**Содержание**

Введение……………………………………………………………..…..3

ГЛАВА 1. Актуальность изучения методов статистической обработки результатов измерений………………………………………………………...6

1.1. Статистический метод в психологии……………………….……..6

1.2. Статистическое распределение……………………………….……9

Глава 2. Использование статистического анализа в интерпретации результатов методики изучения агрессии Басса-Дарки……………………..11

2.1. Исследование уровня и рода враждебности школьников….….11

2.2. Построение регрессионной модели……………………………..14

2.3. Анализ регрессионной модели………………………………..…19

Заключение……………………………………………………………….31

Приложения……………………………………………………………….35

**Введение**

Теоретические методы исследования в науке дают возможность раскрыть качественные характеристики изучаемых явлений. Эти характеристики будут полнее и глубже, если накопленный эмпирический материал подвергнуть количественной обработке. Однако проблема количественных измерений, в частности, в рамках психолого-педагогических исследований очень сложна. Эта сложность заключается, прежде всего, в субъективно-причинном многообразии педагогической деятельности и ее результатов, в самом объекте измерения, находящемся в состоянии непрерывного движения и изменения. Вместе с тем введение в исследование количественных показателей стало сегодня необходимым и обязательным компонентом получения объективных данных о результатах труда. С этой целью при исследовании проблем психологии применяются методы математической статистики. С их помощью решаются различные задачи: обработка фактического материала, получение новых, дополнительных данных, обоснование научной организации исследования и др.

Правильное применение статистики позволяет экспериментатору:

• строить статистические предсказания;

• обобщать данные эксперимента;

• находить зависимость между экспериментальными данными;

• строго обосновывать экспериментальные планы;

• доказывать правильность и обоснованность используемых методических приемов и методов.

Нельзя забывать, однако, что сами по себе методы статистики – это только инструментарий, помогающий экспериментатору эффективно разбираться в сложном исследуемом материале. Наиболее важным при проведении любого эксперимента является четкая постановка задачи, тщательное планирование эксперимента, построение непротиворечивых гипотез.

Методы математической статистики в руках исследователя могут и должны быть мощным инструментом, позволяющим не только успешно лавировать в море экспериментальных данных, но и, прежде всего, способствовать становлению его объективного мышления.

Актуальность данного исследования означена востребованностью статистической обработки экспериментальных данных в психолого- педагогических исследованиях.

Цель: проведение регрессионного анализа статистических данных психологического эксперимента для выявления уровня враждебности школьников в зависимости от уровней обиды и подозрительности (диагностика состояния враждебности Басса-Дарки).

Объект исследования: процесс статистической обработки данных психологического эксперимента.

Предмет исследования: зависимость уровня враждебности от таких психологических факторов личности как обида и подозрительность.

Задачи:

1. Проанализировать научную, учебную, специальную литературу по теме исследования;

2. Изучить теоретические аспекты разновидностей регрессионного анализа;

 3. Выявить методы и средства статистического анализа данных психологического эксперимента;

4. Обработать статистические данные с помощью специальных функций, встроенных в табличный процессор Excel;

5. Провести аппроксимацию данных проведенного эксперимента.

Для решения поставленных задач используются следующие методы:

1. Теоретические:

• анализ литературы;

• систематизация изученного материала;

• обобщение.

2. Эмпирические:

• наблюдение;

• анкетирование(опрос).

**ГЛАВА 1. Актуальность изучения методов статистической обработки результатов измерений**

**1.1. Статистический метод в психологии**

При изучении психических явлений психология не может ограничиться исследованием единичных фактов, как бы интересны они сами по себе ни были. Психические процессы — массовое явление, и поэтому присущие им закономерности могут быть выявлены лишь путем исследования массовых фактов. Например, нельзя сделать научно обоснованные выводы об особенностях мышления детей данного возраста или о процессе формирования двигательных навыков в данном виде спорта, опираясь на исследование этих процессов только у одного какого-нибудь испытуемого. Только исследование достаточного числа лиц позволяет делать выводы, характеризующие действительные, а не случайные особенности и закономерности изучаемого явления. Необходимость изучать массовые явления заставляет психологию прибегать к статистическим методам исследования [5].

Статистические методы применяются при анализе полученных в исследовании материалов. Однако правильное применение методов статистики выдвигает серьезные требования и к методике собирания материала в процессе исследования. Обоснованные статистические выводы могут быть сделаны только при наличии достаточного количества подмеченных фактов и при достаточной однородности собранного материала.

Статистический метод опирается на законы больших чисел, а это в ряде случаев требует не менее ста наблюдений.Полученные при статистической обработке материала выводы могут отличаться различной степенью вероятности (объективной возможности повторения или обнаружения при данных условиях). Вероятность сделанного вывода зависит от количества произведенных наблюдений и колеблется от 0 (невозможность данного факта) до 1 (полная достоверность, обязательное наступление при определенных условиях). Математически выведенная зависимость степени вероятности вывода от числа наблюдений представлена в следующей таблице.

Статистические методы исследования теряют свое научное значение, если они применяются для обработки качественно неоднородного материала, например если исследователь будет объединять в одну группу данные, полученные при изучении лиц разного возраста, разной квалификации, разных специальностей и т. д. Статистические методы требуют предварительной группировки собранных фактов по качественно однородным признакам.

При статистических методах чаще всего применяют следующие способы обработки материала [7].

