получаемых при испытаниях по психологической диагностике, это будет информация индивидуально-психологических особенностях испытуемых.

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. Речь идет о таких закономерностях статистического характера, которые существуют между изучаемыми в эксперименте переменными величинами.

Некоторые из методов математико-статистического анализа позволяют вычислять так называемые элементарные математические статистики, характеризующие выборочное распределение данных, например, выборочное среднее, 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, скажем, корреляционного анализа, факторного анализа, методов сравнения выборочных данных, можно достоверно судить о статистических связях, существующих между переменными величинами, которые исследуют в данном эксперименте.

Все методы математико-статистического анализа условно делятся на первичные и вторичные. Первичными называют методы, с помощью которых можно получить показатели, непосредственно отражающие результаты производимых в эксперименте измерений. Соответственно под первичными статистическими показателями имеются в виду те, которые применяются в самих психодиагностических методиках и являются итогом начальной статистической обработки результатов психодиагностики. К первичным методам статистической обработки относят, например, определение выборочной средней величины, выборочной дисперсии, выборочной моды и выборочной медианы. В число вторичных методов обычно включают корреляционный анализ, регрессионный анализ, методы сравнения первичных статистик у двух или нескольких выборок.

Рассмотрим методы вычисления элементарных математических статистик, начав с выборочного среднего.

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 психодиагностическому обследованию. Сравнивая непосредственно средние значения двух или нескольких выборок, мы можем судить об относительной степени развития у людей, составляющих эти выборки, оцениваемого качества. Выборочное среднее определяется при помощи следующей формулы:

где - выборочная средняя величина или среднее арифметическое значение по выборке; n количество испытуемых в выборке или частных психодиагностических показателей, на основе которых вычисляется средняя величина; хk частные значения показателей у отдельных испытуемых. Всего таких показателей n, поэтому индекс k данной переменной принимает значения от 1 до n; принятый в математике знак суммирования величин тех переменных, которые находятся справа от этого знака.

Дисперсия как статистическая величина характеризует, на сколько частные значения отклоняются от средней величины в данной выборке. Чем больше дисперсия, тем больше отклонения или разброс данных. Иногда вместо дисперсии для выявления разброса частных данных относительно средней используют производную от дисперсии величину, называемую выборочное отклонение. Оно равно квадрат ному корню, извлекаемому из дисперсии, и обозначается тем же самым знаком, что и дисперсия, только без квадрата - :

Медианой называется значение изучаемого признака, которое делит выборку, упорядоченную по величине данного признака, пополам. Справа и слева от медианы в упорядоченном ряду остается по одинаковому количеству признаков.

Мода еще одна элементарная математическая статистика и характеристика распределения опытных данных. Модой называют количественное значение исследуемого признака, наиболее часто встречающееся в выборке. Иногда исходных частных первичных данных, которые подлежат статистической обработке, бывает довольно много, и они требуют проведения огромного количества элементарных арифметических операций. Для того чтобы сократить их число и вместе с тем сохранить нужную точность расчетов, иногда прибегают к замене исходной выборки частных эмпирических данных на интервалы. Интервалом называется группа упорядоченных по величине значений признака, заменяемая в процессе расчетов сред ним значением.

## 1.3 Вторичные методы обработки материалов психологических исследований

С помощью вторичных методов статистической обработки экспериментальных данных непосредственно проверяются, доказываются или опровергаются гипотезы, связанные с экспериментом. Эти методы, как правило, сложнее, чем методы первичной статистической обработки, и требуют от исследователя хорошей подготовки в области элементарной математики и статистики. Обсуждаемую группу методов можно разделить на несколько подгрупп:

1. Регрессионное исчисление.

2. Методы сравнения между собой двух или нескольких элементарных статистик (средних, дисперсий и т.п.), относящихся к разным выборкам.

3. Методы установления статистических взаимосвязей между переменными, например их корреляции друг с другом.

4. Методы выявления внутренней статистической структуры эмпирических данных (например, факторный анализ).

Регрессионное исчисление - это метод математической статистики, позволяющий свести частные, разрозненные данные к некоторому линейному графику, приблизительно отражающему их внутреннюю взаимосвязь, и получить возможность по значению одной из переменных приблизительно оценивать вероятное значение другой переменной.

Следующий метод вторичной статистической обработки, посредством которого выясняется связь или прямая зависимость между двумя рядами экспериментальных данных, носит название метод корреляций. Он показывает, каким образом одно явление влияет на другое или связано с ним в своей динамике. Подобного рода зависимости существуют, к примеру, между вели чинами, находящимися в причинно-следственных связях друг с другом. Если выясняется, что два явления статистически достоверно коррелируют друг с другом и если при этом есть уверенность в том, что одно из них может выступать в качестве причины другого явления, то отсюда определенно следует вывод о наличии между ними причинно-следственной зависимости.

