Министерство науки и высшего образования РФ

ФГБОУ ВО «Удмуртский государственный университет»

Институт математики, информационных технологий и физики

Кафедра математического анализа

**Курсовая работа**

**«Ранговые корреляции профессиональной направленности результатов ЕГЭ по Удмуртской республике за 2015-2020 гг.»**

Выполнил: студент гр. ОПБ-44.03.05-41м

Ульянова Анастасия Андреевна

Научный руководитель:

К.ф.-м.н., доцент Латыпова Наталья Владимировна

 Ижевск,

 2020 год

Содержание

[Введение 3](#_Toc59407023)

[ЕГЭ и его структура 5](#_Toc59407024)

[Анализ результатов ЕГЭ по профильной математике и русскому языку за 2020 год в МБОУ «СОШ № 70» г. Ижевска 8](#_Toc59407025)

[Применение закона Ципфа-Мандельброта для определения ранговой корреляции профессиональной направленности ЕГЭ- респондентов в УР за 2015-2019 гг. 11](#_Toc59407026)

[Заключение 21](#_Toc59407027)

[Список используемой литературы 22](#_Toc59407028)

[Приложение 1 24](#_Toc59407029)

[Приложение 2 25](#_Toc59407030)

# Введение

Статистика – наука, изучающая количественную сторону массовых общественных явлений с целью установления закономерностей в неразрывной связи с их качественной стороной в конкретных условиях места и времени в их взаимосвязи и взаимозависимости [1]. В каждой сфере жизнедеятельности разработаны свои формы и способы обработки статистических данных. Систематический анализ статистических данных (мониторинг) проводится с целью количественной оценки текущей ситуации и с целью прогнозирования развития ситуации в будущем для принятия основанных на фактах управленческих решениях. Так, в образовательной деятельности важнейшим аспектом является система контроля качества знаний.

В Российской системе образования существует такая система контроля знаний как единый государственный экзамен (ЕГЭ), который обеспечивает равные условия при поступлении в Вуз и сдаче выпускных экзаменов в школе.

Опыт монопольного введения ЕГЭ для обеспечения госконтроля уровня подготовки школьников показывает, что его избранная модель не всегда отвечает поставленным целям, является чрезмерно жесткой и нуждается в корректировке.

Гипотезой исследования является предположение о возможности использования законов о нормальном распределении и Ципфа-Мандельброта для обработки данных результатов ЕГЭ.

**Актуальность** работы обусловлена тем, что в обществе существует потребность в повышении качества образования, для этого необходима обработка данных по результатам ЕГЭ с использованием математической статистики.

**Цель работы** – используя законы математической статистики, обработать данные ЕГЭ по УР за 2015-2020 гг.

**Задачи:**

**-**Проверить результаты ЕГЭ по русскому языку и профильной математике на нормальное распределение с помощью MSEcxel;

-Провести ранжирование значимости предметов по числу респондентов, избравших данный предмет;

-Определить, какие предметы наиболее популярны у респондентов;

-Определить корреляцию профессиональной направленности ЕГЭ-респондентов по закону Ципфа-Мандельброта.

В качестве данных, использующихся в исследовании рассматриваются:

-Количество ЕГЭ-респондентов по Удмуртской республике за 2015-2019 гг. (Таблица 3);

-Результаты ЕГЭ по профильной математике и русскому языку в МБОУ «СОШ №70» г. Ижевска за 2020г *(см. Приложение 1)*.

# ЕГЭ и его структура

[2] Единый государственный экзамен (ЕГЭ) — это форма государственной итоговой аттестации по образовательным программам среднего общего образования (ГИА). ЕГЭ проводится по 15 учебным предметам: русский язык, математика, физика, химия, история, обществознание, информатика и информационно-коммуникационные технологии (ИКТ), биология, география, английский язык, немецкий язык, французский язык, испанский язык, китайский язык, литература.

Для получения аттестата выпускники текущего года сдают обязательные предметы — русский язык и математику. Другие предметы ЕГЭ участники сдают на добровольной основе.

Особенностью ЕГЭ является то, что результаты выступают выпускными со средней ступени образования и одновременно востребованы при поступлении в ВУЗы. После прохождения аттестации, ученику вручается сертификат ЕГЭ с количеством набранных баллов. Такая система оценки знаний введена с 2009 года. Это сделано для того, чтобы уравнять шансы на поступление всех желающих абитуриентов, чтобы будущие студенты могли свободно выбирать любые ВУЗы страны и отправлять туда свои результаты по ЕГЭ заочно — через Интернет или по почте.