1. Кривая распределения. Собранный материал располагают в определенном порядке, по степени возрастания или убывания полученных показателей. Подсчитывают частоту каждого показателя и полученные данные изображают в виде диаграммы, где на абсциссе в последовательном порядке отмечаются показатели, а на ординате — частота каждого показателя. Полученные кривые наглядно характеризуют степень однородности собранного материала и типичное для данной группы наблюдений распределение показателей по их частоте; в частности, по вершине кривой, соответствующей значению наиболее частого показателя, судят о характерной для данного явления закономерности.



Рисунок 1. Кривая частоты времени сложной реакции

2. Среднее арифметическое получается путем деления суммы всех полученных измерений на их число.

Сравнивая средние арифметические, полученные при исследовании разных групп испытуемых, можно судить о преимущественном значении для них того или другого признака. Например, при измерении скорости реакции у лыжников и боксеров получены соответственно следующие средние показатели латентного периода сложной реакции: 0,357 и 0,293 сек. Сравнение этих показателей позволяет прийти к выводу, что занятия боксом в большей степени содействуют формированию способности быстрого реагирования в сложных условиях, чем занятия лыжным спортом. Этот вывод имеет значение общей статистической закономерности, характеризующей явление в целом, несмотря на то, что у отдельных боксеров скорость реакции может быть значительно меньшей, чем у лыжников [14].

3. Мера точности измерения, определяемая путем вычисления среднего квадратического отклонения. Для этого определяют отклонение каждого отдельного измерения от среднего арифметического, полученные отклонения возводят в квадрат и из их общей суммы вычисляют квадратный корень.

Чем больше среднее квадратичное отклонение, тем не однороднее состав испытуемых, тем менее точно произведено исследование. Чем оно меньше, тем достовернее полученные исследователем результаты.

Правильное применение статистических методов необходимо при всех психологических исследованиях массового характера [12].

**1.2. Статистическое распределение**

Любая совокупность данных характеризуется какими-либо параметрами. Нечисловые совокупности представлены набором качественных оценок исследуемого явления и имеют, по крайней мере, один параметр – объем совокупности (выборки). Кроме того, нечисловые совокупности обычно структурированы – в них представлены, по крайней мере, два типа данных (дихотомические совокупности) или более. Поэтому в качестве параметров совокупности выступают выборочные оценки представленности (встречаемости, вероятности появления и т.п.) того или иного типа данных в выборке. Другими словами, даже в нечисловых совокупностях целесообразно говорить о статистическом распределении результатов измерений.

Числовые совокупности характеризуются значительно большим количеством параметров. Это значения каждого измерения (интенсивность или степень выраженности признака), объем выборки, размах выборки, среднее арифметическое, мода, асимметрия, эксцесс и т.д. Важным параметром выборки является совокупность интервальных частот встречаемости, или распределение величины [3].

Любое исследование всегда ограничено количеством измерений (объёмом выборки). При этом результаты анализа выборочных данных должны адекватно отражать особенности изучаемого объекта, т.е. выборочная совокупность должна быть репрезентативна по отношению к генеральной совокупности, типична для неё. Основная цель исследования – сделать заключение по имеющейся выборке обо всей совокупности, из которой она взята. Так, например, учащиеся 4 «а» класса конкретной школы являются представителями специфической возрастной группы младших подростков, вовлечённых в процесс социализации методами современной педагогики. Исследуя то или иное психическое явление на примере этого класса, с определёнными оговорками можно выдвинуть предположения о психологических характеристиках подростков этого возраста.

Процесс формирования прогноза особенностей выборки, выполненный на основе знания свойств генеральной совокупности, можно назвать прямой задачей психологии. Необходимым условием возможности решения прямой задачи является наличие теории явления, перекрывающей пределами своего применения всю совокупность независимых переменных, характерных для рассматриваемой выборки. Выявление свойств генеральной совокупности на основе знания особенностей выборочной совокупности представляет собой обратную задачу психологии. Большинство задач практической психодиагностики может быть отнесено именно к этому типу задач [5].

**Глава 2. Использование статистического анализа в интерпретации результатов методики изучения агрессии Басса-Дарки**

**2.1. Исследование уровня и рода враждебности школьников**

Статистические методы раскрывают связи между изучаемыми явлениями. Однако необходимо твердо знать, что, как бы ни была высока вероятность таких связей, они не дают права исследователю признать их причинно-следственными отношениями.

Чтобы подтвердить или отвергнуть существование причинно-следственных отношений, исследователю зачастую приходится продумывать целые серии экспериментов. Если они будут правильно построены и проведены, то статистика поможет извлечь из результатов этих экспериментов информацию, которая необходима исследователю, чтобы либо обосновать и подтвердить свою гипотезу, либо признать ее недоказанной [7].

В работе с подростковой аудиторией педагогу и психологу всегда приходится учитывать особенности агрессии у подростков. А для выявления уровня и рода агрессии детей существуют различные методики. Одна из них – диагностика состояния агрессии (опросник Басса-Дарки). Данный опросник состоит из 75 утверждений, на которые испытуемый отвечает "да" или "нет".

Создавая свой опросник, дифференцирующий проявления агрессии и враждебности, А. Бассе и А. Дарки выделили следующие виды реакций:

1. Физическая агрессия – использование физической силы против другого лица.

2. Косвенная агрессия – агрессия, окольным путем направленная на другое лицо или ни на кого не направленная.

3. Раздражение – готовность к проявлению негативных чувств при малейшем возбуждении (вспыльчивость, грубость).

