МИНИСТЕРСТВО ПРОСВЕЩЕНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ

ВЫСШЕГО ОБРАЗОВАНИЯ

«РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ ПЕДАГОГИЧЕСКИЙ УНИВЕРСИТЕТ им. А.И. ГЕРЦЕНА»



Направление подготовки

09.03.01 – Информатика и вычислительная техника

Профиль «Технологии разработки программного обеспечения»

**Лабораторная работа №7. Часть 1**

**“Корреляционный анализ. Вычисление коэффициентов корреляции”**

|  | Работу выполнили:  Балаев Жамал,  Васильева Марина,  Иванов Никита,  Шардт Максим,  Рожков Максим  очная форма обучения  курс: 2; группа: ИВТ-1.1 |
| --- | --- |
|  | Научный руководитель:  Профессор Власова Елена Зотиковна |

Санкт-Петербург

2023

## **Содержание**

[ЛР №7. Часть 1. Выполнил Балаев Ж. Б. 3](#_nnr3ru9fslfs)

[ЛР №7. Часть 1. Выполнила Васильева М. А. 21](#_wrjrxwnpyekz)

[ЛР №7. Часть 1. Выполнил Иванов Н. Р. 39](#_7xthnbmrkkt)

[ЛР №7. Часть 1. Выполнил Рожков М. В. 57](#_kz7uunep9wm0)

[ЛР №7. Часть 1. Выполнил Шардт М. А. 75](#_dlwjjjo7k8mx)

## 

## 

## **Лабораторная работа №7. Часть 1**

*Выполнил Балаев Ж. Б.*

**Цель работы:** Вычисление коэффициентов корреляции.

**Оборудования:** ПК, Microsoft Excel.

*Использованные формулы:*

1. Коэффициент корреляции — параметр, который характеризует степень линейной взаимосвязи между двумя выборками, рассчитывается по формуле:

1. Коэффициент ранговой корреляции Спирмена:

,

где - квадраты разности рангов,

n - число наблюдений.

1. Статистика гипотезы Стьюдента:

*Задание 2:*

Построить корреляционное поле. Сделать предположение о форме и направлении взаимосвязи двух исследуемых показателей. Вычислить коэффициент линейной корреляции Пирсона и коэффициент ранговой корреляции Спирмена.

*Пример 1:*

10 школьникам были даны тесты на наглядно-образное и вербальное мышление. Измерялось среднее время решения заданий теста в секундах. Исследователя интересует вопрос: существует ли взаимосвязь между временем решения этих задач? Переменная X — обозначает среднее время решения наглядно-образных, а переменная Y— среднее время решения вербальных заданий тестов.

| № испытуемых | X | Y |
| --- | --- | --- |
| 1 | 19 | 17 |
| 2 | 32 | 7 |
| 3 | 33 | 17 |
| 4 | 44 | 28 |
| 5 | 28 | 27 |
| 6 | 35 | 31 |
| 7 | 39 | 20 |
| 8 | 39 | 17 |
| 9 | 44 | 35 |
| 10 | 44 | 43 |

*Решение:*

Исходя из исходных данных сделаем вспомогательную таблицу:

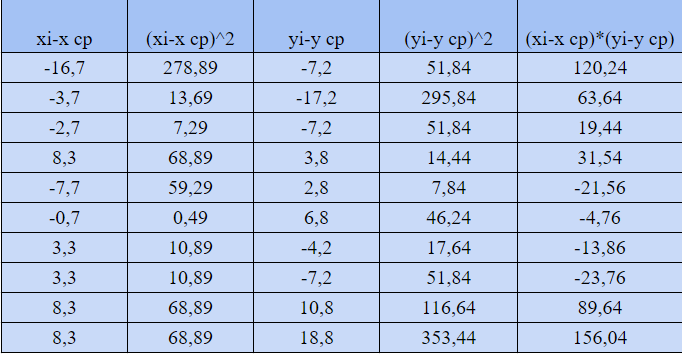


Табл.1. Вспомогательная таблица с промежуточными вычислениями

1. Нулевая гипотеза – взаимосвязь между временем решения существует.
2. Конкурирующая гипотеза – взаимосвязь между временем решения отсутствует.

Найдем значение коэффициента корреляции по формуле (1):



Далее по таблице “Критические значения корреляции rzy Пирсона” определим критические значения для полученного коэффициента корреляции, где k = n - 2 = 10 - 2 = 8, P = 0,05 (уровень значимости):



Для автоматического поиска критического значения по таблице используем функции ПОИСК и ИНДЕКС.

= 0,63 > 0,54 , следовательно, гипотеза Н1 отвергается и принимается гипотеза H0.

Построим корреляционное поле (рис.1):

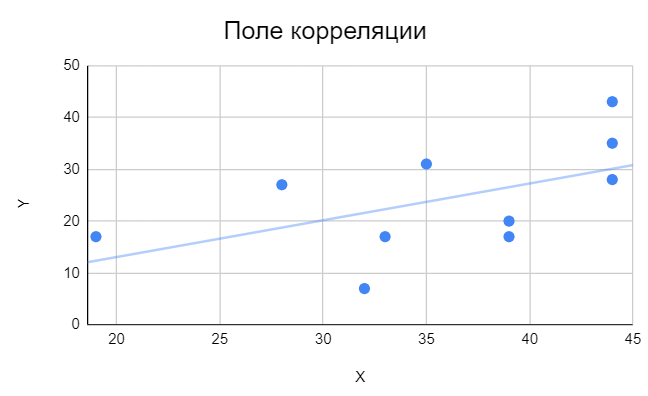


Рис.1. Корреляционное поле

**Ответ:** на уровне значимости 0,05 нулевую гипотезу Н0 принимаем.

*Пример 2:*

Преподавателю и студенту было предложено расположить 10 профессий в порядке их общественной значимости. Ответы перечислены в таблице 1:

***Таблица 1.***

| Оценка преподавателя, xi | Профессии | Оценка студента, yi |
| --- | --- | --- |
| 3 | профессор | 2 |
| 1 | врач | 1 |
| 4 | учитель школы | 7 |
| 2 | директор магазина | 4 |
| 8 | бухгалтер | 5 |
| 6 | банкир | 3 |
| 9 | водитель | 9 |
| 5 | журналист | 8 |
| 10 | ди-джей | 10 |

Какова корреляция рангов между двумя рядами оценок? Одинаково ли мнение преподавателя и студента по этому вопросу?

*Решение:*

Исходя из исходных данных составим вспомогательную таблицу разности рангов, их квадраты и суммы:

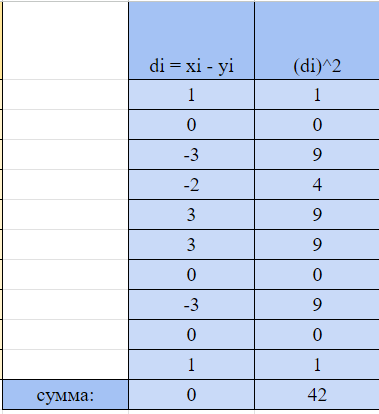


Табл. 2. Вспомогательная таблица с промежуточными вычислениями

Найдем коэффициент ранговой корреляции Спирмена по формуле (2):



Далее необходимо проверить, существует ли положительная корреляционная связь между мнениями преподавателя и студента. Для этого (при n ≥ 10) используем формулу (3) с k = (n – 2) степенями свободы:



1. Нулевая гипотеза – коэффициент корреляции не является статистически значимым ().
2. Конкурирующая гипотеза – существует положительная корреляционная зависимость ().

Далее по таблице “Критические точки t-распределения Стьюдента” определим критическое значение статистики Стьюдента, где k = n - 2 = 10 - 2 = 8, уровень значимости α = 0.05 для односторонней (правосторонней) критической области:



Для автоматического поиска критического значения по таблице используем функции ПОИСК и ИНДЕКС.

t расч > t кр(3,16>1,86). Следовательно, связь между мнениями преподавателя и студента является статистически значимой при 5%-ном уровне значимости.