Имеется несколько разновидностей данного метода: линейный, ранговый, парный и множественный. Линейный корреляционный анализ позволяет устанавливать прямые связи между переменными величинами по их абсолютным значениям. Эти связи графически выражаются прямой линией, отсюда название "линейный". Ранговая корреляция определяет зависимость не между абсолютными значениями переменных, а между порядковыми местами, или рангами, занимаемыми ими в упорядочен ном по величине ряду. Парный корреляционный анализ включает изучение корреляционных зависимостей только между парами переменных, а множественный, или многомерный, между многими переменными одновременно.

## Глава 2. Практическая часть

## 2.1 Ранговая корреляция

В психологии часто возникает потребность анализа связи между переменными, которые не могут быть измерены в интервальной или реляционных шкалах, но тем не менее поддаются упорядочению и могут быть проранжированы по степени убывания или возрастания признака. Для определения тесноты связи между признаками, измеренными в порядковых шкалах, применяются методы ранговой корреляции. К ним относятся: коэффициенты ранговой корреляции Спирмена и Кендалла (используются для определения тесноты связи между двумя величинами) и коэффициент конкордации (устанавливает статистическую связь между несколькими признаками). Использование коэффициента линейной корреляции Пирсона в случае, когда о законе распределения и о типе измерительной шкалы отсутствует сколько-нибудь надежная информация, может привести к существенным ошибкам.

Методы ранговой корреляции могут быть использованы для определения тесноты связи не только между количественными переменными, но и между качественными признаками при условии, что их значения можно упорядочить и проранжировать. Эти методы также могут быть использованы применительно к признакам, измеренным в интервальных и реляционных шкалах, однако их эффективность в этом случае всегда будет ниже.

Коэффициент ранговой корреляции Спирмена. Каждая из двух совокупностей располагается в виде вариационного ряда с присвоением каждому члену ряда соответствующего порядкового номера (ранга), выраженного натуральным числом. Одинаковым значениям ряда присваивают среднее ранговое число.

Сравниваемые признаки можно ранжировать в любом направлении:

как в сторону ухудшения качества (ранг 1 получает самый большой, быстрый, умный и т.д. испытуемый), так и наоборот. Главное, чтобы обе переменные были проранжированы одинаковым способом.

Коэффициент ранговой корреляции Спирмена находится по формуле n

6 ⋅ ∑ d i2

rS = 1 − i =1, n −n3

где di - разность рангов для каждой i-пары из n наблюдений.

Если в вариационных рядах для X и Y встречаются члены ряда с одинаковыми ранговыми числами, то в формулу для коэффициента корреляции Спирмена необходимо внести поправки Tx и Ty на одинаковые ранги:

n

6 ⋅ ∑ d i2 l

rS = 1 − i =1, T = ∑ (t k − t k).

3

1

(n 3 − n) − (Tx + T y) k =1

2

Здесь l - число групп в вариационном ряду с одинаковыми ранговыми числами; tk - число членов в каждой из l групп.

Ранговый коэффициент корреляции Спирмена, как и линейный, изменяется от -1 до +1, однако значение рангового коэффициента корреляции Спирмена всегда меньше значения коэффициента линейной корреляции Пирсона: rS < r.

Проверка гипотезы о значимости коэффициента ранговой корреляции Спирмена проводится по-разному в зависимости от объема выборки.

1. Объем выборки больше 30 (n > 30).

Проверка нулевой гипотезы h0: с = 0 при альтернативной h1: с ≠ 0 осуществляется с помощью критерия Стьюдента и заключается в вычислении величины rS

t = ⋅ n−2,1 − rS2

имеющей распределение Стьюдента с df = n - 2 степенями свободы. Эмпирическое значение сравнивается с критическими значениями tб (n - 2).

Нулевая гипотеза с = 0 не отвергается, если эмпирическое значение попадает в область допустимых значений:

| t | ≤ t0,05 (df), df = n - 2.

Коэффициент ранговой корреляции Спирмена значимо отличается от нуля, если эмпирическое значение попадает в критическую область:

| t | > t0,01 (df), df = n - 2.

2. Очень малый объем выборки (n ≤ 30).

Проверка нулевой гипотезы осуществляется путем сравнения вычисленного коэффициента rS с критическими значениями rб (n), взятым из статистических таблиц для выбранного уровня значимости б и числа пар наблюдений n (табл.3.1). Нулевая гипотеза с = 0 не отвергается, если эмпирическое значение попадает в область допустимых значений:

| rS | ≤ r0,05 (n).

Коэффициент ранговой корреляции Спирмена значимо отличается от нуля, если вычисленное значение попадает в критическую область:

| rS | > r0,01 (n).