Сама идея ЕГЭ первый раз была апробирована в 2000 году, в ряде регионов. В обязательном порядке ЕГЭ стал обязательным в 2009 году.

Экзаменационные задания ЕГЭ — контрольные измерительные материалы (КИМ) представляют собой комплексы заданий стандартизированной формы, выполнение которых позволяет установить уровень освоения федерального государственного образовательного стандарта.

КИМ разрабатываются Федеральным институтом педагогических измерений (ФИПИ).

При проведении ГИА в форме ЕГЭ (за исключением ЕГЭ по математике базового уровня) используется стобалльная система оценки.

По каждому предмету ЕГЭ установлено минимальное количество баллов, преодоление которого подтверждает освоение образовательной программы среднего общего образования [2].

[3] По словам министра образования Удмуртии, Рособрнадзор каждый год оценивает качество проведения итоговой аттестации и составляет рейтинг регионов. Удмуртия по итогам ЕГЭ 2020 года поднялась с 17-го на 7-е место среди 85 регионов.

*Рис.1 Средние тестовые баллы выпускников ОУ УР*

По данному рисунку можно заметить, что наибольшую сложность у выпускников УР вызывают такие предметы как профильная математика, и биология. Успешность выполнения заданий работы на экзамене обусловлена, во-первых, хорошими знаниями по предмету и, во-вторых, правильной подготовкой к данному экзамену.

По данным отчета министра образования и науки Удмуртской республики Болотниковой Светланы Михайловны за последние 3 года наблюдается рост количества стобалльников: 2018г.- 42 человека, 2019г.- 52 человека, 2020г. —- 60 человек. За 2020 год: 19 человек по русскому языку, 13 по литературе, по 7 человек по информатике и химии, 6 по истории, по 3 человека по профильной математике и географии, 2 человека по обществознанию и 1 по физике. Технические предметы вызывают у выпускников наибольшие затруднения и доля стобалльников по данным предметам мала [3].

# Анализ результатов ЕГЭ по профильной математике и русскому языку за 2020 год в МБОУ «СОШ № 70» г. Ижевска

[4] Известно, что если значения, которые принимает случайная величина, зависят от большого числа различных факторов, каждый из которых, взятый в отдельности, влияет на ту величину сравнительно мало, то можно приближённо считать, что рассматриваемая случайная величина подчиняется нормальному закону распределения.

Будем считать, что тестовый балл, получаемый участником ЕГЭ, является случайной величиной Х, зависящей от множества факторов: популярность предмета, уровень подготовленности к предмету, сложность теста, продолжительность подготовки к тестированию, психологическое состояние участника во время процедуры сдачи ЕГЭ и т.п., т.е. зависит от влияния большого числа разнородных факторов. Следовательно, можно предположить, что наша случайная величина Х должна приближённо подчиняться нормальному (Гаусса) закону [4].

Используя данные *Приложения 1*, с помощью встроенных пакетов Microsoft Excel построим полигоны эмпирических частот и график вероятности для нормального распределения по русскому языку:

*Рис.2 Полигон эмпирических частот и график вероятности для нормального распределения по русскому языку*

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| X | 5 | 15 | 25 | 35 | 45 | 55 | 65 | 75 | 85 | 95 |
| Вероятность,  | 0,0001 | 0,0009 | 0,0061 | 0,0273 | 0,0846 | 0,1722 | 0,2447 | 0,2283 | 0,0544 | 0,0636 |

*Таблица 1 Вероятности для нормального распределения по русскому языку*

Данные статистического анализа результатов ЕГЭ показывают, что выборочная плотность распределения измеряемой случайной величины (количество ЕГЭ-респондентов данного ранга) получена из генеральной совокупности с нормальным законом распределения.

По русскому языку наблюдается смещение максимума (красная кривая) в сторону более высоких показателей. Это означает, что данная система оценки качества (ЕГЭ) по русскому языку адаптирована в МБОУ «СОШ № 70».

Рассмотрим статистику ЕГЭ математике. Аналогично построим полигоны эмпирических частот и вероятность для нормального распределения по математике:

*Рис.3 Полигон эмпирических частот и график вероятности для нормального распределения по математике*

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| X | 5 | 15 | 25 | 35 | 45 | 55 | 65 | 75 | 85 | 95 |
| Вероятность,  | 0,0169 | 0,0537 | 0,1241 | 0,2047 | 0,2316 | 0,1921 | 0,1105 | 0,0444 | 0,0131 | 0,0027 |

*Таблица 2 Вероятности для нормального распределения по математике*

Видно, что выборочная плотность распределения измеряемой случайной величины (количество ЕГЭ-респондентов данного ранга) получена из генеральной совокупности с нормальным законом распределения.