Построим корреляционное поле (рис.2):

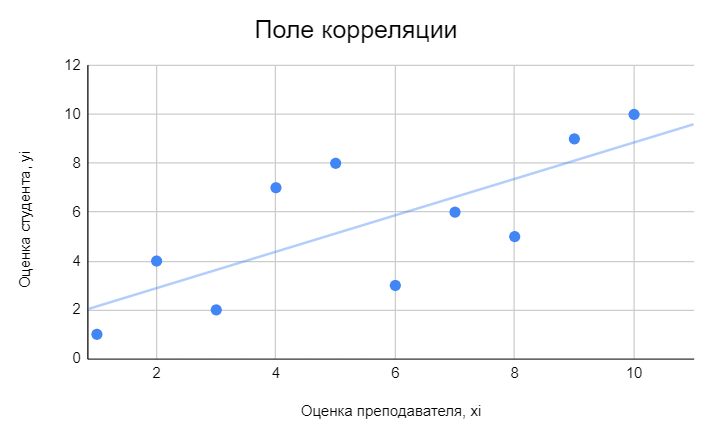


Рис.2. Корреляционное поле

**Ответ:** на уровне значимости 0,05 нулевую гипотезу отвергаем.

*Задание 3:*

Построить корреляционное поле. Сделать предположение о форме и направлении взаимосвязи двух исследуемых показателей. Найти значения коэффициентов ранговой корреляции Спирмена.

*Задача 3.1*

С помощью коэффициента ранговой корреляции установить зависимость между стажем практической работы и временем решения контрольной задачи у 10 программистов на основе следующих данных:

| Номера испытуемых | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Стаж (в мес.) | 32 | 15 | 16 | 18 | 20 | 28 | 21 | 29 | 23 | 17 |
| Время решения (в мин.) | 12 | 24 | 23 | 21 | 20 | 9 | 11 | 10 | 15 | 16 |

*Решение:*

Сформулируем статистические гипотезы:

1. Нулевая гипотеза – коэффициент корреляции не является статистически значимым ().
2. Конкурирующая гипотеза – существует положительная корреляционная зависимость ().

Определим ранги оценкам стажа xi, располагая оценки в возрастающем порядке и сохраняя связь с оценками времени решения. Ранги пронумеруем от 1 до 10.

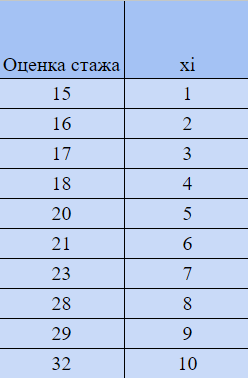


Табл. 3. Вспомогательная таблица с промежуточными вычислениями

Аналогично определим ранги оценкам времени решения yi. И также пронумеруем их от 1 до 10, располагая по возрастанию.

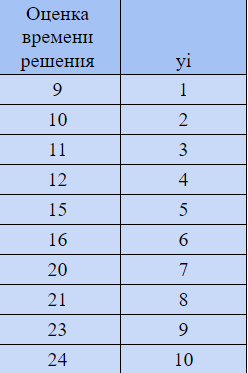


Табл. 4. Вспомогательная таблица с промежуточными вычислениями

Для пояснения процесса вычисления рангов yi возьмем ранг y1. Мы рассматриваем оценку времени решения программиста, который по оценке стажа занимает 1 место (xi = 1; 15 мес.) Теперь найдем оценку времени решения этого программиста (24 мин.) и видим, что это время решения получило ранг yi=10. В сводную таблицу заносим xi = 1, yi=10. Проделываем аналогичные действия для остальных значений и получаем сводную таблицу:

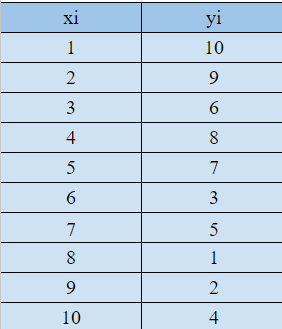


Табл. 5. Сводная таблица результатов

Исходя из полученных данных составим вспомогательную таблицу разности рангов, их квадраты и суммы:

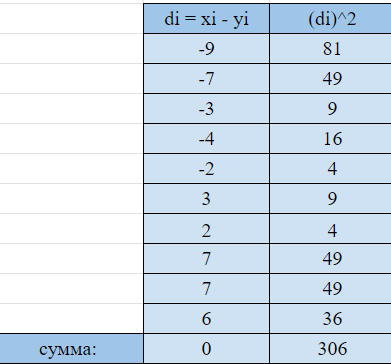


Табл. 6. Вспомогательная таблица с промежуточными вычислениями

Найдем коэффициент ранговой корреляции Спирмена по формуле (2):



Далее необходимо проверить, существует ли положительная корреляционная связь между оценкой стажа и оценкой времени решения задач. Для этого (при n ≥ 10) используем формулу (3) с k = (n – 2) степенями свободы:



Далее по таблице “Критические точки t-распределения Стьюдента” определим критическое значение статистики Стьюдента, где k = n - 2 = 10 - 2 = 8, уровень значимости α = 0.05 для односторонней (правосторонней) критической области:



Для автоматического поиска критического значения по таблице используем функции ПОИСК и ИНДЕКС.

|t расч |> t кр(4,65>1,86). Следовательно, связь между оценкой стажа и оценкой времени решения является статистически значимой при 5%-ном уровне значимости. Также можно сказать, что связь достаточно сильная и обратная по направлению (видно по корреляционному полю).

Построим корреляционное поле (рис.3):

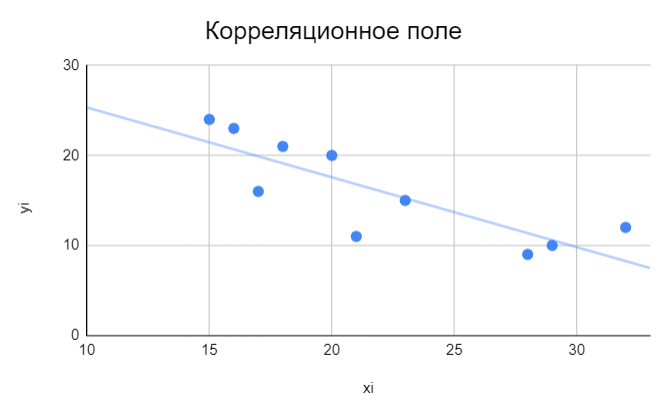


Рис.3. Корреляционное поле

**Ответ:** на уровне значимости 0,05 нулевую гипотезу отвергаем.

*Задача 3.2*

Три арбитра оценили мастерство 10 спортсменов, в итоге были получены три последовательности рангов (в первой строке приведены ранги арбитра А, во второй – ранги арбитра В, в третьей – ранги арбитра С):

| xi | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| yi | 3 | 10 | 7 | 2 | 8 | 5 | 6 | 9 | 1 | 4 |
| zi | 6 | 2 | 1 | 3 | 9 | 4 | 5 | 7 | 10 | 8 |

Определить пару арбитров, оценки которых наиболее согласуются, используя коэффициент ранговой корреляции Спирмена. Построить корреляционное поле.

*Решение:*

Для определения коэффициента корреляции Спирмена найдем разности рангов, их квадраты и сумму для оценок каждой из пар арбитров.

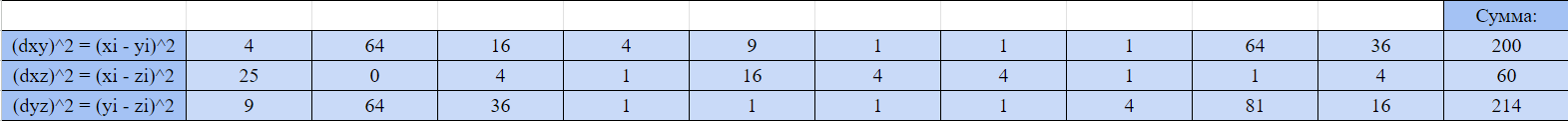


Табл. 7. Вспомогательная таблица с промежуточными вычислениями

Также для каждой пары арбитров вычислим коэффициент Спирмена и определим t-критерий Стьюдента tвыч по формулам (2) и (3). Далее по таблице “Критические точки t-распределения Стьюдента” определим критическое значение статистики Стьюдента, где k = n - 2 = 10 - 2 = 8, уровень значимости α = 0.05 для односторонней (правосторонней) критической области:

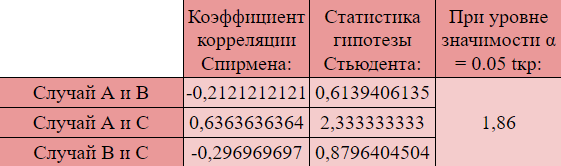


Табл. 8. Таблица с итоговыми результатами вычислений.