4. Негативизм – оппозиционная манера в поведении от пассивного сопротивления до активной борьбы против установившихся обычаев и законов.

5. Обида – зависть и ненависть к окружающим за действительные и вымышленные действия.

6. Подозрительность – в диапазоне от недоверия и осторожности по отношению к людям до убеждения в том, что другие люди планируют и приносят вред.

7. Вербальная агрессия – выражение негативных чувств как через форму (крик, визг), так и через содержание словесных ответов (проклятия, угрозы).

8. Чувство вины – выражает возможное убеждение субъекта в том, что он является плохим человеком, что поступает зло, а также ощущаемые им угрызения совести.

Обработка результатов: Обработка опросника Басса-Дарки производится при помощи индексов различных форм агрессивных и враждебных реакций, которые определяются суммированием полученных ответов. Физическая агрессия, косвенная агрессия, раздражение и вербальная агрессия вместе образуют суммарный индекс агрессивных реакций, а обида и подозрительность – индекс враждебности.

 Данная методика была апробирована 28.10.15 г. в 9а классе МАОУ СОШ № 5 г. Москва. В исследовании приняли участие 20 учащихся. Результаты опроса (значения параметров) представлены в сводной таблице (Приложение 1).

Для полной реализации сути опросника Басса-Дарки необходимо представить суммарный индекс агрессивных реакций и суммарный индекс враждебности (Приложение 2).

Перед началом регрессионного анализа осуществляется отбор факторов. Сначала отбираются факторы, связанные с изучаемым явлением, на основе данных теоретического исследования (психологическая теория, заключения экспериментатора и т.д.). При этом для построения множественной регрессии отбираются факторы, которые могут быть количественно измерены.

Проблему данного исследования составило рассмотрение и анализ уровня враждебности, вследствие этого регрессионный анализ экспериментальных данных методики Басса-Дарки будет проведен по индексу враждебности (зависимая переменная y), получающийся суммированием выявленных уровней обиды и подозрительности (независимые переменные x и z, соответственно).

**2.2. Построение регрессионной модели**

Регрессионный анализ экспериментальных данных методики Басса- Дарки будет проведен по индексу враждебности (зависимая переменная y), получающийся суммированием выявленных уровней обиды и подозрительности (независимые переменные x и z, соответственно).

Как будет варьировать индекс враждебности испытуемого, если будут изменяться уровни обиды и подозрительности? Ответ на этот вопрос психолог получит с помощью использования метода множественной регрессии. Данные для анализа представлены в таблице 1, в которой произведены предварительные вычисления.

Таблица 1. Исходные данные

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| № | Фамилия ученика | *xi* | *zi* | *yi* | *xi \** *xi* | *xi* \**zi* | *xi* \* *yi* | *zi* \* *zi* | *yi* \* *zi* |
| 1 | Бакиева | 5 | 8 | 13 | 25 | 40 | 65 | 64 | 104 |
| 2 | Гатауллин | 1 | 4 | 5 | 1 | 4 | 5 | 16 | 20 |
| 3 | Гатин | 2 | 2 | 4 | 4 | 4 | 8 | 4 | 8 |
| 4 | Долженко | 5 | 4 | 9 | 25 | 20 | 45 | 16 | 36 |
| 5 | Жарова | 4 | 7 | 11 | 16 | 28 | 44 | 49 | 77 |
| 6 | Жуйкова | 6 | 3 | 9 | 36 | 18 | 54 | 9 | 27 |
| 7 | Корикова | 5 | 7 | 12 | 25 | 35 | 60 | 49 | 84 |
| 8 | Костерина | 7 | 7 | 14 | 49 | 49 | 98 | 49 | 98 |
| 9 | Курманалиева | 4 | 7 | 11 | 16 | 28 | 44 | 49 | 77 |
| 10 | Летунов | 3 | 2 | 5 | 9 | 6 | 15 | 4 | 10 |
| 11 | Мороков | 4 | 5 | 9 | 16 | 20 | 36 | 25 | 45 |
| 12 | Перовских В. | 4 | 9 | 13 | 16 | 36 | 52 | 81 | 117 |
| 13 | Перовских М. | 4 | 7 | 11 | 16 | 28 | 44 | 49 | 77 |
| 14 | Смирнова | 4 | 8 | 12 | 16 | 32 | 48 | 64 | 96 |
| 15 | Солосина | 7 | 8 | 15 | 49 | 56 | 105 | 64 | 120 |
| 16 | Тимирова | 1 | 2 | 3 | 1 | 2 | 3 | 4 | 6 |
| 17 | Трухин | 2 | 4 | 6 | 4 | 8 | 12 | 16 | 24 |
| 18 | Филиппов | 4 | 6 | 10 | 16 | 24 | 40 | 36 | 60 |
| 19 | Хабисов | 6 | 3 | 9 | 36 | 18 | 54 | 9 | 27 |
| 20 | Цыпанов | 0 | 2 | 2 | 0 | 0 | 0 | 4 | 4 |
| **Суммы:** | **78** | **105** | **183** | **376** | **456** | **832** | **621** | **1117** |

С помощью решения системы уравнений (3.1) необходимо найти уравнение регрессии y на x, т.е. определить коэффициенты a, b и c, и таким образом ответить на поставленный вопрос.