Система оценки качества (ЕГЭ) по математике адаптирована в МБОУ «СОШ № 70», но смещения в сторону более высоких баллов не наблюдается.

# Применение закона Ципфа-Мандельброта для определения ранговой корреляции профессиональной направленности ЕГЭ- респондентов в УР за 2015-2019 гг.

 *«Математики верят в закон Ципфа потому,*

 *что лингвисты постановили считать его лингвистическим законом,*

*а лингвисты со своей стороны верят в него потому,*

*что математики постановили считать его*

*математическим законом» (Густав Хердан, 1962).*

Корреляционный анализ — метод обработки статистических данных, заключающийся в изучении коэффициентов (корреляции) между переменными. При этом сравниваются коэффициенты корреляции между одной парой или множеством пар признаков для установления между ними статистических взаимосвязей.

Задачи описываемого метода: получить информацию об одной из искомых переменных с помощью другой; определить тесноту связи между исследуемыми переменными.

Корреляционный анализ предполагает определение зависимости между изучаемыми признаками, в связи с чем задачи корреляционного анализа можно дополнить следующими [5]:

* выявление факторов, оказывающих наибольшее влияние на результативный признак;
* выявление неизученных ранее причин связей;
* построение корреляционной модели с ее параметрическим анализом;
* исследование значимости параметров связи и их интервальная оценка.

Частотно-ранговые распределения широко используются при количественном описании объектов в информатике, лингвистике, биологии и многих других областях. Рассмотрим закон Ципфа-Мандельброта

Закон Ципфа-Мандельброта — эмпирическая закономерность распределения частоты слов естественного языка: если все слова языка (или просто достаточно длинного текста) упорядочить по убыванию частоты их использования, то частота n-го слова в таком списке окажется приблизительно обратно пропорциональной его порядковому номеру n (так называемому рангу этого слова). Например, второе по используемости слово встречается примерно в два раза реже, чем первое, третье — в три раза реже, чем первое, и так далее.

Допустим, имеется текст длиной  слов и словарь объемом  слов с указанием частоты появления слова в тексте. Слова в словаре расположены в порядке убывания их по частоте и проранжированы от 1 до . Ранг, равный 1, присваивается слову, частота появления которого наибольшая; ранг, равный , – наименее употребимому слову. Тогда:

где  – относительная частота появления слова в тексте;  – абсолютная частота появления слова   ранга в тексте определенной длины;  – число слов в тексте;  – ранг слова, где .

Если умножить вероятность или относительную частоту обнаружения слова в тексте на ранг  слова, то получим:

,

где   – константа; .

Если преобразовать формулу, то получим: , т.е. функцию типа , график которой – равносторонняя гипербола.

Таким образом, на основе анализа полученных зависимостей Ципф предложил эмпирическую формулу, устанавливающую связь между частотой появления слов в тексте и его рангом в словаре:

где  –  эмпирически определяемая константа, изменяющаяся для разных текстов.

Ранговое распределение имеет ступенчатый характер (ряд слов появляется с одинаковой частотой), но при аппроксимации (замены одних математических объектов другими, в том или ином смысле близкими к исходным) можно считать распределение Ципфа гиперболой.



*Рис. 3 Закон Ципфа-Мандельброта при анализе частот слов*

Законы Ципфа универсальны. В принципе, они применимы не только к текстам. В аналогичную форму выливается, например, зависимость количества городов от числа проживающих в них жителей. Характеристики популярности узлов в сети Интернет – тоже отвечают законам Ципфа.

Законы Ципфа проявляются и при исследовании документальных информационных потоков (ДИП). Степенные статистики описывают распространение вирусов в Интернете и ВИЧ-инфицированного населения.

Ципф обнаружил, что это же правило действует в распределении доходов людей в стране: самый богатый человек имеет вдвое больше денег, чем следующий богач и так далее [6].

Позже стало понятно, что этот закон также работает в отношении размера городов. Город с самым большим населением в любой стране в два раза больше, чем следующий по размеру город и так далее. Невероятно, но закон Ципфа действовал абсолютно во всех странах мира на протяжении прошлого столетия.