Для автоматического поиска критического значения по таблице используем функции ПОИСК и ИНДЕКС.

Построим поле корреляции для графического изображения зависимостей каждой пары арбитров (рис.4-рис.6):

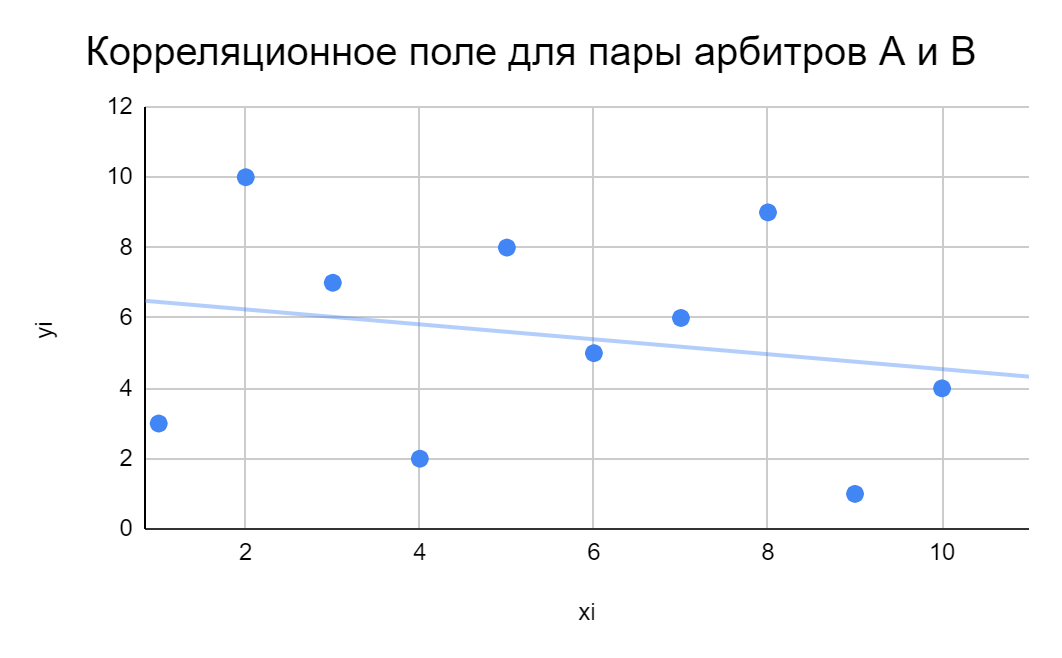


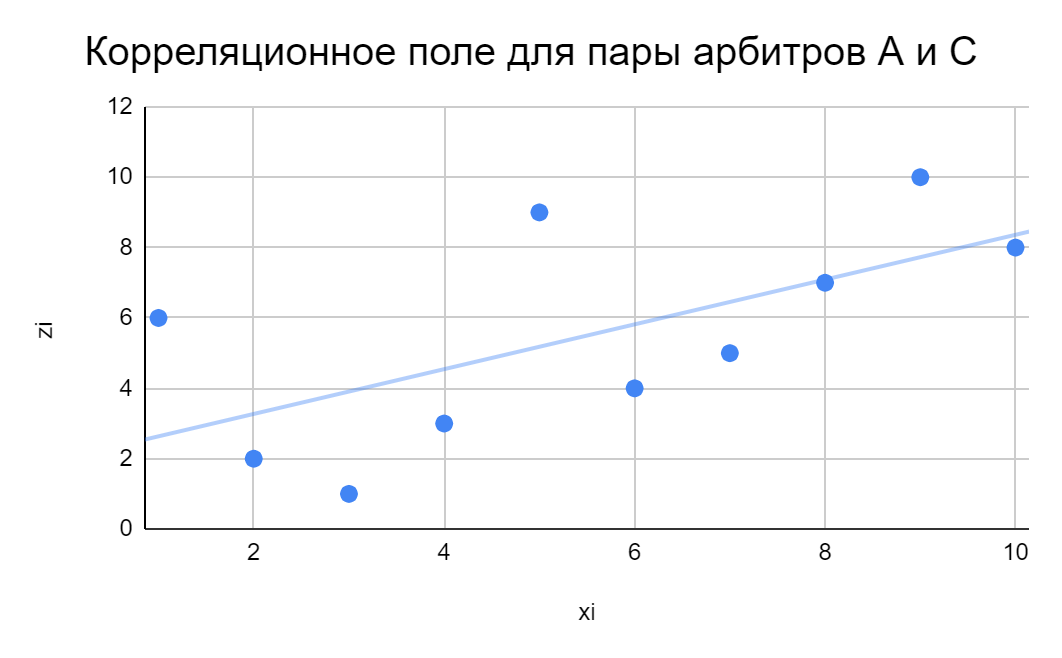
Рис.4. Корреляционное поле

Рис.5. Корреляционное поле

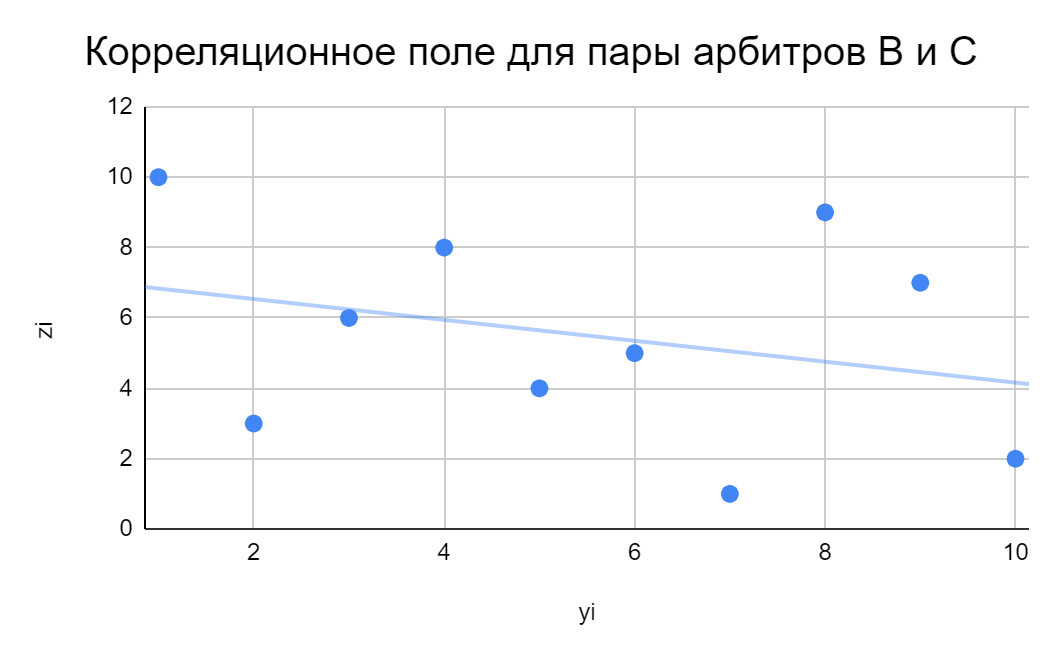


Рис.6. Корреляционное поле

Сравнивая значения с критическим, видим, что наиболее согласуются оценки арбитров А и С. Данный коэффициент ранговой корреляции наибольший по величине.