Чтобы получить и решить уравнение множественной линейной регрессии (3.1), необходимо найти a, b и c. Для этого используется система уравнений (3.4). Благодаря вычислениям, приведенным в таблице 3, известны все необходимые величины сумм. Перепишем систему уравнений (3.4), учитывая N = 20, поскольку в эксперименте участвовало 20 человек, и учитывая данные таблицы 3:



Получили систему линейных уравнений (СЛУ) с тремя неизвестными. Решается данная система несколькими способами: по правилу Крамера, методом Гаусса и с помощью обратной матрицы.

В СЛУ (3.8) число уравнений равно числу неизвестных, поэтому целесообразно для нахождения неизвестных применить метод Крамера. Для начала составляется матрица третьего порядка:



Здесь последний столбец – это столбец свободных членов.

*Теорема* ***(правило Крамера).*** Пусть ∆ – определитель матрицы СЛУ, а

∆ *j* - определитель, полученный из определителя ∆ заменой *j*-го столбца

столбцом свободных членов. Тогда если ∆ ≠0 , то система линейных

уравнений имеет единственное решение, определяемое по формулам:

*x j* =∆ *j*

∆ , где *j* = 1,2,…,*n*

Формулы вычисления неизвестных (3.10) – решения системы линейных уравнений (3.8) – носят название *формул Крамера.*

Составляется и вычисляется главный определитель матрицы:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 20 | 78 | 105 |  |
|   | 78 | 376 | 456 |  56916 |
|  | 105 | 456 | 621 |  |

Так как вычисления данного определителя очень громоздкие, то целесообразно осуществлять все расчеты с помощью «Мастера функций» MS Excel. Для этого используется встроенная математическая функция МОПРЕД. Порядок вычисления следующий:

Так как вычисления данного определителя очень громоздкие, то целесообразно осуществлять все расчеты с помощью «Мастера функций» MS Excel. Для этого используется встроенная математическая функция МОПРЕД. Порядок вычисления следующий:

1) введите в упорядоченные ячейки электронной таблице исходные элементы определителя, сохраняя порядок следования элементов;

2) активируйте Мастер функций любым из способов:

а) в главном меню выберите команду Вставка/Функция;

 б) на панели инструментов Стандартная щелкните по кнопке Вставка функции;

3) в появившемся диалоговом окне «Мастер функций – шаг 1 из 2» в поле Категории выберите Математические, в окне Функция – МОПРЕД. Щелкните по кнопке ОК;

4) в появившемся окне Аргументы функции необходимо указать диапазон ячеек от первого элемента исходного определителя до последнего (например, А1:С3);

5) щелкните по кнопке ОК.

После выполнения данного алгоритма на экране компьютера появится результат – определитель.

Как видно, полученный определитель (∆ = 56916) отличен от нуля, стало быть, СЛУ (3.8) имеет единственное решение, которое вычисляется по формулам:



Чтобы применить формулы, необходимо составить определители ∆ 1, ∆ 2 , ∆ 3 по правилу Крамера и произвести их расчеты с помощью

«Мастера функций» MS Excel. Все расчеты представлены ниже»



Теперь, когда известны все определители, можно применить формулы:



Решив систему уравнений, получилось a = - 3,34, b = 1,82, c = 1,02. следовательно, искомое уравнение регрессии y на x примет вид:



где y – зависимая переменная, –3,34 - свободный член, 1,82 и 1,02 – параметры уравнения.

Уравнение дает ответ на поставленный ранее вопрос: Как будет варьировать индекс враждебности испытуемого, если будут изменяться уровни обиды и подозрительности? Так, при увеличении величины уровня обиды x на 1 балл, количественная величина индекса враждебности y увеличится на 1,82, при постоянной величине уровня подозрительности z. А при постоянной величине уровня обиды и при увеличении величины уровня подозрительности на 1 балл количественная величина индекса враждебности увеличится в среднем на 1,02 балла.

Полученное уравнение многофакторной регрессии имеет еще одно приложение. Так, подставляя в него значения переменных x и z, можно определить ожидаемую величину переменной y (уровня враждебности) [11].

**2.3. Анализ регрессионной модели**

В предыдущем параграфе была вычислена модель множественной регрессии (3.14): y = -3,34 + 1,82 x + 1,02 \* z ,

где y – значение зависимой переменной,

x и z – значения зависимых переменных,

–3,34 – свободный член,

1,82 и 1,02 – параметры уравнения (коэффициенты при независимых переменных).

Для многофакторной регрессионной модели имеют место следующие предпосылки:

1) Зависимые переменные – величины неслучайные;

2) Математическое ожидание случайной составляющей в любом наблюдении равно нулю:



3) Дисперсия случайной составляющей постоянна для всех наблюдений:



4) Отсутствие систематической связи между значениями случайной составляющей в любых двух наблюдениях:





Факторы, включенные во множественную регрессию, количественно измерены и не сильно коррелируют друг с другом (корреляция – связь между собой двух и более переменных в одной или нескольких изучаемых группах). Кроме того, каждый фактор тесно связан с результатом.

Многофакторная регрессия представляет регрессию результативного признака с двумя и большим числом независимых переменных вида:



В уравнении регрессии случайная (зависимая) переменная y зависит не только от значений независимых переменных x и z, но и от ряда други факторов, влияющих на y, которые не могут быть проконтролированы.

В связи с этим случайная величина, характеризующая отклонения результативного признака от теоретического, найденного по уравнению регрессии.

При исследовании зависимости результативного признака y в многофакторной модели необходимо решать такие же задачи, что и при однофакторной модели:

• определение вида регрессии;

• оценка параметров;

• определение тесноты связи.