Просто взгляните на список самых больших городов Соединённых Штатов. Итак, в соответствии с переписью 2010-го года население самого большого города США, Нью-Йорка, составляет 8 175 133 человека. Номер два – Лос-Анджелес с населением в 3 792 621 человек. Следующие три города, Чикаго, Хьюстон и Филадельфия, могут похвастаться населением в 2 695 598, 2 100 263 и 1 526 006 человек соответственно. Очевидно, эти числа неточны, но, тем не менее, они удивительно соответствуют закону Ципфа [7].

Проиллюстрируем проявление закона Ципфа-Мандельброта в учебном процессе.

Рассмотрим данные о профессиональной направленности ЕГЭ-респондентов, полученные по результатам проведения ЕГЭ в Удмуртский республике за 2015-2019 гг., посредством ранжирования значимости предметов по числу респондентов, избравших данный профильный предмет

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Количество респондентов | Предмет | Количество респондентов | Предмет | Количество респондентов | Предмет | Количество респондентов | Предмет | Количество респондентов | Предмет |
|  | 2015 | 2016 | 2017 | 2018 | 2019 |
| 1 | 4961 (33,7 %) | Проф. математика | 4546(31,4%) | Проф.математика | 4633(30%) | Проф. математика | 4805(29,7%) | Проф. Математика | 4220(27,6%) | Проф. Математика |
| 2 | 3197 (21,7%) | Обществознание | 3017(20,9%) | Обществознание | 3114(20,2%) | Обществознание | 3120(19,3%) | Обществознание | 2860(18,7%) | Обществознание |
| 3 | 1855(12,6%) | Физика | 1959(13,5%) | Физика | 1992(12,9%) | Физика | 1978(12,2%) | Физика | 1745(11,4%) | Физика |
| 4 | 1200(8,2%) | Биология | 1287(8,9%) | Биология | 1431(9,3%) | Биология | 1562(9,7%) | Биология | 1565(10,2%) | Биология |
| 5 | 1084(7,4%) | История | 1105(7,6%) | История | 1262(8,2%) | История | 1177(7,3%) | История | 1106(7,2%) | История |
| 6 | 697(4,7%) | Химия | 702(4,9%) | Химия | 769(5%) | Химия | 931(5,7%) | ИКТ | 1066(7%) | ИКТ |
| 7 | 519(3,51%) | Литература | 574(4%) | Литература | 713(4,6%) | ИКТ | 909(5,6%) | Химия | 950(6,2%) | Химия |
| 8 | 512(3,49%) | ИКТ | 573(3,9%) | ИКТ | 682(4,4%) | Литература | 656(4,1%) | Литература | 713(4,7%) | Литература |
| 9 | 472(3,2%) | Англ. язык | 489(3,4%) | Англ. язык | 553(3,6%) | Англ. язык | 631(4%) | Англ. Язык | 676(4,5%) | Англ.язык |
| 10 | 209(1,4%) | География | 207(1,4%) | География | 268(1,7%) | География | 358(2,2%) | География | 354(2,3%) | География |
| 11 | 11 (0,1%) | Немецкий язык | 13(0,1%) | Немецкий язык | 17(0,1%) | Немецкий язык | 27(0,2%) | Немецкий язык | 26(0,2%) | Немецкий язык |
| Всего | 14717 (100%) |  | 14472(100%) |  | 15434(100%) |  | 16154(100%) |  | 15281(100%) |  |

*Таблица 3 Проранжированные данные ЕГЭ*

Анализируя эти данные можно сказать, что ведущими предметами за эти года являются профильная математика, обществознание, также в лидерах находятся физика и биология. Профильная математика занимает первенство, потому что практически на все специальности Ижевского государственного технического университета имени М.Т. Калашникова, Ижевской государственной сельскохозяйственной академии, многие специальности Удмуртского государственного университета обязательным является результат ЕГЭ по ней.

Высокий ранг обществознания при этом объясняется условиями приема в вузы, где данные ЕГЭ по обществознанию востребованы чаще остальных профильных предметов, что далеко не всегда оправдано. По поводу физики можно сказать, что респонденты выбирают ее чаще всего, потому что она является ведущим предметом для приема в ИжГТУ имени М.Т. Калашникова, на большинство направлений подготовки ИжГСХА, а также в институт математики, информационных технологий и физики в УдГУ. Биология является обязательным предметом при приеме в Ижевская государственная медицинскую академию, некоторые направления подготовки ИжГСХА, а также институт естественных наук УдГУ. Также популярным предметом является история, которая необходима для поступления на многие гуманитарные направления подготовки Российской правовой академии и УдГУ. Популярность такого предмета как химия объясняется тем, что это основной предмет для поступления в ИГМА и ИЕН УдГУ. Можно заметить, что все большую популярность набирает информатика и ИКТ, потому что в настоящее время востребованы IT-специалисты, и ее результаты учитываются на некоторые направления подготовки.