Таблица 3.1

Критические значения коэффициента ранговой корреляции Спирмена

|  |
| --- |
| α α α |
| n n n |
| 0,05 0,01 0,05 0,01 0,05 0,01 |
| 7 0,745 0,893 15 0,518 0,654 23 0,415 0,531 |
| 8 0,690 0,857 16 0,500 0,632 24 0,406 0,520 |
| 9 0,663 0,817 17 0,485 0,615 25 0,398 0,510 |
| 10 0,636 0,782 18 0,472 0,598 26 0,389 0,500 |
| 11 0,609 0,754 19 0,458 0,582 27 0,383 0,491 |
| 12 0,580 0,727 20 0,445 0,568 28 0,375 0,483 |
| 13 0,555 0,698 21 0,435 0,555 29 0,368 0,474 |
| 14 0,534 0,675 22 0,424 0,543 30 0,362 0,466 |

В методике С.А. Будаси испытуемому предлагается проранжировать 20 качеств по степени желательности (ранг 20 присуждается самому желаемому качеству). Затем в другой колонке его просят проранжировать эти же качества по степени выраженности у него в данный момент (ранг 20 получает самое характйрное качество). На основе расчета коэффициента ранговой корреляции Спирмена делается вывод об уровне самооценки испытуемого. Результаты испытуемого С. О-ва приведены в таблице 3.2 Требуется рассчитать коэффициент корреляции Спирмена между выраженностью качеств у обследуемого испытуемого в данный момент и его идеальным представлением.

Решение:

Составляем расчетную таблицу, в которую заносим две ранговые последовательности (желаемую N и реальную N'), разности рангов d и d2.

Таблица 3.2

Расчет коэффициента ранговой корреляции Спирмена

|  |  |
| --- | --- |
| Качества | N N’ d = N - N’ d2 |
| уступчивость | 14 15 - 1 1 |
| смелость | 15 18 - 3 9 |
| вспыльчивость | 2 16 - 14 196 |
| настойчивость | 13 13 0 0 |
| нервозность | 1 7 - 6 36 |
| терпеливость | 17 10 7 49 |
| увлекаемость | 12 20 - 8 64 |
| пассивность | 8 2 6 36 |
| холодность | 10 19 - 9 81 |
| энтузиазм | 9 17 - 8 64 |
| осторожность | 16 4 12 144 |
| капризность | 3 1 2 4 |
| медлительность | 18 6 12 144 |
| нерешительность | 7 11 - 4 16 |
| энергичность | 20 12 8 64 |
| жизнерадостность | 19 8 11 121 |
| мнительность | 4 3 1 1 |
| упрямство | 5 9 - 4 16 |
| беспечность | 11 14 - 3 9 |
| застенчивость | 6 5 1 1 |
|  |  1056 |
|  |

Значение коэффициента корреляции Спирмена подсчитываем по формуле

6 ⋅ 1056

rS = 1 − 3 = 0, 206.

20 − 20

Вследствие малого n (меньше 30) гипотезу о значимости коэффициента корреляции проверяем с помощью статистических таблиц. Для n = 20 имеем (см. табл.3.1):

h0? h1

⎯⎯|⎯⎯⎯⎯⎯|⎯⎯⎯⎯⎯⎯|⎯⎯⎯⎯⎯⎯→

0, 206 0,445 0,568 r

Значение коэффициента корреляции rS = 0, 206 попадает в область допустимых

значений, что не позволяет отвергнуть нулевую гипотезу. Коэффициент корреляции не отличается от нуля.

Вывод:

Отсутствует связь между выраженностью качеств у обследуемого испытуемого в данный момент и идеальным представлением.

## Заключение

В нашей работе мы рассмотрели методы математической обработки экспериментальных данных. Также выполнили практическую часть, произвели расчеты по методу ранговой корреляции Спирмена, высчитали коэффициент корреляции Спирмена между выраженностью качеств у обследуемого испытуемого в данный момент и его идеальным представлением. На основе расчетов, сделали вывод об отсутствии связи между выраженностью качеств у обследуемого испытуемого в данный момент и идеальным представлением.

Психология - это наука, которая исследует, наблюдает, анализирует. Она постоянно ищет свой путь в выявлении новых закономерностей и фактов. Математические методы обработки данных оказывают на этом пути дают неоценимую помощь. Они используют математические приемы, формулы, способы качественных расчетов, с помощью которых показатели, получаемые в ходе эксперимента, можно обобщать, приводить в систему, выявляя скрытые в них закономерности.

## Литература

1. Сидоренко Е.В. Методы математической обработки в психологии/Е.В. Сидоренко. - СПб.: - 2002.

2. Немов Р.С. Психология / Р.С. Немов. - М.: - 2005. - 630

3. Харченко М.А. Корреляционный анализ: Учебное пособие для ВУЗов / Л.М. Носилова. - Воронеж.: - 2008.