Однако наряду с этими задачами необходимо рассматривать и ряд задач, характерных лишь для многофакторной регрессии.

К таким задачам относится отбор факторов, существенно влияющих на фактор y, при наличии возможностей внутренней взаимосвязи между зависимыми переменными x и z. Такой отбор требует, прежде всего, глубокого теоретического и практического знания качественной стороны рассматриваемых психологических явлений.

Интерпретация результатов

До сих пор мы употребляли абстрактный математический язык. Перевод модели на язык экспериментатора называется интерпретацией модели. Задача интерпретации весьма сложна.

Устанавливается, в какой мере каждый из факторов влияет на параметр оптимизации. Величина коэффициента регрессии – количественная мера этого влияния. Чем больше коэффициент, тем сильнее влияет фактор. О характере влияния факторов говорят знаки коэффициентов. Знак плюс свидетельствует о том, что с увеличением значения фактора растет величина параметра оптимизации, а при знаке минус – убывает.

Анализируя сущность уравнения регрессии, следует отметить следующие положения. Рассмотренный подход не обеспечивает раздельной (независимой) оценки коэффициентов – изменение значения одного коэффициента влечет изменение значений других. Полученные коэффициенты не следует рассматривать как вклад соответствующего параметра в значение показателя. Уравнение регрессии является всего лишь хорошим аналитическим описанием имеющихся экспериментальных данных, а не законом, описывающим взаимосвязи параметров и показателя. Это уравнение применяют для расчета значений показателя в заданном диапазоне изменения параметров. Оно ограниченно пригодно для расчета вне этого диапазона, т.е. его можно применять для решения задач интерполяции и в ограниченной степени для экстраполяции.

 В настоящее время регрессионный анализ широко используется в дифференциальной психологии и психодиагностике. С его помощью можно разрабатывать тесты, устанавливать структуру связей между отдельными психологическими характеристиками, измеряемыми набором тестов или заданиями теста.

Регрессионный анализ используется также для стандартизации тестовых методик, которая проводится на репрезентативной выборке испытуемых.

**2.4. Аппроксимация экспериментальных данных**

На практике часто приходится сталкиваться с задачей сглаживания экспериментальных зависимостей или задачей аппроксимации.

Аппроксимацией называется процесс подбора эмпирической формулы  для установленной из опыта функциональной зависим Эмпирические формулы служат для аналитического представления опытных данных.

Другими словами, аппроксимация, или приближение – это научный метод, состоящий в замене одних объектов другими, в том или ином смысле близкими к исходным, но более простыми. Аппроксимация позволяет исследовать числовые характеристики и качественные свойства объекта, сводя задачу к изучению более простых или более удобных объектов (например, таких, характеристики которых легко вычисляются, или свойства которых уже известны).

Одна независимая переменная.

Обычно задача аппроксимации распадается на две части. Сначала устанавливают вид зависимости  и, соответственно, вид эмпирической формулы, т.е. решают, является ли она линейной, квадратичной, логарифмической или какой-либо другой. После этого определяются численные значения неизвестных параметров выбранной эмпирической формулы, для которых приближение к заданной функции оказывается наилучшим. Если нет каких-либо теоретических соображений для подбора вида формулы, обычно выбирают функциональную зависимость из числа наиболее простых, сравнивая их графики с графиком заданной функции. После выбора вида формулы определяют ее параметры. Для наилучшего выбора параметров задают меру близости аппроксимации экспериментальных данных. Во многих случаях, в особенности, если функция  задана графиком или таблицей (на дискретном множестве точек), для оценки степени приближения рассматривают разности .

Существуют различные меры близости и, соответственно, способы решения этой задачи. Некоторые из них очень просты, быстро приводят к результату, но результат этот является сильно приближенным, другие более точными, но более сложными. Обычно определение параметров при известном виде зависимости осуществляют по методу наименьших квадратов. При этом функция  считается наилучшим приближением к, если для нее сумма квадратов отклонений  «теоретических» значений , найденных по эмпирической формуле, от соответствующих опытных значений yi , имеет наименьшее значение по сравнению с другими функциями, из числа которых выбирается искомое приближение.

Используя методы дифференциального исчисления, метод наименьших квадратов формулирует аналитические условия достижения суммой квадратов отклонений своего наименьшего значения.

В простейшем случае задача аппроксимации экспериментальных данных выглядит следующим образом.

Пусть экспериментальные данные, полученные практическим путем, которые можно представить парами чисел зависимость между которыми отражает таблица.

На основе данных требуется подобрать функцию, которая наилучшим образом сглаживала бы экспериментальную зависимость между переменными и, по возможности, точно отражала общую тенденцию зависимости между x и y, исключая погрешности измерений и случайные отклонения. Это значит, что отклонения бы наименьшими. в каком-то смысле были бы наименьшими.

 Выяснить вид функции можно либо из теоретических соображений, либо анализируя расположение точек на координатной плоскости.

 Расположение экспериментальных точек может иметь самый различный вид, и каждому соответствует конкретный тип функции.

Построение эмпирической функции сводится к вычислению входящих в нее параметров, так чтобы их всех функций такого вида выбрать ту, которая лучше других описывает зависимость между изучаемыми величинами. То есть сумма квадратов разности между табличными значениями функции в некоторых точках и значениями, вычисленными по полученной формуле, должна быть минимальна.

