МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение высшего образования

«Тульский государственный университет»

Институт прикладной математики и компьютерных наук

Кафедра вычислительной техники

Отчет по лабораторной работе № 2

«ОСНОВНЫЕ РАСПРЕДЕЛЕНИЯ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ»

по дисциплине «Математическая статистика»

Выполнил студент группы 220681:

Илюхин Д.М.

Проверил:

Асс. Демидова А. В.

Тула 2020

**1. Цель работы**

Исследовать основные распределения, используемые в математической статистике: нормальное распределение, распределение хи-квадрат, распределения Стьюдента и Фишера.

**2. Задание на лабораторную работу**

1. Предприятие изготавливает трубы, средний внешний диаметр которых равен 30,20 мм, а стандартное отклонение равно 0,33 мм, Согласно ТУ трубы признаются годными, если диаметр находится в пределах 29,70+/-0,41 мм. Какая доля изготовленных труб соответствует ТУ? Обработать данные, построить график плотности нормального распределения. Сделать выводы.

2. Директор школы хочет узнать, действительно ли то, что учителя более предвзято относятся к мальчикам, чем к девочкам, т.е. более склонны хвалить девочек. Для этого им были проанализированы характеристики учеников, написанные учителями, на предмет частоты встречаемости трех слов: «пассивный», «активный», «старательный», «дисциплинированный», синонимы слов так же подсчитывались. Данные о частоте встречаемости слов были занесены в таблицу:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Пассивный | Активный | Старательный | Дисциплинированный |
| Мальчики | 7 | 6 | 6 | 9 |
| Девочки | 6 | 8 | 7 | 8 |

Обработать данные с использованием критерия хи-квадрат. Сделать выводы.

3. Необходимо сравнить между собой результаты выполнения логических задач до и после курса обучения дисциплины «Математическая статистика». Чтобы узнать различаются ли результаты до курса обучения и после необходимо вычислить критерий Стьюдента, построить график и сделать выводы по этой задаче

|  |  |  |
| --- | --- | --- |
| № | Результаты выполнения логических задач до курса (мин.) | Результаты выполнения логических задач после курса (мин.) |
| 1 | 10 | 9 |
| 2 | 7 | 7 |
| 3 | 8 | 7 |
| 4 | 10 | 9 |
| 5 | 11 | 11 |
| 6 | 9 | 8 |
| 7 | 9 | 8 |
| 8 | 10 | 7 |
| 9 | 7 | 6 |

4. В двух третьих классах проводилось тестирование умственного развития по тесту ТУРМШ десяти учащихся. Полученные значения величин средних достоверно не различались, однако психолога интересует вопрос – есть ли различия в степени однородности показателей умственного развития между классами. Обработать данные с использованием критерия Фишера. Сделать выводы.

|  |  |  |
| --- | --- | --- |
| № учащегося | Первый класс А | Второй класс Б |
| 1 | 90 | 35 |
| 2 | 29 | 30 |
| 3 | 39 | 54 |
| 4 | 79 | 76 |
| 5 | 88 | 78 |
| 6 | 53 | 54 |
| 7 | 34 | 41 |
| 8 | 40 | 87 |
| 9 | 75 | 41 |
| 10 | 79 | 93 |

**3. Ход работы**

1. Чтобы найти долю правильных труб, нужно сначала определить верхнюю и нижнюю границу нормы. После необходимо найти значения функции НОРМ.РАСП от двух границ, среднего диаметра, стандартного отклонения и ИСТИНА (интегральная функция распределения). Результатом будет вероятность того, что случайное значение окажется меньше границы. Находим разность вероятностей двух границ и получаем процент правильных труб (Рисунок 1).

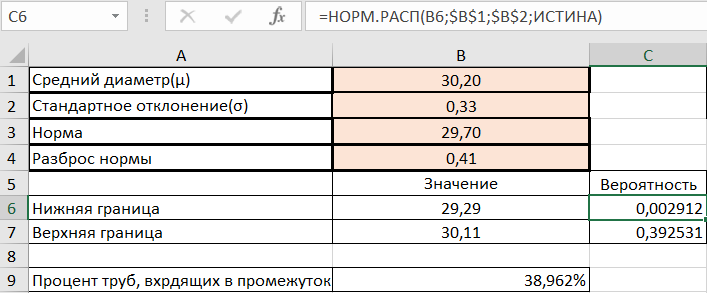


Рисунок 1. Нахождения процента правильных труб.

