

Министерство цифрового развития, связи и
массовых коммуникаций Российской Федерации

Федеральное государственное бюджетное образовательное учреждение высшего
образования «Сибирский государственный университет телекоммуникаций и
информатики» (СибГУТИ)

Кафедра вычислительных систем

ОТЧЕТ
по лабораторной работе №1
по дисциплине «**Прикладные задачи Теории Вероятностей**»

Выполнил:

студент гр. ИС-142

«__» сентября 2023 г.

/Григорьев Ю.В./

Проверил:

профессор кафедры В.С., д.т.н.

«__» сентября 2023 г.

/Родионов А.С./

Оценка « _____ »

Новосибирск 2023

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ	3
ОСНОВНАЯ ЧАСТЬ	3
ЗАКЛЮЧЕНИЕ	5
ПРИЛОЖЕНИЯ	5

ВВЕДЕНИЕ

Целью данной лабораторной работы является анализ представленных статистических данных и выдвижение, доказательство любой гипотезы из области математической статистики.

ОСНОВНАЯ ЧАСТЬ

Для начала я визуализировал данные из Таблицы 1 в Microsoft Excel:



Исходя из данного графика, можно выдвинуть гипотезу о том, что данные под заголовками «Яремная область» и «Сердце» могут иметь корреляцию между собой — их линии часто имеют похожее поведение и будто «движутся друг за другом». Также с увеличением шкалы «Возраст» можно заметить, что «резких падений» линий графика становится больше — появляется повод проверить корреляции между «Яремной областью» и «Возрастом» и «Сердцем» и «Возрастом».

Для доказательства гипотез, я использовал метод корреляции Пирсона, который позволяет измерить степень линейной взаимосвязи между двумя количественными переменными.

Метод корреляции Пирсона

Коэффициент корреляции Пирсона (также известный как корреляция Пирсона) - это статистическая мера, используемая для измерения степени линейной связи между двумя переменными. Коэффициент корреляции Пирсона вычисляется следующим образом:

$$r_{xy} = \frac{\sum (x_i - M_x) (y_i - M_y)}{\sqrt{\sum (x_i - M_x)^2 \cdot \sum (y_i - M_y)^2}}$$

где:

- X_i и Y_i - значения переменных x и y соответственно.

- M_x и M_y - средние значения переменных x и y соответственно.

Коэффициент корреляции Пирсона находится в диапазоне от -1 до 1. Значение -1 указывает на полную отрицательную линейную корреляцию, значение 1 - на положительную линейную корреляцию, а значение 0 - на отсутствие линейной корреляции.

Важно понимать, как интерпретировать значения этого коэффициента:

Корреляция близка к 1 ($> +0.8$):

Высокая положительная линейная корреляция. Это означает, что две переменные сильно и положительно взаимосвязаны: при увеличении одной переменной другая тоже увеличивается, и наоборот. Например, если корреляция между уровнем образования и заработной платой составляет +0.9, это может указывать на то, что люди с более высоким уровнем образования обычно имеют более высокую заработную плату.

Корреляция близка к 0 (< 0.15):

Отсутствие или очень слабая линейная корреляция. Это говорит о том, что две переменные не имеют значимой линейной связи. Изменения в одной переменной не влияют на изменения в другой.

Корреляция близка к -1 (< -0.8):

Высокая отрицательная линейная корреляция. Это означает, что две переменные сильно и отрицательно взаимосвязаны: при увеличении одной переменной другая уменьшается, и наоборот. Например, если корреляция между количеством часов сна и уровнем стресса составляет -0.85, это может указывать на то, что более долгий сон связан с меньшим уровнем стресса.

Корреляция близка к 0.5 или -0.5:

Средняя (умеренная) корреляция. Означает, что с увеличением (уменьшением) значений одной переменной, значения другой переменной склонны увеличиваться (уменьшаться), но это происходит не очень сильно.

Важно отметить, что оценка корреляции Пирсона не дает информации о причинно-следственной связи между переменными. Высокая корреляция между двумя переменными может быть вызвана влиянием третьей переменной или быть случайным статистическим явлением. Для более глубокого анализа необходимо проводить дополнительные исследования и учитывать контекст.

Произведение вычислений

Для проведения доказательных вычислений я написал программу на языке программирования C++ (Приложение 1), которая вычисляет коэффициент корреляции Пирсона, исходя из массивов представленных данных (в данном случае используется выборка для мужского населения) (Таблица 1).

После проведения вычислений с использованием метода корреляции Пирсона на основе предоставленных данных, я получил следующие результаты (Приложение 2):

1. Корреляция между "Яремной областью" и "Сердцем" составляет 0.940709.
2. Корреляция между "Яремной областью" и "Возрастом" составляет -0.401191.