Степень близости аппроксимации экспериментальных данных выбранной функцией оценивается коэффициентом детерминации (R2). Он показывает, какая доля дисперсии результативного признака объясняется влиянием независимых переменных. Таким образом, если есть несколько подходящих вариантов типов аппроксимации функций, можно выбрать функцию с большим коэффициентом детерминации (стремящимся к 1).

Таблица 2. Показатели тесноты связи

|  |  |
| --- | --- |
| Количественная мера тесноты связи | Качественная характеристика силы связи |
| 0,1-0,3 | Слабая |
| 0,3-0,5 | Умеренная |
| 0,5-0,7 | Заметная |
| 0,7-0,9 | Высокая |
| 0,9-0,99 | Весьма высокая |

Таким образом, функциональная связь возникает при значении равном 1,

а отсутствие связи – 0. При значениях показателей тесноты связи меньше 0,7 величина коэффициента детерминации всегда будет ниже 50%. Это означает, что на долю вариации факторных признаков приходится меньшая часть по сравнению с остальными неучтенными в модели факторами, влияющими на изменение результативного показателя. Построенные при таких условиях регрессионные модели имеют низкое практическое значение.

В MS Excel аппроксимация экспериментальных данных осуществляется путем построения графика – линии тренда (x, y – заданные величины).

Тренд – тенденция изменения показателей временного ряда. Тренды могут быть описаны различными функциями. Тип тренда устанавливают на основе данных временного ряда, путем осреднения показателей динамики ряда, на основе статистической проверки гипотезы о постоянстве параметров графика. Возможны следующие варианты функций:

1. Линейная – . Обычно применяется в простейших случаях, когда экспериментальные данные возрастают или убывают с постоянной скоростью.

2. Полиномиальная – 

где до шестого порядка включительно - константы. Используется для описания экспериментальных данных, попеременно возрастающих и убывающих. Степень полинома определяется количеством экстремумов (максимумов и минимумов) кривой. Полином второй степени может описать только один максимум или минимум, полином третей степени может дать один или два экстремума, четвертой степени – не более трех экстремумов и т.д.

3. Логарифмическая –– константы, ln x – функция натурального логарифма. Функция применяется для описания экспериментальных данных, которые вначале быстро растут или убывают, а затем постепенно стабилизируются.

 4. Степенная –– константы. Аппроксимация степенной функцией используется для экспериментальных данных с постоянно увеличивающейся (или убывающей) скоростью роста. Данные не должны иметь нулевых или отрицательных значений.

 5. Экспоненциальная – – константы, e – основание натурального логарифма. Применяется для описания экспериментальных данных, которые быстро растут или убывают, а затем постепенно стабилизируются. Часто ее использование вытекает из теоретических соображений.

Для осуществления аппроксимации на диаграмме экспериментальных данных необходимо щелчком правой кнопки мыши вызвать всплывающее меню и выбрать пункт Добавить линию тренда. В появившемся диалоговом окне Линия тренда на вкладке Тип выбирается вид аппроксимирующей функции, а на вкладке Параметры устанавливаются флажки в полях показывать уравнение на диаграмме и поместить на диаграмму величину достоверности аппроксимации (R2). После чего нужно щелкнуть по кнопке ОК. В результате получим на диаграмме аппроксимирующую кривую.

Проделав данную операцию несколько раз, можно представить линейную зависимость индексов с уравнениями линий тренда и коэффициентом детерминации – для анализа полной достоверности результатов исследуемых показателей (рис.2, рис.3).

Зависимость индекса враждебности от индекса обиды

20

18

16

14

12

10

8

6

4

2

0

0 1 2 3 4 5 6 7

Yi Линейный (Yi) Полиномиальный (Yi) Экспоненциальный (Yi)

Рис.2 Зависимость индекса враждебности от индекса обиды

***Зависимость индекса враждебности от индекса подозрительности***

18

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |
|  |
|  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |

16



14 

12

10

 

8

6

4

2

0

0 1 2 3 4 5 6 7 8 9

Yi Линейный (Yi) Полиномиальный (Yi) Экспоненциальный (Yi)

Рис.3 Зависимость индекса враждебности от индекса подозрительности

Как видно, из рис.2, Несколько независимых переменных

В тех случаях, когда аппроксимируемая переменная y зависит от нескольких независимых переменных x1 , x2 ,..., xn , подход с построением линии тренда не дает решения. Здесь могут быть использованы следующие специальные функции MS Excel:

ЛИНЕЙН и ТЕНДЕНЦИЯ для аппроксимации линейных функций вида:



ЛГРФПРИБЛ и РОСТ для аппроксимации показательных функций вида:



Функции ЛИНЕЙН и ЛГРФПРИБЛ служат для вычисления неизвестных коэффициентов a0 , a1 , a2 ,..., an , в указанных выражениях, а также коэффициентов детерминации ( R2 ). Обе функции имеют одинаковые параметры:

ЛИНЕЙН (известные значения y; известные значения x; конст; статистика);

ЛГРФПРИБЛ (известные значения y; известные значения x; конст; статистика).

Здесь:

* известные значения y – множество наблюдаемых значений y из казанных выражений;
* известные значения x – множество наблюдаемых значений

 x1 , x2 ,..., xn .