Теперь построим график области нормального распределения. Для этого создадим множество значений x и найдём множество результатов функции НОРМ.РАСП. Только теперь последней переменной будет являться ЛОЖЬ. Построим по результатам функции график, который и будет являться графиком области нормального распределения (Рисунок 2).



Рисунок 2. График области нормального распределения.

В результате нормальными трубами является всего 39 процентов экземпляров. Также практически все труб располагаются в промежутке [29,2; 31,2] и большая часть труб имеет значение близкое к 30,2.

2. Для использования хи-квадрата сначала нужно подсчитать сумму значений каждого столбца, каждой строки и сумму всех значений. Теперь нужно подсчитать ожидаемые частоты распределения. Ожидаемое значение равно сумме столбца\*сумма строки/сумма всех значений. Далее нужно рассчитать вероятность независимости с помощью функции ХИ2.ТЕСТ, которая принимает все исходные и ожидаемые значения (Рисунок 3).



Рисунок 3. Обработка данных с использованием критерия хи-квадрат.

В результате получается значение 0,923. Если значение стремиться к 0, то между полом и оценкой наблюдается полная зависимость. Если к 1, то между полом и оценкой наблюдается полная независимость. В результате зависимость между полом и оценкой практически отсутствует.

3. Для определения различия нужно использовать функцию СТЬЮДЕНТ.ТЕСТ, которая принимает два массива данных, значение хвоста (равно 1) и тип (равно 3). Если полученное значение стремиться к 0, то значения различаются. Если стремится к 0.5, то значения не различаются (Рисунок 4).

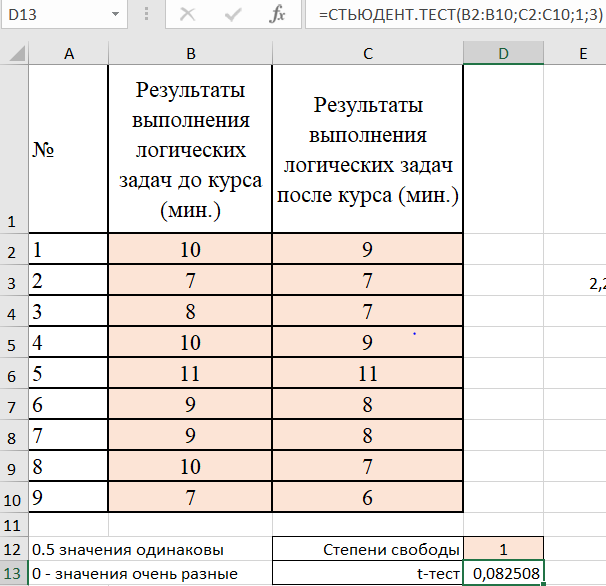


Рисунок 4. Результат t-теста.

Теперь построим график, для чего нужно задать значения x и найти значение функции СТЬЮДЕНТ.РАСП от х, степени свободы (равно 1) и ЛОЖЬ (функция плотности вероятности). После на основе найденных значений нужно построить график (Рисунок 5).

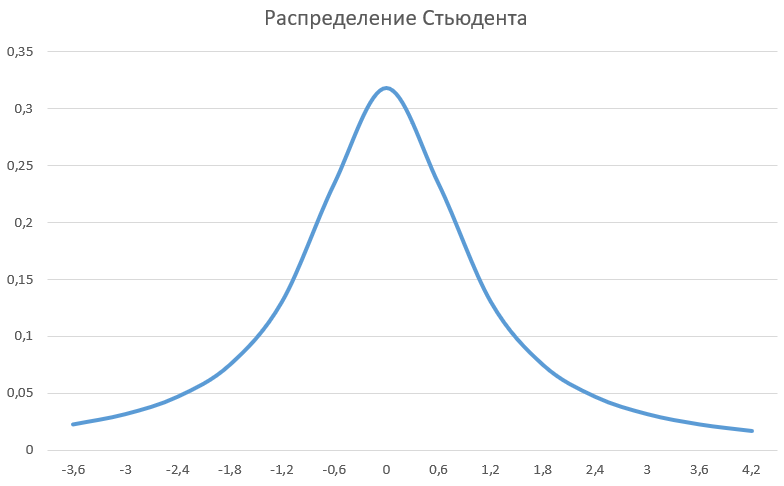


Рисунок 5. Распределение Стьюдента.