Построим график на основе *Таблицы 3:*

*Рис.4 Данные о профессиональной направленности ЕГЭ-респондентов в 2015-2019 гг.*

Если рассмотреть график, то видно, что он принимает вид гиперболы, это значит, что имеют место ранговые корреляции с количеством ЕГЭ-респондентов по профильным предметам, которые аппроксимируются степенной функцией вида:

 (1),

где – количество ЕГЭ-респондентов, избравших предмет i-го ранга; А, α – положительные константы, определяемые по экспериментальным данным следующим образом. Степенная функция (1) в логарифмических координатах представляет прямую:

 (2),

которая проводится методом наименьших квадратов по ранжированным данным, откуда определяются постоянные А, α *(см. Приложение 1)*

|  |  |  |  |
| --- | --- | --- | --- |
|  | Значение А | Значение  | Уравнение кривой Ципфа-Мандельброта |
| 2015 г. | А= |  |  |
| 2016 г. | А=9,190804 | 1,68946 |  |
| 2017 г. | А=9,158483 | 1,58699 |  |
| 2018 г. | А=9,081712 | 1,46284 |  |
| 2019 г. | А=8,947106 | 1,38508 |  |

*Таблица 4 Значения А, и уравнения кривой Ципфа-Мандельброта*

Построим соответствующие кривые Ципфа-Мандельброта:

|  |  |
| --- | --- |
| 2015 г. |  |
| 2016 г. |  |
| 2017 г. |  |
| 2018 г. |  |
| 2019 г. |  |

*Таблица 5 Кривые Ципфа-Мандельброта*

Анализ данных табл. 5 и рис. 4 говорит о том, что при проведении ЕГЭ УР 2015-2019 гг. наблюдались ранговые корреляции профессиональной направленности ЕГЭ-респондентов, аппроксимируемые законом Ципфа-Мальдеброта (1). Легко заметить, что коэффициенты А и за данный период изменились слабо, и «лидирующая» группа предметов *профильная математика- обществознание- физика- биология- история* не изменилась.

В качестве примера была взята статья В. Е. Фирстова, Р. А. Иванова «Из истории стенографии и интерпретации закона Ципфа−Мандельброта в учебном процессе» [8].

# Заключение

В ходе работе поставленные задачи были выполнены:

**-**С помощью Microsoft Excel были проверены результаты ЕГЭ по профильной математике и русскому языку в МБОУ «СОШ №70» на нормальное распределение. Выборочная плотность распределения измеряемых случайных величин (количество ЕГЭ-респондентов данного ранга) получена из генеральной совокупности с нормальным законом распределения. По русскому языку наблюдается смещение в сторону больших баллов, по математике же максимум находится точно по середине. Из этого можно сделать заключение, что в МБОУ «СОШ № 70» ЕГЭ по русскому языку адаптирован более успешно.

-Проведено ранжирование значимости предметов по числу респондентов, избравших данный предмет.

-Определено, какие предметы наиболее популярны у респондентов. На протяжении 5 лет «лидирующая» группа предметов профильная математика- обществознание- физика- биология- история не изменилась.

-Определена корреляция профессиональной направленности ЕГЭ-респондентов по закону Ципфа-Мандельброта.

Поставленная цель достигнута. С помощью законов математической статистики обработаны данные ЕГЭ по УР за 2015-2020 гг.

В дальнейшем планируется, используя законы математический статистики, изучить корреляцию результатов ОГЭ (9 класс) и ЕГЭ (11 класс) по данным МБОУ «СОШ №70» г. Ижевска и на основе полученных данных сделать выводы и составить рекомендации по подготовке к экзаменам.

Список используемой литературы

1. В. В. Полякова, Н. В. Шаброва; М-во образования и науки Рос. Федерации, Урал. федер. ун-т. – 2-е изд., испр. и доп. – Екатеринбург : Изд-во Урал. ун-та, 2015. – 148 с.
2. Основные сведения о ЕГЭ

[Электронный ресурс]// URL: <http://ege.edu.ru/ru/main/main_item/>

(Дата обращения: 14.12.2020)

1. Совещание по итогам ГИА в 2020 году [Электронный ресурс]// URL:

<https://ege.ciur.ru/content/247732/>

(Дата обращения: 14.12.2020)