Причем, если массив известные значения y имеет один столбец, то каждый столбец массива известные значения x интерпретируется как отдельная переменная, а если массив известные значения y имеет одну строку, то тогда каждая строка массива известные значения x интерпретируется как отдельная переменная;

* конст – логическое значение, которое указывает, требуется ли, чтобы константа a0 была равна 0 (для функции ЛИНЕЙН) или 1 (для функции ЛГРФПРИБЛ). При этом, если конст имеет значение

ИСТИНА или опущено, то a0 вычисляется обычным образом, а если конст имеет значение ЛОЖЬ, то a0 полагается равным 0 или 1;

* статистика – логическое значение, которое указывает, требуется ли вычислять дополнительную статистику по регрессии, если введено значение ИСТИНА, то дополнительные параметры вычисляются, если ЛОЖЬ, то – нет.

Функции ТЕНДЕНЦИЯ и РОСТ позволяют находить точки, лежащие на аппроксимирующих кривых, для значений коэффициентов найденных функциями ЛИНЕЙН и ЛГРФПРИБЛ.

Обе функции имеют одинаковые аргументы:

 a0 , a1 , a2 ,..., an ,

 ТЕНДЕНЦИЯ (известные значения y; известные значения x; новые значения x; конст);

РОСТ (известные значения y; известные значения x; новые значения x; конст).

Здесь:

* известные значения y – множество значений y;
* известные значения x – множество значений x;
* новые значения x – те значений x, для которых необходимо определить соответствующие аппроксимирующие или предсказанные значения y. Новые значения x должны содержать столбец (или строку) для каждой независимой переменной, как и известные значения x. Если аргумент новые значения x опущен, то предполагается, что он совпадает с аргументом известные значения x;
* конст – логическое значение, которое указывает, требуется ли, чтобы константа a0 была равна 0 (для функции ТЕНДЕНЦИЯ) или 1 (для функции РОСТ). При этом, если конст имеет значение ИСТИНА или опущено, то a0 вычисляется обычным образом, а если конст имеет значение ЛОЖЬ, то a0 полагается равным 0 или 1.

**Заключение**

Тщательное, скурпулезное проведение эксперимента, несомненно, является главным условием успеха исследования. Это общее правило, и планирование эксперимента не относится к исключениям.

Однако экспериментатору не безразлично, как обработать полученные данные. Необходимо извлечь из них всю информацию и сделать соответствующие выводы. С одной стороны, не извлечь из эксперимента все, что из него следует, - значит пренебречь нелегким трудом экспериментатора. С другой стороны, сделать утверждения, не следующие из эксперимента, - значит создавать иллюзии, заниматься самообманом. Статистические методы обработки результатов эксперимента позволяют не перейти разумной меры риска.

Если данные, полученные в эксперименте, качественного характера, то правильность делаемых на основе их выводов полностью зависит от интуиции, эрудиции и профессионализма исследователя, а также от логики его рассуждений. Если же эти данные количественного типа, то сначала проводят их первичную, а затем вторичную статистическую обработку.

Вторичная статистическая обработка проводится в том случае, если для решения задач или доказательства предложенных гипотез необходимо определить статистические закономерности, скрытые в первичных экспериментальных данных. Приступая к вторичной статистической обработке, исследователь, прежде всего, должен решить, какие из различных вторичных статистик ему следует применить для обработки первичных экспериментальных данных.

 Таким образом, реализована цель данной работы, т.е. разработана методика проведения регрессионного анализа статистических данных психологического эксперимента для прогнозирования исследуемых показателей. Это было достигнуто через реализацию всех поставленных задач с помощью теоретических и эмпирических методов. Таких как анализ различной литературы, систематизация полученной информации (знаний) и ее обобщение; наблюдение и анкетирование (опрос).

Математическая статистика – прикладная отрасль математики, основанная на теории вероятностей и предназначенная в самом общем плане для систематизации и анализа эмпирических (опытных) данных, получаемых при изучении повторяющихся и варьирующихся явлений.

Планирование и анализ экспериментов – это раздел математической статистики, включающий систему методов обнаружения и проверки причинных связей между переменными.

Таким образом, математическая статистика – это точная и полезная наука. Но лишь для думающего исследователя, не пренебрегающего необходимостью вникнуть в существо идей и методов теории вероятностей и математической статистики.

В целом же, статистические методы помогают исследователям описывать данные, делать выводы в отношении больших массивов данных и изучать причинные зависимости.

**Список использованных источников**

1. Дуброва, Т.А. Статистиские методы прогнозирования: Учебное пособие

[Текст]/ Т.А. Дуброва. – М.: ЮНИТИ, 2003. – 204с.

2. Ермолаев О.Ю. Математическая статистика для психологов : Учебник [Текст]/ О.Ю. Ермолаев. – М.: Изд-во Флинта Московского психолого- социального института, 2004. – 335с.

 3. Калинина, В.Н. Теория вероятностей и математическая статистика: Учебное пособие для вузов [Текст]/ В.Н. Калинина. – М.: Дрофа, 2008. – 471с.

4. Калинина, В.Н. Математическая статистика: Учебник для студентов

[Текст]/ В.Н. Калинина, В.Ф. Панкин. – М.: Дрофа, 2002. – 335с.

5. Крамер, Д. Математическая обработка данных в социальных науках: современные методы: Учебное пособие для вузов [Текст]/ Дункан Крамер. – Академия, 2007. – 287с.

6. Красс, М.С. Математика для экономического бакалавриата: Учебник

[Текст]/ М.С. Красс, Б.П. Чупрынов. – М.: Дело, 2005. – 574с.

7. Кричевец, А.Н. Математика для психологов: Учебник [Текст]/ А.Н. Кричевец, Е.В. Шикин, А.Г. Дьячков. – М.: Изд-во Флинта Московского психолого-социального института, 2005. – 371с.