3. Корреляция между "Сердцем" и "Возрастом" составляет -0.334889.

Интерпретация результатов:

1. Между "Яремной областью" и "Сердцем" наблюдается очень сильная положительная линейная корреляция (коэффициент корреляции очень близок к 1). Это означает, что увеличение значений "Яремной области" обычно сопровождается увеличением значений "Сердца", и наоборот.
2. Между "Яремной областью" и "Возрастом" существует несильная отрицательная корреляция. Это указывает на то, что возраст и "Яремная область" имеют некоторую обратную связь, но она не является сильной.
3. Между "Сердцем" и "Возрастом" также существует несильная отрицательная корреляция. Это означает, что возраст и "Сердце" имеют некоторую обратную связь, но она также не является сильной.

ЗАКЛЮЧЕНИЕ

На основе проведённых вычислений и анализа данных, я могу заключить, что существует очень сильная положительная корреляция между массивами данных "Яремная область" и "Сердце". Однако между "Яремной областью" и "Возрастом", а также между "Сердцем" и "Возрастом" наблюдается несильная отрицательная корреляция.

Эти результаты могут быть полезными при дальнейшем анализе данных и принятии решений, связанных с исследованием влияния "Яремной области", "Сердца" и "Возраста" на какие-либо явления или состояния.

Моё субъективное непрофессиональное мнение касательно результатов вычислений можно описать так: анатомически яремная область и сердце имеют сильную взаимосвязь — обе отвечают за кровеносную систему в теле человека и закономерно измерения в области сердца положительно коррелируют с измерениями в яремной области, где находятся многие артерии и вены. Отрицательная корреляция показателей этих областей с возрастом пациентов может свидетельствовать о снижении тонуса кровеносных сосудов в связи со старостью.

ПРИЛОЖЕНИЯ

Таблица 1. Данные для мужского населения, сжатые в горизонтальном виде для более удобного представления

Яремная область

31, 62, 62, 62, 62, 70, 68, 60, 70, 72, 60, 65, 70, 82, 74, 78, 80, 72,
56, 60, 70, 74, 64, 90, 74, 74, 82, 70, 72, 76, 68, 79, 62, 80, 78, 78,
54, 66, 68, 72, 70, 74, 58, 80, 38, 78, 68, 58, 61, 84, 80, 70, 72, 54,
84, 66, 58, 62, 60, 58, 50, 62, 58, 60, 64, 76, 80, 70, 26, 34, 71, 74,
66, 30, 74, 64, 19, 74, 29, 76, 70, 79, 60, 68, 74, 68, 35, 70, 66, 50,

22, 55, 68, 67, 74, 62, 28, 19, 70, 68, 30, 71, 68, 25, 69, 74, 63, 29,
53, 32, 62, 36, 26, 74, 70, 64, 70, 65, 58, 20, 57, 29, 52, 21, 30, 56,
68, 62, 71, 32, 33, 60, 39, 54, 67, 43, 62, 56, 25, 62, 72, 22, 68, 56,
30, 54, 27, 12, 54, 60, 60, 42, 19, 24, 65, 62, 19, 72, 60, 64, 66, 56,
27, 19, 56, 26, 25, 41, 22, 70, 22, 60, 16, 58, 65, 67, 52, 21, 21, 70,
58, 27, 22, 48, 62, 15, 65, 19, 15, 66, 14, 60, 66, 54, 70, 60, 60, 70,
60, 58, 64, 64, 60, 52, 48, 44, 60, 62, 69

Сердце

32, 64, 68, 73, 68, 76, 70, 66, 70, 70, 66, 70, 72, 82, 74, 82, 89, 78,
66, 60, 78, 80, 66, 82, 80, 82, 82, 80, 74, 70, 65, 69, 66, 76, 80, 90,
58, 75, 70, 68, 68, 78, 55, 72, 60, 76, 70, 78, 72, 90, 88, 70, 74, 58,
88, 68, 60, 68, 64, 60, 54, 78, 63, 72, 64, 75, 80, 65, 26, 30, 60, 78,
70, 38, 74, 70, 22, 70, 40, 74, 72, 82, 60, 72, 65, 72, 36, 78, 70, 64,
22, 40, 77, 71, 76, 70, 33, 22, 72, 70, 33, 72, 66, 28, 71, 72, 69, 29,
62, 32, 63, 40, 35, 76, 72, 72, 58, 68, 65, 34, 64, 38, 34, 39, 34, 62,
68, 68, 58, 35, 39, 60, 41, 62, 75, 49, 68, 70, 38, 74, 77, 38, 72, 66,
34, 66, 37, 23, 62, 70, 60, 46, 24, 30, 68, 54, 25, 74, 66, 64, 68, 64,
37, 27, 60, 35, 33, 49, 36, 71, 37, 61, 31, 70, 68, 70, 63, 30, 30, 74,
72, 38, 34, 54, 74, 24, 60, 30, 26, 72, 22, 64, 72, 62, 71, 70, 72, 64,
64, 70, 76, 72, 64, 72, 72, 57, 74, 58, 70