**Ответ:** Пара арбитров, оценки которых наиболее согласуются, это пара A и C.

*Задание 4:*

*Задача 4.1*

Построить корреляционное поле. Сделать предположение о форме и направлении взаимосвязи двух исследуемых показателей. Найти значение коэффициентов линейной корреляции. Необходимо определить взаимосвязь характеристик: агрессивности и IQ у школьников по полученным данным тестирования.

| № п/п | Данные по агрессивности X\_agr | Данные по IQ Y\_iq |
| --- | --- | --- |
| 1 | 24 | 100 |
| 2 | 27 | 115 |
| 3 | 26 | 117 |
| 4 | 21 | 119 |
| 5 | 20 | 134 |
| 6 | 31 | 94 |
| 7 | 26 | 105 |
| 8 | 22 | 103 |
| 9 | 20 | 111 |
| 10 | 18 | 124 |
| 11 | 30 | 122 |
| 12 | 29 | 109 |
| 13 | 24 | 110 |
| 14 | 26 | 86 |

*Решение:*

Сформулируем статистические гипотезы:

1. Нулевая гипотеза – взаимосвязь между агрессивностью и IQ у школьников существует.
2. Конкурирующая гипотеза – взаимосвязь между агрессивностью и IQ у школьников отсутствует.

Исходя из исходных данных сделаем вспомогательную таблицу:

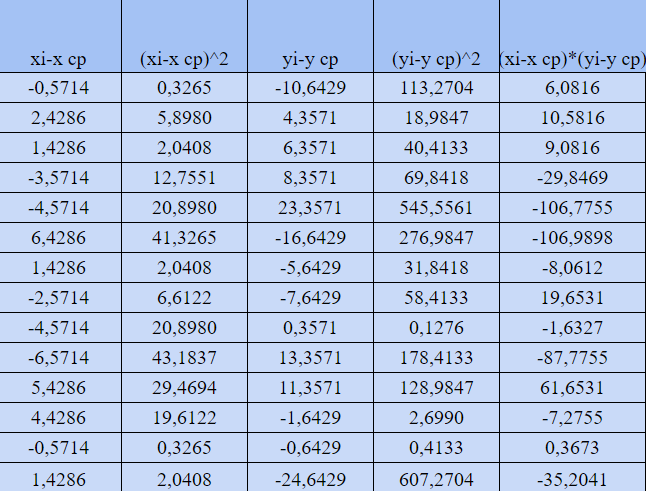


Табл.9. Вспомогательная таблица с промежуточными вычислениями

Найдем значение коэффициента корреляции по формуле (1):



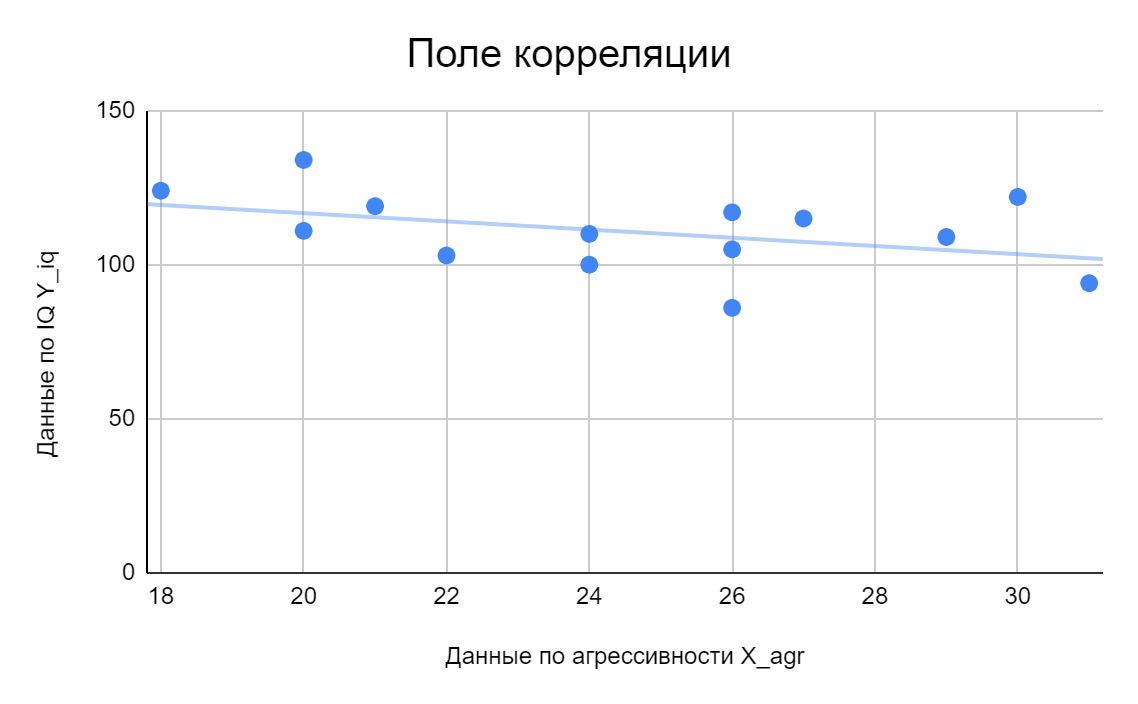
Далее по таблице “Критические значения корреляции rzy Пирсона” определим критические значения для полученного коэффициента корреляции, где k = n - 2 = 14 - 2 = 12, P = 0,05 (уровень значимости):



Для автоматического поиска критического значения по таблице используем функции ПОИСК и ИНДЕКС.

= 0,53 > |-0,42| , следовательно, гипотеза Н1 отвергается и принимается гипотеза H0.

Построим корреляционное поле (рис.7):

Рис.7. Корреляционное поле

**Ответ**: на уровне значимости 0,05 нулевую гипотезу принимаем.

*Задача 4.2*

На основании наблюдений за развивающимся сайтом и изменением его средневзвешенной позиции по основным запросам в поисковой системе необходимо проверить, можно ли говорить о линейной зависимости между позицией сайта и числом посетителей. Построить корреляционное поле.

Исходные данные:

X - число посетителей в сутки;

Y – усредненная позиция сайта в поисковой системе.

В таблице даны значения признаков

| № п/п | Число посетителей в сутки, X | Усредненная позиция сайта в поисковой системе, Y |
| --- | --- | --- |
| 1 | 500 | 5,4 |
| 2 | 790 | 4,2 |
| 3 | 870 | 4 |
| 4 | 1500 | 3,4 |
| 5 | 2300 | 2,5 |
| 6 | 5600 | 1 |
| 7 | 100 | 6,1 |
| 8 | 20 | 8,2 |
| 9 | 5 | 14,6 |

*Решение:*

Сформулируем статистические гипотезы:

1. Нулевая гипотеза – взаимосвязь между позицией сайта и числом посетителей существует.
2. Конкурирующая гипотеза – взаимосвязь между позицией сайта и числом посетителей отсутствует.

Исходя из исходных данных сделаем вспомогательную таблицу:

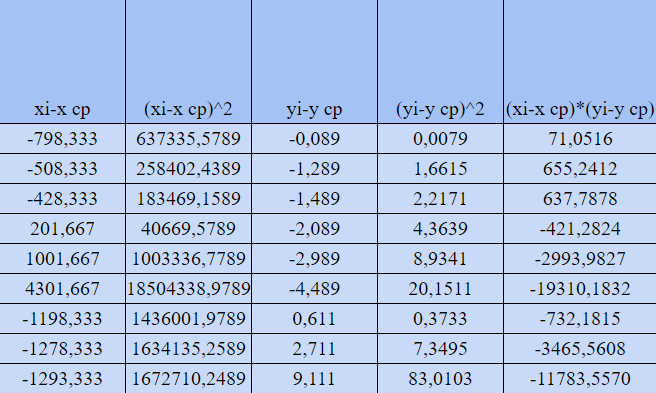


Табл.10. Вспомогательная таблица с промежуточными вычислениями

Найдем значение коэффициента корреляции по формуле (1):



Далее по таблице “Критические значения корреляции rzy Пирсона” определим критические значения для полученного коэффициента корреляции, где k = n - 2 = 9 - 2 = 7, P = 0,05 (уровень значимости):



Для автоматического поиска критического значения по таблице используем функции ПОИСК и ИНДЕКС.