8. Могилев, А.В, Информатика: Учебник [Текст]/ А.В. Могилев, Н.И. Пак, Е.К. Хеннер. – М.: Академия, 2003. – 809с.

9. Немов, Р.С. Психодиагностика. Введение в научное психологическое исследование с элементами математической статистики [Текст]/ Р.С. Немов.

– М.: ВЛАДОС, 1998. – 632 с.

10. Палий, И.А. Прикладная статистика: Учебное пособие для вузов

[Текст]/ И.А. Палий. – М.: Высшая школа, 2004. – 175с.

11. Рубинштейн, С.Л. Основы общей психологии [Текст]/ С.Л. Рубинштейн.

– СПб.: Питер, 2008. – 705с.

12. Симонович, С.В. Специальная информатика: Учебное пособие [Текст]/

С.В. Симонович, Г.А. Евсеев, А.Г. Алексеев. – М., 2002. – 479с.

13. Созонова, М.С. Математические методы в психологии: Учебное пособие [Текст]/ М.С. Созонова. – Тобольск: ТГСПА им. Д.И. Менделеева, 2006. – 172с.

14. Фадеев, М.А. Элементарная обработка результатов эксперимента: Учебное пособие [Текст]/ М.А. Фадеев. – СПб, М., Краснодар: Лань, 2008. – 117с.

**Приложение 1**

Таблица 1. Набранные индексы по видам реакций, их сумма.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| № | Фамилия ученика | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | Сумма |
| 1 | Бакиева | 9 | 7 | 6 | 4 | 5 | 8 | 10 | 7 | *56* |
| 2 | Гатауллин | 8 | 6 | 3 | 2 | 1 | 4 | 10 | 5 | *34* |
| 3 | Гатин | 7 | 2 | 3 | 2 | 2 | 2 | 7 | 6 | *31* |
| 4 | Долженко | 6 | 2 | 8 | 0 | 5 | 4 | 9 | 5 | *39* |
| 5 | Жарова | 9 | 4 | 7 | 3 | 4 | 7 | 9 | 5 | *48* |
| 6 | Жуйкова | 5 | 5 | 7 | 4 | 6 | 3 | 3 | 8 | *41* |
| 7 | Корикова | 9 | 4 | 3 | 4 | 5 | 7 | 7 | 8 | *47* |
| 8 | Костерина | 10 | 9 | 8 | 4 | 7 | 7 | 11 | 9 | *65* |
| 9 | Курманалиева | 10 | 7 | 7 | 4 | 4 | 7 | 13 | 8 | *60* |
| 10 | Летунов | 8 | 6 | 6 | 4 | 3 | 2 | 10 | 7 | *46* |
| 11 | Мороков | 9 | 7 | 9 | 4 | 4 | 5 | 10 | 7 | *55* |
| 12 | Перовских В. | 10 | 8 | 8 | 4 | 4 | 9 | 11 | 7 | *61* |
| 13 | Перовских М. | 8 | 2 | 5 | 2 | 4 | 7 | 9 | 8 | *45* |
| 14 | Смирнова | 5 | 5 | 9 | 5 | 4 | 8 | 10 | 7 | *53* |
| 15 | Солосина | 4 | 7 | 7 | 1 | 7 | 8 | 9 | 7 | *50* |
| 16 | Тимирова | 9 | 6 | 4 | 3 | 1 | 2 | 11 | 6 | *42* |
| 17 | Трухин | 7 | 6 | 3 | 4 | 2 | 4 | 12 | 2 | *40* |
| 18 | Филиппов | 8 | 3 | 5 | 4 | 4 | 6 | 10 | 8 | *48* |
| 19 | Хабисов | 8 | 3 | 6 | 3 | 6 | 3 | 9 | 5 | *43* |
| 20 | Цыпанов | 6 | 0 | 0 | 2 | 0 | 2 | 3 | 3 | 16 |
| Максимальный набор индексов | *10* | *9* | *11* | *5* | *8* | *10* | *13* | *9* | *75* |

1 - физическая агрессия

2 - косвенная агрессия

3 - раздражение

4 - негативизм

5 - обида

6 - подозрительность

7 - вербальная агрессия

8 - угрызение совести, чувство вины

**Приложение 2**

Таблица 2. Индексы.

|  |  |  |  |
| --- | --- | --- | --- |
| № | Фамилия ученика | Индекс агрессии | Индекс враждебности |
| 1 | Бакиева | 32 | 13 |
| 2 | Гатауллин | 27 | 5 |
| 3 | Гатин | 19 | 4 |
| 4 | Долженко | 25 | 9 |
| 5 | Жарова | 29 | 11 |
| 6 | Жуйкова | 20 | 9 |
| 7 | Корикова | 23 | 12 |
| 8 | Костерина | 38 | 14 |
| 9 | Курманалиева | 37 | 11 |
| 10 | Летунов | 30 | 5 |
| 11 | Мороков | 35 | 9 |
| 12 | Перовских В. | 37 | 13 |
| 13 | Перовских М. | 24 | 11 |
| 14 | Смирнова | 29 | 12 |
| 15 | Солосина | 27 | 15 |
| 16 | Тимирова | 30 | 3 |
| 17 | Трухин | 28 | 6 |
| 18 | Филиппов | 26 | 10 |
| 19 | Хабисов | 26 | 9 |
| 20 | Цыпанов | 9 | 2 |
| **Максимальный набор индексов** | **43** | **18** |