1. М.Б. Булакина, Е.Е. Якивчук, Анализ результатов ЕГЭ 2009 на основе закона нормального распределения // Труды XVII Всероссийской научно-методической конференции, Санкт-Петербург 21–24 июня 2010 года – Санкт-Петербург: Государственный научно-исследовательский институт информационных технологий и телекоммуникаций «Информика» – с. 80;
2. Метод корреляционного анализа: пример [Электронный ресурс]// URL:

<https://fb-ru.turbopages.org/fb.ru/s/article/341341/metod-korrelyatsionnogo-analiza-primer-korrelyatsionnyiy-analiz---eto>(Дата обращения: 14.12.2020)

1. Закономерности информетрии и их применение для исследования информационных потоков [Электронный ресурс]// URL:

<https://studme.org/34366/informatika/zakonomernosti_informetrii_primenenie_dlya_issledovaniya_informatsionnyh_potokov>

(Дата обращения: 14.12.2020)

1. Закон Ципфа в городе [Электронный ресурс]// URL:

<https://masterok-livejournal-com.turbopages.org/masterok.livejournal.com/s/2824971.html>

(Дата обращения: 14.12.2020)

1. В. Е. Фирстов, Р. А. Иванов, Из истории стенографии и интерпретации закона Ципфа−Мандельброта в учебном процессе // Ярославский педагогический вестник – 2013 – № 4 – Том III (Естественные науки) – c. 14-20.

# Приложение 1

Результаты ЕГЭ по русскому языку в МБОУ «СОШ № 70» г. Ижевска за 2020 г.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 |
| Баллы | 82 | 82 | 89 | 71 | 71 | 72 | 43 | 82 | 59 | 89 | 73 | 64 | 49 | 82 | 34 | 45 | 66 | 78 | 62 | 80 | 64 | 62 |

Результаты ЕГЭ по математике в МБОУ «СОШ № 70» г. Ижевска за 2020 г.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 |
| Баллы | 50 | 62 | 45 | 50 | 50 | 74 | 62 | 39 | 50 | 78 | 50 | 39 | 18 | 45 | 39 | 33 | 45 | 76 | 45 | 39 | 39 | 18 |

Представленные выше данные разбиваются на интервалы с шагом 10 баллов и вычисляется частота попадания в каждый интервал.

|  |  |
| --- | --- |
| Математика | Русский язык |
| Группы | Середина интервала | Частота попадания | Группы | Середина интервала | Частота попадания |
| 0-10 | 5 | 0 | 0-10 | 5 | 0 |
| 10-20 | 15 | 2 | 10-20 | 15 | 0 |
| 20-30 | 25 | 0 | 20-30 | 25 | 0 |
| 30-40 | 35 | 6 | 30-40 | 35 | 1 |
| 40-50 | 45 | 9 | 40-50 | 45 | 3 |
| 50-60 | 55 | 0 | 50-60 | 55 | 1 |
| 60-70 | 65 | 2 | 60-70 | 65 | 5 |
| 70-80 | 75 | 3 | 70-80 | 75 | 6 |
| 80-90 | 85 | 0 | 80-90 | 85 | 6 |
| 90-100 | 95 | 0 | 90-100 | 95 | 0 |
|  |  | 22 |  |  | 22 |

Далее с помощью пакетов Microsoft Excel строятся полигоны эмпирических частот и графики вероятностей для нормального распределения *(Рис.3, Рис.4)*

# Приложение 2

**Нахождение констант A и для 2015 года и построение кривой Ципфа:**

По данным *Таблицы 3* заполняем следующую таблицу (𝑝(𝑖) – количество ЕГЭ-респондентов, избравших предмет 𝑖 -го ранга):

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| № |   | ln (p(i)) | ln(i)\*ln(p(i)) |  |  |  A | ln A |
| 1 | 0 | 8,509363 | 0 | 0 | -1,76002 | 9,265638 | 2,226313 |
| 2 | 0,693147 | 8,069968 | 5,593676 | 0,480453 |  |  |  |
| 3 | 1,098612 | 7,52564 | 8,267761 | 1,206949 |  |  |  |
| 4 | 1,386294 | 7,090077 | 9,828934 | 1,921812 |  |  |  |
| 5 | 1,609438 | 6,988413 | 11,24742 | 2,59029 |  |  |  |
| 6 | 1,791759 | 6,546785 | 11,73026 | 3,210402 |  |  |  |
| 7 | 1,94591 | 6,251904 | 12,16564 | 3,786566 |  |  |  |
| 8 | 2,079442 | 6,238325 | 12,97223 | 4,324077 |  |  |  |
| 9 | 2,197225 | 6,156979 | 13,52827 | 4,827796 |  |  |  |
| 10 | 2,302585 | 5,342334 | 12,30118 | 5,301898 |  |  |  |
| 11 | 2,397895 | 2,397895 | 5,749902 | 5,749902 |  |  |  |
| всего | 17,50231 | 71,11768 | 103,3853 | 33,40015 |  |  |  |