= 0,67 > |-0,65| , следовательно, гипотеза Н1 отвергается и принимается гипотеза H0.

Построим корреляционное поле (рис.8):

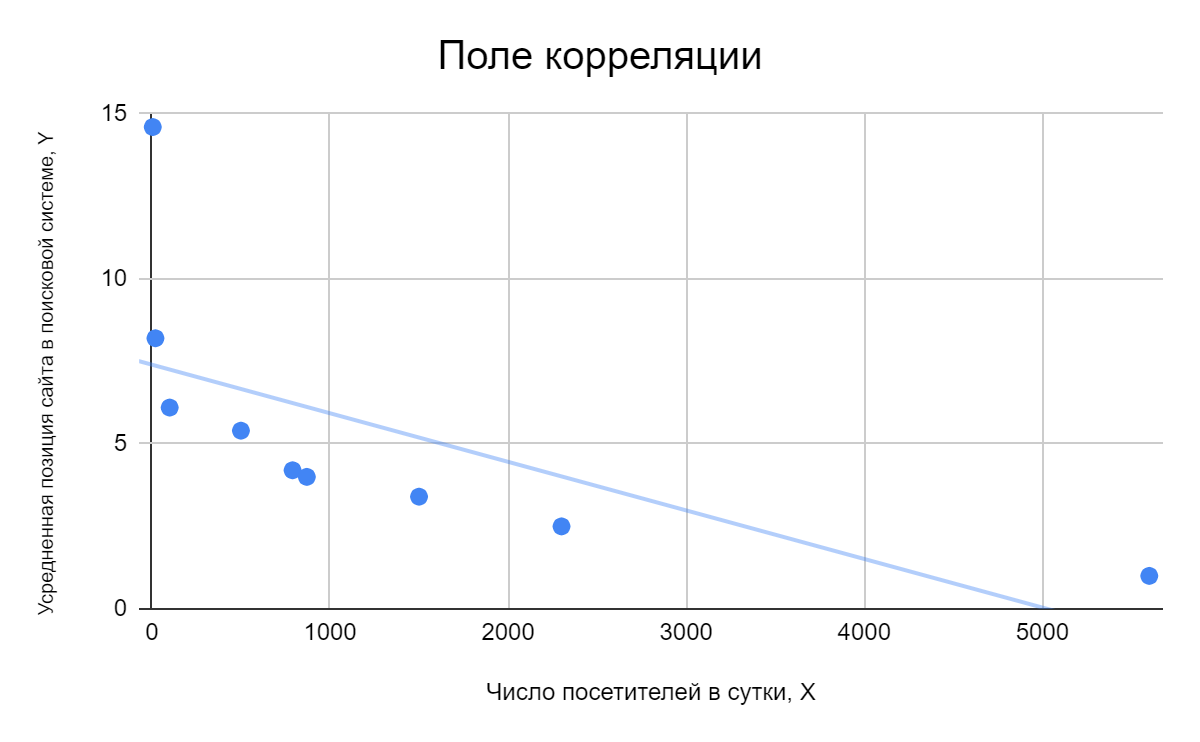


Рис.8. Корреляционное поле

**Ответ**: на уровне значимости 0,05 нулевую гипотезу принимаем.

## **Лабораторная работа №7. Часть 1**

*Выполнила Васильева М. А.*

**Цель работы:** Вычисление коэффициентов корреляции.

**Оборудования:** ПК, Microsoft Excel.

*Использованные формулы:*

1. Коэффициент корреляции — параметр, который характеризует степень линейной взаимосвязи между двумя выборками, рассчитывается по формуле:

1. Коэффициент ранговой корреляции Спирмена:

,

где - квадраты разности рангов,

n - число наблюдений.

1. Статистика гипотезы Стьюдента:

*Задание 2:*

Построить корреляционное поле. Сделать предположение о форме и направлении взаимосвязи двух исследуемых показателей. Вычислить коэффициент линейной корреляции Пирсона и коэффициент ранговой корреляции Спирмена.

*Пример 1:*

10 школьникам были даны тесты на наглядно-образное и вербальное мышление. Измерялось среднее время решения заданий теста в секундах. Исследователя интересует вопрос: существует ли взаимосвязь между временем решения этих задач? Переменная X — обозначает среднее время решения наглядно-образных, а переменная Y— среднее время решения вербальных заданий тестов.

| № испытуемых | X | Y |
| --- | --- | --- |
| 1 | 19 | 17 |
| 2 | 32 | 7 |
| 3 | 33 | 17 |
| 4 | 44 | 28 |
| 5 | 28 | 27 |
| 6 | 35 | 31 |
| 7 | 39 | 20 |
| 8 | 39 | 17 |
| 9 | 44 | 35 |
| 10 | 44 | 43 |

*Решение:*

Исходя из исходных данных сделаем вспомогательную таблицу:

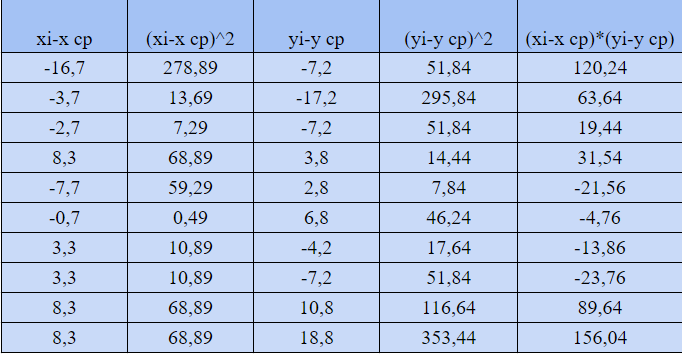


Табл.1. Вспомогательная таблица с промежуточными вычислениями

1. Нулевая гипотеза – взаимосвязь между временем решения существует.
2. Конкурирующая гипотеза – взаимосвязь между временем решения отсутствует.

Найдем значение коэффициента корреляции по формуле (1):



Далее по таблице “Критические значения корреляции rzy Пирсона” определим критические значения для полученного коэффициента корреляции, где k = n - 2 = 10 - 2 = 8, P = 0,05 (уровень значимости):



Для автоматического поиска критического значения по таблице используем функции ПОИСК и ИНДЕКС.

= 0,63 > 0,54 , следовательно, гипотеза Н1 отвергается и принимается гипотеза H0.

Построим корреляционное поле (рис.1):

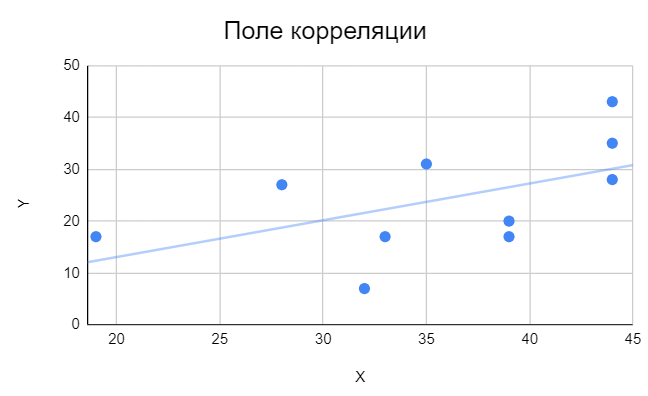


Рис.1. Корреляционное поле

**Ответ:** на уровне значимости 0,05 нулевую гипотезу Н0 принимаем.

*Пример 2:*

Преподавателю и студенту было предложено расположить 10 профессий в порядке их общественной значимости. Ответы перечислены в таблице 1:

***Таблица 1.***

| Оценка преподавателя, xi | Профессии | Оценка студента, yi |
| --- | --- | --- |
| 3 | профессор | 2 |
| 1 | врач | 1 |
| 4 | учитель школы | 7 |
| 2 | директор магазина | 4 |
| 8 | бухгалтер | 5 |
| 6 | банкир | 3 |
| 9 | водитель | 9 |
| 5 | журналист | 8 |
| 10 | ди-джей | 10 |

Какова корреляция рангов между двумя рядами оценок? Одинаково ли мнение преподавателя и студента по этому вопросу?