Возраст

7, 10, 10, 10, 10, 10, 11, 11, 11, 11, 11, 11, 11, 11, 12, 12, 12, 12,
12, 12, 12, 12, 12, 12, 12, 12, 12, 13, 13, 13, 13, 13, 13, 13, 13, 13,
13, 13, 13, 13, 13, 13, 13, 13, 13, 13, 13, 14, 14, 14, 14, 14, 14, 14,
14, 14, 14, 14, 14, 14, 14, 14, 14, 14, 15, 15, 16, 16, 16, 16, 17, 17, 18,
20, 21, 21, 21, 21, 22, 24, 24, 24, 25, 27, 28, 28, 28, 28, 28, 29, 30,
30, 30, 30, 30, 31, 31, 32, 32, 33, 33, 33, 33, 33, 34, 34, 34, 35, 35, 35,
36, 36, 36, 37, 37, 37, 37, 38, 38, 38, 38, 38, 39, 39, 39, 39, 39, 40,
40, 40, 40, 40, 40, 40, 40, 41, 41, 41, 41, 42, 42, 42, 42, 42, 42, 42,
42, 42, 42, 42, 43, 43, 43, 43, 43, 44, 44, 44, 44, 45, 45, 46, 46, 46,
46, 46, 47, 47, 47, 48, 48, 48, 48, 49, 49, 49, 49, 49, 50, 50, 50, 50,
50, 50, 50, 51, 51, 51, 52, 52, 52, 52, 52, 53, 53, 54, 54, 54, 54, 56,
56, 56, 57, 58, 59, 59, 59, 60, 60, 60, 62

Приложение 1. Исходный код программы — файл **pearson.cpp**

```
#include <cmath>
#include <iostream>
#include <vector>

// среднее значение (M / Mean)
double calculateMean(const std::vector<double>& data) {
    double sum = 0.0;
    for (const double& value : data) {
        sum += value;
    }
    return sum / data.size();
}
```

```

double calculatePearsonCorrelation(const std::vector<double>& x,
                                   const std::vector<double>& y) {
    // Проверка, что обе выборки имеют одинаковый размер
    if (x.size() != y.size()) {
        throw std::invalid_argument("Выборки должны иметь одинаковый размер");
    }

    double meanX = calculateMean(x);
    double meanY = calculateMean(y);
    double sumProduct = 0.0;
    double sumXSquare = 0.0;
    double sumYSquare = 0.0;

    for (size_t i = 0; i < x.size(); ++i) {
        double deltaX = x[i] - meanX;
        double deltaY = y[i] - meanY;
        sumProduct += deltaX * deltaY;
        sumXSquare += deltaX * deltaX;
        sumYSquare += deltaY * deltaY;
    }

    double correlation = sumProduct / (sqrt(sumXSquare) * sqrt(sumYSquare));
    return correlation;
}

int main() {
    // Данные из таблицы
    std::vector<double> jugular_area = {
        31, 62, 62, 62, 62, 70, 68, 60, 70, 72, 60, 65, 70, 82, 74, 78, 80, 72,
        56, 60, 70, 74, 64, 90, 74, 74, 82, 70, 72, 76, 68, 79, 62, 80, 78, 78,
        54, 66, 68, 72, 70, 74, 58, 80, 38, 78, 68, 58, 61, 84, 80, 70, 72, 54,
        84, 66, 58, 62, 60, 58, 50, 62, 58, 60, 64, 76, 80, 70, 26, 34, 71, 74,
        66, 30, 74, 64, 19, 74, 29, 76, 70, 79, 60, 68, 74, 68, 35, 70, 66, 50,
        22, 55, 68, 67, 74, 62, 28, 19, 70, 68, 30, 71, 68, 25, 69, 74, 63, 29,
        53, 32, 62, 36, 26, 74, 70, 64, 70, 65, 58, 20, 57, 29, 52, 21, 30, 56,
        68, 62, 71, 32, 33, 60, 39, 54, 67, 43, 62, 56, 25, 62, 72, 22, 68, 56,
        30, 54, 27, 12, 54, 60, 60, 42, 19, 24, 65, 62, 19, 72, 60, 64, 66, 56,
        27, 19, 56, 26, 25, 41, 22, 70, 22, 60, 16, 58, 65, 67, 52, 21, 21, 70,
        58, 27, 22, 48, 62, 15, 65, 19, 15, 66, 14, 60, 66, 54, 70, 60, 60, 70,
        60, 58, 64, 64, 60, 52, 48, 44, 60, 62, 69};
    std::vector<double> heart = {
        32, 64, 68, 73, 68, 76, 70, 66, 70, 70, 66, 70, 72, 82, 74, 82, 89, 78,
        66, 60, 78, 80, 66, 82, 80, 82, 82, 80, 74, 70, 65, 69, 66, 76, 80, 90,
        58, 75, 70, 68, 68, 78, 55, 72, 60, 76, 70, 78, 72, 90, 88, 70, 74, 58,
        88, 68, 60, 68, 64, 60, 54, 78, 63, 72, 64, 75, 80, 65, 26, 30, 60, 78,
        70, 38, 74, 70, 22, 70, 40, 74, 72, 82, 60, 72, 65, 72, 36, 78, 70, 64,
        22, 40, 77, 71, 76, 70, 33, 22, 72, 70, 33, 72, 66, 28, 71, 72, 69, 29,
        62, 32, 63, 40, 35, 76, 72, 72, 58, 68, 65, 34, 64, 38, 34, 39, 34, 62,
        68, 68, 58, 35, 39, 60, 41, 62, 75, 49, 68, 70, 38, 74, 77, 38, 72, 66,
        34, 66, 37, 23, 62, 70, 60, 46, 24, 30, 68, 54, 25, 74, 66, 64, 68, 64,
        37, 27, 60, 35, 33, 49, 36, 71, 37, 61, 31, 70, 68, 70, 63, 30, 30, 74,