Находим необходимые значения (α, А) через встроенную в Excel функцию построения линейной регрессии. На плоскости откладываем точки с координатами: по оси абсцисс lni, а по оси ординат lnp(i), далее используем встроенную функцию построения прямой линейной регрессии методом наименьших квадратов, выводим уравнение этой прямой (на рисунке это пунктирная прямая). Тогда угловой коэффициент прямой равен – α, а свободный член ln A.

, откуда получаем, что А=,

Далее полученные данные подставляются в уравнение кривой Ципфа:

Строим кривую Ципфа:

|  |  |
| --- | --- |
| 1 | 9,265638 |
| 2 | 2,735633 |
| 3 | 1,340092 |
| 4 | 0,807682 |
| 5 | 0,545353 |
| 6 | 0,395655 |
| 7 | 0,301641 |
| 8 | 0,238464 |
| 9 | 0,193818 |
| 10 | 0,161013 |
| 11 | 0,136147 |

**Нахождение констант A и для 2016 года и построение кривой Ципфа:**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| № |   | ln (p(i)) | ln(i)\*ln(p(i)) |  |  |  A | ln A |
| 1 | 0 | 8,422003 | 0 | 0 | -1,68946 | 9,190804 | 2,218203 |
| 2 | 0,693147 | 8,012018 | 5,553508 | 0,480453 |  |  |  |
| 3 | 1,098612 | 7,580189 | 8,327689 | 1,206949 |  |  |  |
| 4 | 1,386294 | 7,160069 | 9,925964 | 1,921812 |  |  |  |
| 5 | 1,609438 | 7,007601 | 11,2783 | 2,59029 |  |  |  |
| 6 | 1,791759 | 6,553933 | 11,74307 | 3,210402 |  |  |  |
| 7 | 1,94591 | 6,352629 | 12,36165 | 3,786566 |  |  |  |
| 8 | 2,079442 | 6,350886 | 13,2063 | 4,324077 |  |  |  |
| 9 | 2,197225 | 6,192362 | 13,60601 | 4,827796 |  |  |  |
| 10 | 2,302585 | 5,332719 | 12,27904 | 5,301898 |  |  |  |
| 11 | 2,397895 | 2,564949 | 6,15048 | 5,749902 |  |  |  |
| Всего: | 17,50231 | 71,52936 | 104,432 | 33,40015 |  |  |  |

Строим прямую методом наименьших квадратов:

Получаем А=9,190804, 1,68946

Уравнение кривой:

Строим кривую Ципфа:

|  |  |
| --- | --- |
| 1 | 9,190804 |
| 2 | 2,84954 |
| 3 | 1,436399 |
| 4 | 0,883478 |
| 5 | 0,605996 |
|  6 | 0,445345 |
| 7 | 0,343236 |
| 8 | 0,273916 |
| 9 | 0,22449 |
| 10 | 0,187885 |
| 11 | 0,159941 |

**Нахождение констант A и α для 2017 года и построение кривой Ципфа:**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| № |   | ln (p(i)) | ln(i)\*ln(p(i)) |  |  |  A | ln A |
| 1 | 0 | 8,44096 | 0 | 0 | -1,58699 | 9,158483 | 2,214681 |
| 2 | 0,693147 | 8,043663 | 5,575443 | 0,480453 |  |  |  |
| 3 | 1,098612 | 7,596894 | 8,346042 | 1,206949 |  |  |  |
| 4 | 1,386294 | 7,266129 | 10,07299 | 1,921812 |  |  |  |
| 5 | 1,609438 | 7,140453 | 11,49212 | 2,59029 |  |  |  |
| 6 | 1,791759 | 6,645091 | 11,9064 | 3,210402 |  |  |  |
| 7 | 1,94591 | 6,569481 | 12,78362 | 3,786566 |  |  |  |
| 8 | 2,079442 | 6,52503 | 13,56842 | 4,324077 |  |  |  |
| 9 | 2,197225 | 6,315358 | 13,87626 | 4,827796 |  |  |  |
| 10 | 2,302585 | 5,590987 | 12,87372 | 5,301898 |  |  |  |
| 11 | 2,397895 | 2,833213 | 6,793749 | 5,749902 |  |  |  |
| Всего  | 17,50231 | 72,96726 | 107,2888 | 33,40015 |  |  |  |