*Решение:*

Исходя из исходных данных составим вспомогательную таблицу разности рангов, их квадраты и суммы:

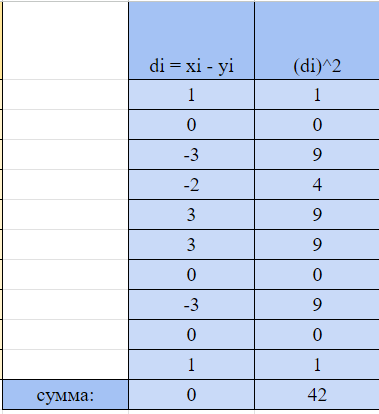


Табл. 2. Вспомогательная таблица с промежуточными вычислениями

Найдем коэффициент ранговой корреляции Спирмена по формуле (2):



Далее необходимо проверить, существует ли положительная корреляционная связь между мнениями преподавателя и студента. Для этого (при n ≥ 10) используем формулу (3) с k = (n – 2) степенями свободы:



1. Нулевая гипотеза – коэффициент корреляции не является статистически значимым ().
2. Конкурирующая гипотеза – существует положительная корреляционная зависимость ().

Далее по таблице “Критические точки t-распределения Стьюдента” определим критическое значение статистики Стьюдента, где k = n - 2 = 10 - 2 = 8, уровень значимости α = 0.05 для односторонней (правосторонней) критической области:



Для автоматического поиска критического значения по таблице используем функции ПОИСК и ИНДЕКС.

t расч > t кр(3,16>1,86). Следовательно, связь между мнениями преподавателя и студента является статистически значимой при 5%-ном уровне значимости.

Построим корреляционное поле (рис.2):

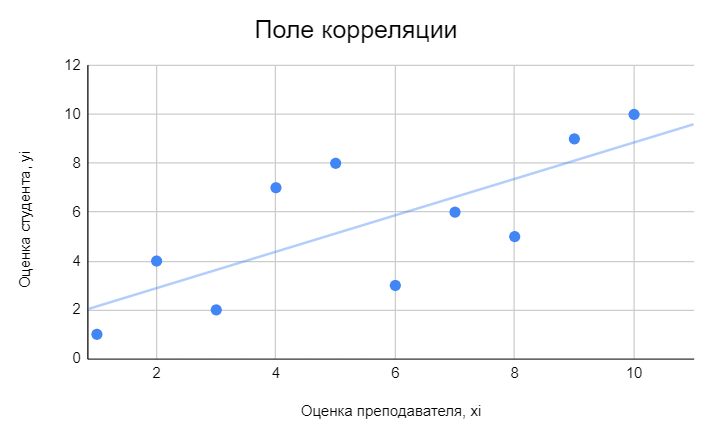


Рис.2. Корреляционное поле

**Ответ:** на уровне значимости 0,05 нулевую гипотезу отвергаем.

*Задание 3:*

Построить корреляционное поле. Сделать предположение о форме и направлении взаимосвязи двух исследуемых показателей. Найти значения коэффициентов ранговой корреляции Спирмена.

*Задача 3.1*

С помощью коэффициента ранговой корреляции установить зависимость между стажем практической работы и временем решения контрольной задачи у 10 программистов на основе следующих данных:

| Номера испытуемых | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Стаж (в мес.) | 32 | 15 | 16 | 18 | 20 | 28 | 21 | 29 | 23 | 17 |
| Время решения (в мин.) | 12 | 24 | 23 | 21 | 20 | 9 | 11 | 10 | 15 | 16 |

*Решение:*

Сформулируем статистические гипотезы:

1. Нулевая гипотеза – коэффициент корреляции не является статистически значимым ().
2. Конкурирующая гипотеза – существует положительная корреляционная зависимость ().

Определим ранги оценкам стажа xi, располагая оценки в возрастающем порядке и сохраняя связь с оценками времени решения. Ранги пронумеруем от 1 до 10.

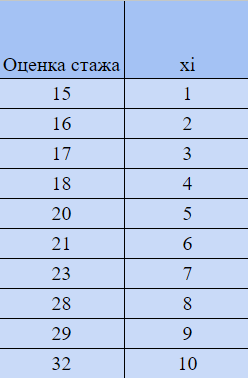


Табл. 3. Вспомогательная таблица с промежуточными вычислениями

Аналогично определим ранги оценкам времени решения yi. И также пронумеруем их от 1 до 10, располагая по возрастанию.

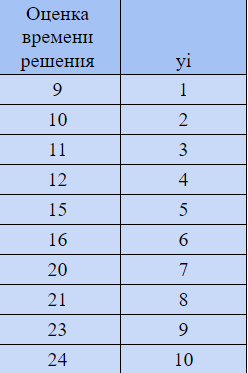


Табл. 4. Вспомогательная таблица с промежуточными вычислениями

Для пояснения процесса вычисления рангов yi возьмем ранг y1. Мы рассматриваем оценку времени решения программиста, который по оценке стажа занимает 1 место (xi = 1; 15 мес.) Теперь найдем оценку времени решения этого программиста (24 мин.) и видим, что это время решения получило ранг yi=10. В сводную таблицу заносим xi = 1, yi=10. Проделываем аналогичные действия для остальных значений и получаем сводную таблицу:

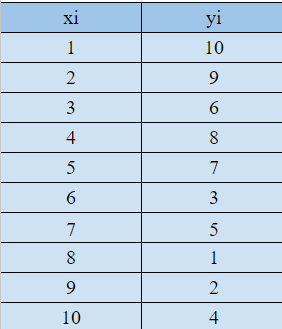


Табл. 5. Сводная таблица результатов

Исходя из полученных данных составим вспомогательную таблицу разности рангов, их квадраты и суммы:

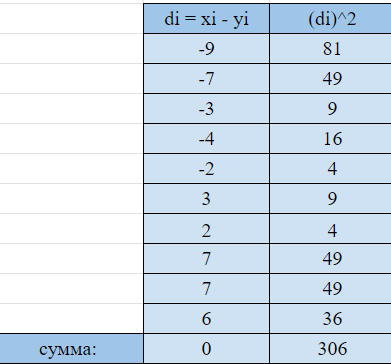


Табл. 6. Вспомогательная таблица с промежуточными вычислениями

Найдем коэффициент ранговой корреляции Спирмена по формуле (2):



Далее необходимо проверить, существует ли положительная корреляционная связь между оценкой стажа и оценкой времени решения задач. Для этого (при n ≥ 10) используем формулу (3) с k = (n – 2) степенями свободы:



Далее по таблице “Критические точки t-распределения Стьюдента” определим критическое значение статистики Стьюдента, где k = n - 2 = 10 - 2 = 8, уровень значимости α = 0.05 для односторонней (правосторонней) критической области:



Для автоматического поиска критического значения по таблице используем функции ПОИСК и ИНДЕКС.

|t расч |> t кр(4,65>1,86). Следовательно, связь между оценкой стажа и оценкой времени решения является статистически значимой при 5%-ном уровне значимости. Также можно сказать, что связь достаточно сильная и обратная по направлению (видно по корреляционному полю).

Построим корреляционное поле (рис.3):

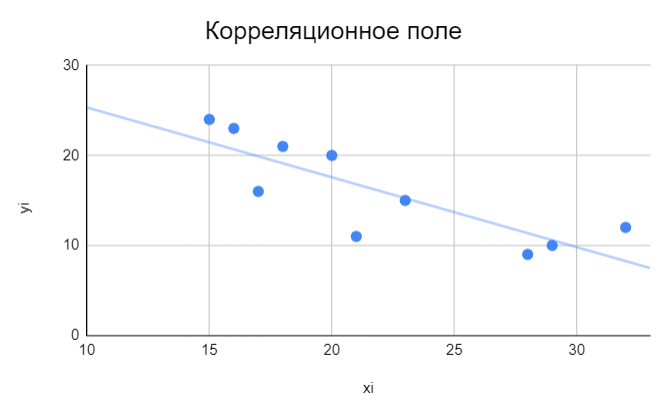


Рис.3. Корреляционное поле

**Ответ:** на уровне значимости 0,05 нулевую гипотезу отвергаем.

*Задача 3.2*

Три арбитра оценили мастерство 10 спортсменов, в итоге были получены три последовательности рангов (в первой строке приведены ранги арбитра А, во второй – ранги арбитра В, в третьей – ранги арбитра С):

| xi | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| yi | 3 | 10 | 7 | 2 | 8 | 5 | 6 | 9 | 1 | 4 |
| zi | 6 | 2 | 1 | 3 | 9 | 4 | 5 | 7 | 10 | 8 |

Определить пару арбитров, оценки которых наиболее согласуются, используя коэффициент ранговой корреляции Спирмена. Построить корреляционное поле.