В результате t-теста было получено значение 0.082, которое очень близко к нулю. Значит результаты теста до и после обучения различаются.

4. Для обработки данных с помощью критерия Фишера нужно использовать функцию F.ТЕСТ, которая принимает два массива значений (Рисунок 6). Если получено значение 1, то разница между дисперсиями отсутствует. Если 0, то дисперсии сильно отличаются друг от друга.

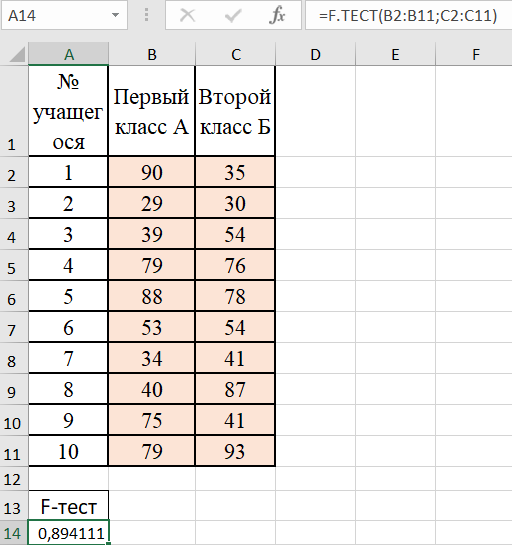


Рисунок 6. Результат f-теста.

В результате f-теста было получено значение 0.89, из чего делаем вывод, что дисперсии близки друг к другу и показатели умственных способностей однородны.

1. **Ответы на контрольные вопросы**
   * + 1. Что такое плотность нормального распределения и в каких случаях она используется?

Нормальное распределение — распределение вероятностей, которое в одномерном случае задаётся функцией плотности вероятности, совпадающей с функцией Гаусса.

Нормальное распределение играет важнейшую роль во многих областях знаний, особенно в статистической физике. Физическая величина, подверженная влиянию значительного числа независимых факторов, способных вносить с равной погрешностью положительные и отрицательные отклонения, вне зависимости от природы этих случайных факторов, часто подчиняется нормальному распределению, поэтому из всех распределений в природе чаще всего встречается нормальное.

* + - 1. Критерий хи-квадрат: дать определение и рассказать в каких случаях он используется.

Распределение хи-квадрат — это распределение суммы квадратов k независимых стандартных нормальных случайных величин. Используется при оценивании дисперсии, при проверке гипотез согласия, однородности, и во многих других задачах статистического анализа данных.

* + - 1. Критерий Стьюдента: дать определение и рассказать в каких случаях он используется.

Распределение Стьюдента — это однопараметрическое семейство абсолютно непрерывных распределений. Играет важную роль в статистическом анализе и используется, например, для оценки статистической значимости разности двух выборочных средних, при построении доверительного интервала для математического ожидания нормальной совокупности при неизвестной дисперсии, а также в линейном регрессионном анализе.

* + - 1. Критерий Фишера: дать определение и рассказать в каких случаях он используется.

Критерий Фишера — статистический критерий, тестовая статистика которого при выполнении нулевой гипотезы имеет распределение Фишера (F-распределение). Используется для проверки равенства дисперсий, проверки ограничений на параметры регрессии, проверки значимости линейной регрессии и для проверки гетероскедастичности.

**5. Выводы по проделанной работе**

В ходе данной лабораторной работы были изучены основные распределения, используемые в математической статистике: нормальное распределение, распределение хи-квадрат, распределения Стьюдента и Фишера.