Строим прямую методом наименьших квадратов:

Получаем А=9,158483,

Уравнение кривой:

Строим кривую Ципфа:

|  |  |
| --- | --- |
| 1 | 9,158483 |
| 2 | 3,048533 |
| 3 | 1,601899 |
| 4 | 1,014748 |
| 5 | 0,712136 |
| 6 | 0,533215 |
| 7 | 0,417502 |
| 8 | 0,337774 |
| 9 | 0,280186 |
| 10 | 0,237045 |
| 11 | 0,20377 |

**Нахождение констант A и α для 2018 года и построение кривой Ципфа:**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| № |   | ln (p(i)) | ln(i)\*ln(p(i)) |  |  |  A | ln A |
| 1 | 0 | 8,477412 | 0 | 0 | -1,46284 | 9,081712 | 2,206263 |
| 2 | 0,693147 | 8,045588 | 5,576777 | 0,480453 |  |  |  |
| 3 | 1,098612 | 7,589842 | 8,338293 | 1,206949 |  |  |  |
| 4 | 1,386294 | 7,353722 | 10,19442 | 1,921812 |  |  |  |
| 5 | 1,609438 | 7,070724 | 11,37989 | 2,59029 |  |  |  |
| 6 | 1,791759 | 6,836259 | 12,24893 | 3,210402 |  |  |  |
| 7 | 1,94591 | 6,812345 | 13,25621 | 3,786566 |  |  |  |
| 8 | 2,079442 | 6,486161 | 13,48759 | 4,324077 |  |  |  |
| 9 | 2,197225 | 6,447306 | 14,16618 | 4,827796 |  |  |  |
| 10 | 2,302585 | 5,880533 | 13,54043 | 5,301898 |  |  |  |
| 11 | 2,397895 | 3,295837 | 7,903072 | 5,749902 |  |  |  |
| Всего  | 17,50231 | 74,29573 | 110,0918 | 33,40015 |  |  |  |

Строим прямую методом наименьших квадратов:

Получаем А=9,081712,

Уравнение кривой:

Строим кривую Ципфа:

|  |  |
| --- | --- |
| 1 | 9,081712 |
| 2 | 3,294644 |
| 3 | 1,820602 |
| 4 | 1,195224 |
| 5 | 0,862354 |
| 6 | 0,660474 |
| 7 | 0,527137 |
| 8 | 0,433601 |
| 9 | 0,364974 |
| 10 | 0,312843 |
| 11 | 0,272129 |

**Нахождение констант A и α для 2019 года и построение кривой Ципфа:**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| № |   | ln (p(i)) | ln(i)\*ln(p(i)) |  |  |  A | ln A |
| 1 | 0 | 8,34759 | 0 | 0 | -1,38508 | 8,947106 | 2,19133 |
| 2 | 0,693147 | 7,958577 | 5,516465 | 0,480453 |  |  |  |
| 3 | 1,098612 | 7,46451 | 8,200602 | 1,206949 |  |  |  |
| 4 | 1,386294 | 7,355641 | 10,19708 | 1,921812 |  |  |  |
| 5 | 1,609438 | 7,008505 | 11,27975 | 2,59029 |  |  |  |
| 6 | 1,791759 | 6,971669 | 12,49155 | 3,210402 |  |  |  |
| 7 | 1,94591 | 6,856462 | 13,34206 | 3,786566 |  |  |  |
| 8 | 2,079442 | 6,569481 | 13,66085 | 4,324077 |  |  |  |
| 9 | 2,197225 | 6,516193 | 14,31754 | 4,827796 |  |  |  |
| 10 | 2,302585 | 5,869297 | 13,51456 | 5,301898 |  |  |  |
| 11 | 2,397895 | 3,258097 | 7,812574 | 5,749902 |  |  |  |
| Всего | 17,50231 | 74,17602 | 110,333 | 33,40015 |  |  |  |

Строим прямую методом наименьших квадратов:

Получаем А=8,947106, α=1,38508

Уравнение кривой:

Строим кривую Ципфа:

|  |  |
| --- | --- |
| 1 | 8,947106 |
| 2 | 3,425557 |
| 3 | 1,953575 |
| 4 | 1,311534 |
| 5 | 0,962834 |
| 6 | 0,747961 |
| 7 | 0,60416 |
| 8 | 0,502144 |
| 9 | 0,426558 |
| 10 | 0,368638 |
| 11 | 0,323048 |