*Решение:*

Для определения коэффициента корреляции Спирмена найдем разности рангов, их квадраты и сумму для оценок каждой из пар арбитров.

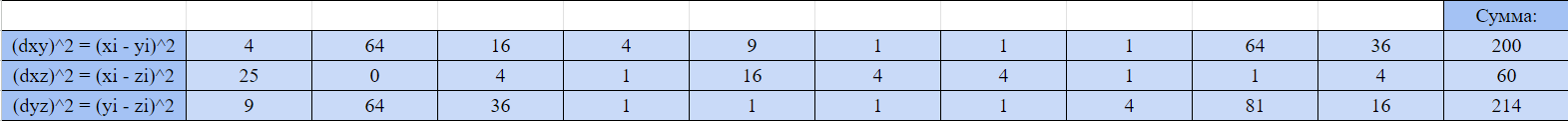


Табл. 7. Вспомогательная таблица с промежуточными вычислениями

Также для каждой пары арбитров вычислим коэффициент Спирмена и определим t-критерий Стьюдента tвыч по формулам (2) и (3). Далее по таблице “Критические точки t-распределения Стьюдента” определим критическое значение статистики Стьюдента, где k = n - 2 = 10 - 2 = 8, уровень значимости α = 0.05 для односторонней (правосторонней) критической области:

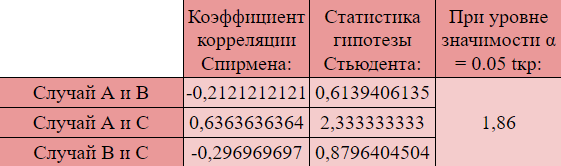


Табл. 8. Таблица с итоговыми результатами вычислений.

Для автоматического поиска критического значения по таблице используем функции ПОИСК и ИНДЕКС.

Построим поле корреляции для графического изображения зависимостей каждой пары арбитров (рис.4-рис.6):