```

```

72, 38, 34, 54, 74, 24, 60, 30, 26, 72, 22, 64, 72, 62, 71, 70, 72, 64,
64, 70, 76, 72, 64, 72, 72, 57, 74, 58, 70};
std::vector<double> age = {
7, 10, 10, 10, 10, 10, 11, 11, 11, 11, 11, 11, 11, 12, 12, 12, 12,
12, 12, 12, 12, 12, 12, 12, 12, 13, 13, 13, 13, 13, 13, 13, 13,
13, 13, 13, 13, 13, 13, 13, 13, 13, 13, 14, 14, 14, 14, 14, 14, 14,
14, 14, 14, 14, 14, 14, 14, 14, 15, 15, 16, 16, 16, 16, 17, 17, 18,
20, 21, 21, 21, 21, 22, 24, 24, 24, 25, 27, 28, 28, 28, 28, 29, 30,
30, 30, 30, 30, 31, 31, 32, 32, 33, 33, 33, 33, 34, 34, 34, 35, 35, 35,
36, 36, 36, 37, 37, 37, 37, 38, 38, 38, 38, 38, 39, 39, 39, 39, 40,
40, 40, 40, 40, 40, 40, 40, 41, 41, 41, 41, 42, 42, 42, 42, 42, 42, 42,
42, 42, 42, 42, 43, 43, 43, 43, 43, 44, 44, 44, 44, 45, 45, 46, 46, 46,
46, 46, 47, 47, 47, 48, 48, 48, 48, 49, 49, 49, 49, 49, 50, 50, 50, 50,
50, 50, 50, 51, 51, 51, 52, 52, 52, 52, 52, 53, 53, 54, 54, 54, 54, 56,
56, 56, 57, 58, 59, 59, 59, 60, 60, 60, 62};

// Вычисление корреляции между "Яремной областью" и "Сердцем"
double correlation_jugular_area_heart =
    calculatePearsonCorrelation(jugular_area, heart);
std::cout << "Корреляция между \"Яремной областью\" и \"Сердцем\": \"
    << correlation_jugular_area_heart << '\n';

// Вычисление корреляции между "Яремной областью" и "Возрастом"
double correlation_jugular_area_age =
    calculatePearsonCorrelation(jugular_area, age);
std::cout << "Корреляция между \"Яремной областью\" и \"Возрастом\": \"
    << correlation_jugular_area_age << '\n';

// Вычисление корреляции между "Сердцем" и "Возрастом"
double correlation_heart_age =
    calculatePearsonCorrelation(heart, age);
std::cout << "Корреляция между \"Сердцем\" и \"Возрастом\": \"
    << correlation_heart_age << '\n';

return 0;
}

```

Приложение 2. Демонстрация компиляции программы и результатов её выполнения

```

allenvox@MacBook-Pro-Yuriy terver1 % make
g++ -pedantic -std=c++17 -o pearson.cppout pearson.cpp
allenvox@MacBook-Pro-Yuriy terver1 % ./pearson.cppout
Корреляция между "Яремной областью" и "Сердцем:" 0.940709
Корреляция между "Яремной областью" и "Возрастом": -0.401191
Корреляция между "Сердцем" и "Возрастом": -0.334889
allenvox@MacBook-Pro-Yuriy terver1 %

```