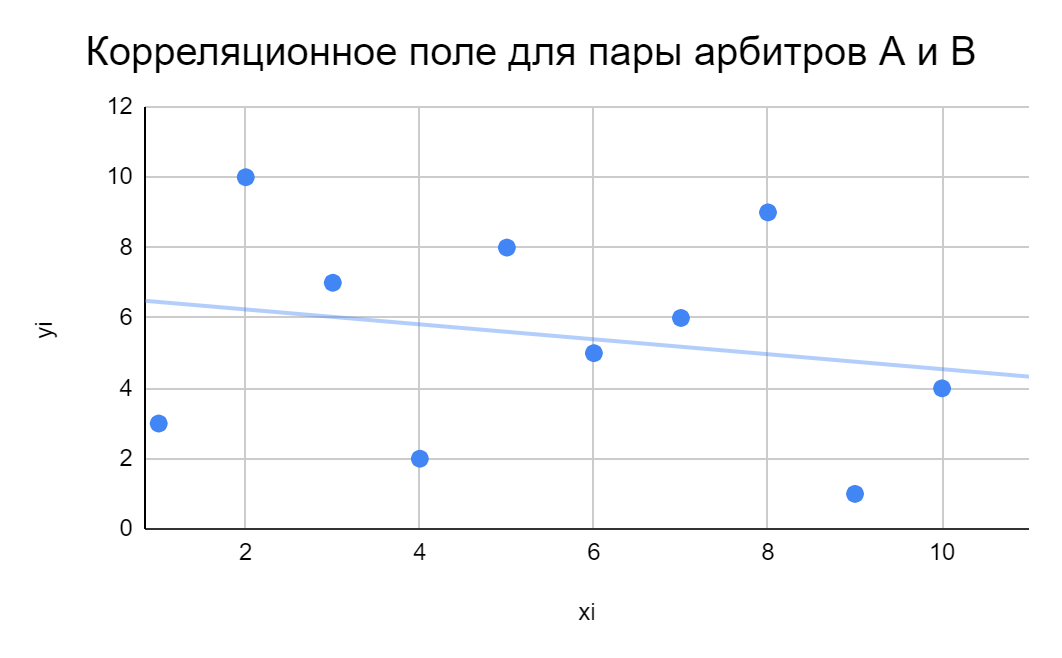


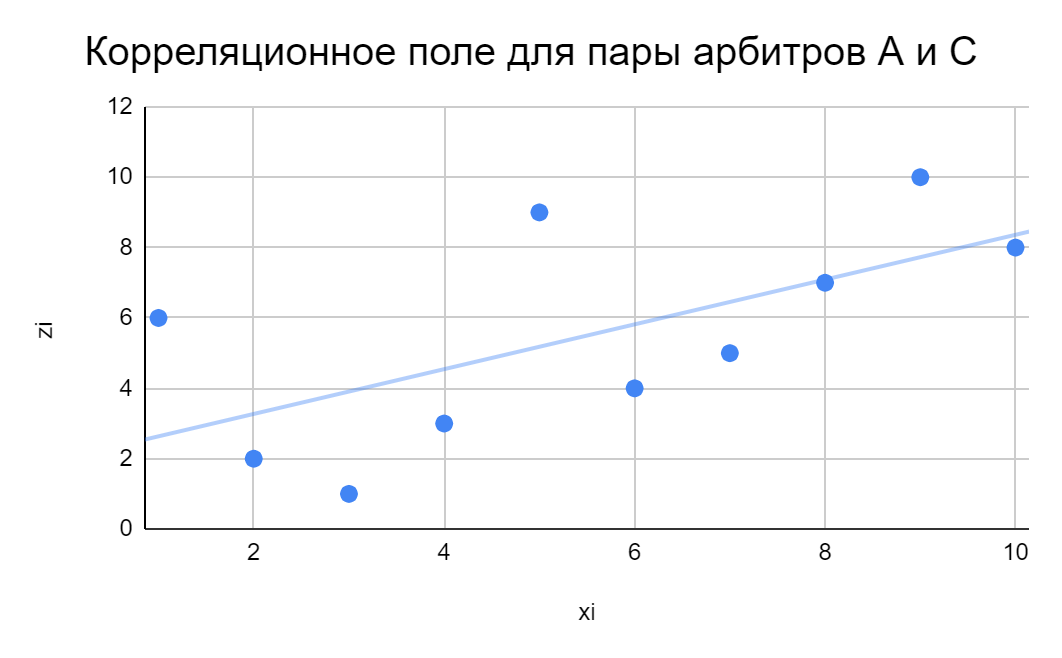
Рис.4. Корреляционное поле

Рис.5. Корреляционное поле

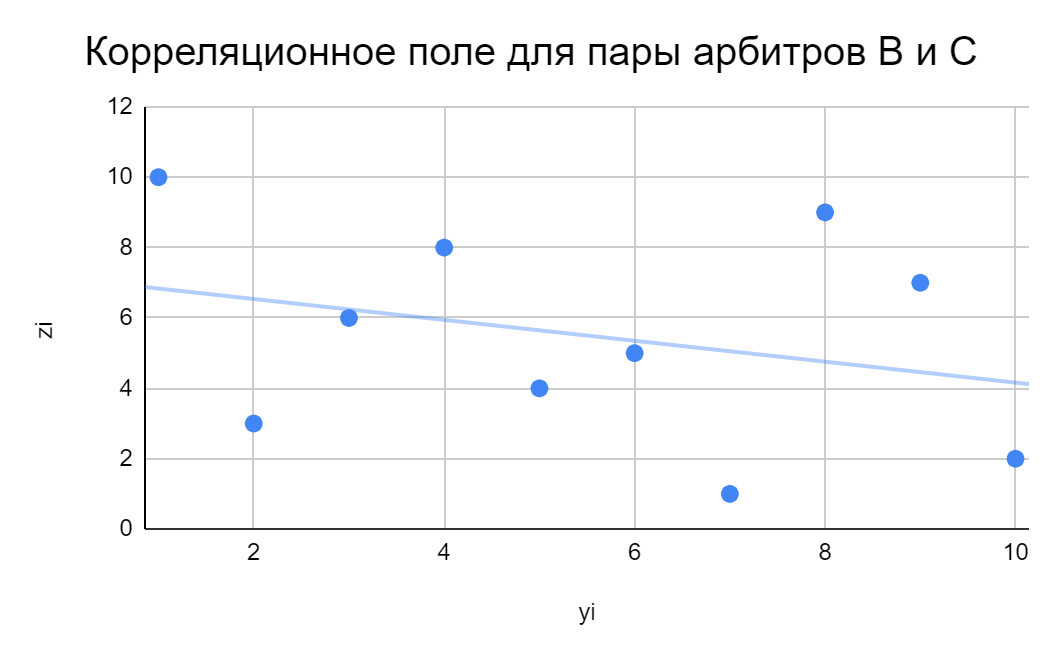


Рис.6. Корреляционное поле

Сравнивая значения с критическим, видим, что наиболее согласуются оценки арбитров А и С. Данный коэффициент ранговой корреляции наибольший по величине.

**Ответ:** Пара арбитров, оценки которых наиболее согласуются, это пара A и C.

*Задание 4:*

*Задача 4.1*

Построить корреляционное поле. Сделать предположение о форме и направлении взаимосвязи двух исследуемых показателей. Найти значение коэффициентов линейной корреляции. Необходимо определить взаимосвязь характеристик: агрессивности и IQ у школьников по полученным данным тестирования.

| № п/п | Данные по агрессивности X\_agr | Данные по IQ Y\_iq |
| --- | --- | --- |
| 1 | 24 | 100 |
| 2 | 27 | 115 |
| 3 | 26 | 117 |
| 4 | 21 | 119 |
| 5 | 20 | 134 |
| 6 | 31 | 94 |
| 7 | 26 | 105 |
| 8 | 22 | 103 |
| 9 | 20 | 111 |
| 10 | 18 | 124 |
| 11 | 30 | 122 |
| 12 | 29 | 109 |
| 13 | 24 | 110 |
| 14 | 26 | 86 |

*Решение:*

Сформулируем статистические гипотезы:

1. Нулевая гипотеза – взаимосвязь между агрессивностью и IQ у школьников существует.
2. Конкурирующая гипотеза – взаимосвязь между агрессивностью и IQ у школьников отсутствует.

Исходя из исходных данных сделаем вспомогательную таблицу:

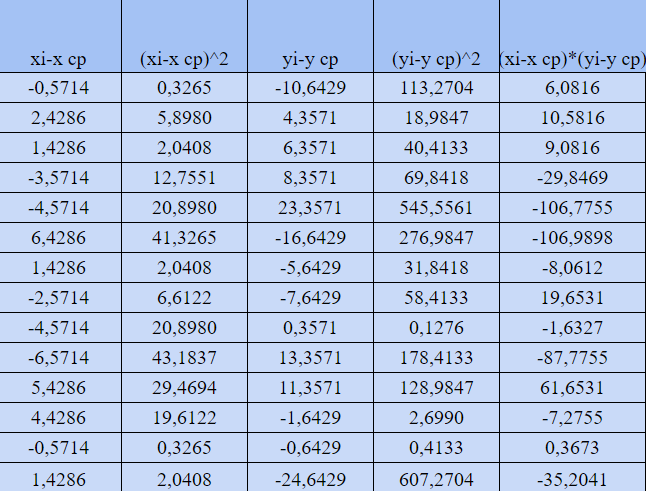


Табл.9. Вспомогательная таблица с промежуточными вычислениями

Найдем значение коэффициента корреляции по формуле (1):



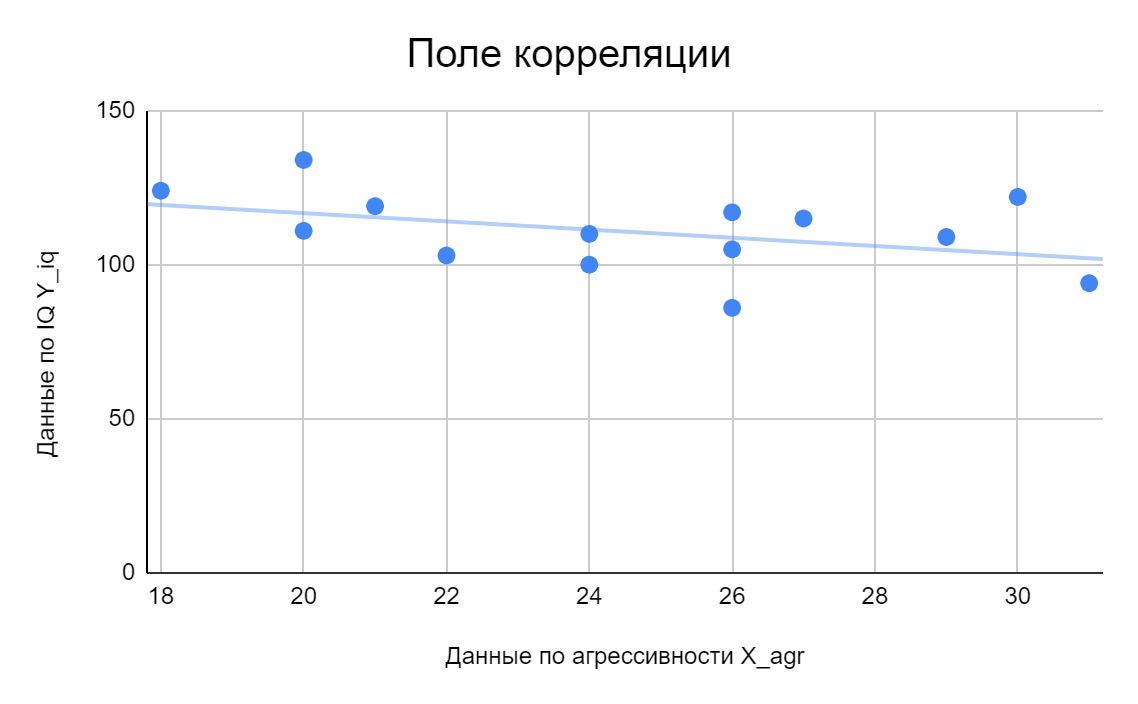
Далее по таблице “Критические значения корреляции rzy Пирсона” определим критические значения для полученного коэффициента корреляции, где k = n - 2 = 14 - 2 = 12, P = 0,05 (уровень значимости):



Для автоматического поиска критического значения по таблице используем функции ПОИСК и ИНДЕКС.

= 0,53 > |-0,42| , следовательно, гипотеза Н1 отвергается и принимается гипотеза H0.

Построим корреляционное поле (рис.7):

Рис.7. Корреляционное поле

**Ответ**: на уровне значимости 0,05 нулевую гипотезу принимаем.

*Задача 4.2*

На основании наблюдений за развивающимся сайтом и изменением его средневзвешенной позиции по основным запросам в поисковой системе необходимо проверить, можно ли говорить о линейной зависимости между позицией сайта и числом посетителей. Построить корреляционное поле.

Исходные данные:

X - число посетителей в сутки;

Y – усредненная позиция сайта в поисковой системе.

В таблице даны значения признаков

| № п/п | Число посетителей в сутки, X | Усредненная позиция сайта в поисковой системе, Y |
| --- | --- | --- |
| 1 | 500 | 5,4 |
| 2 | 790 | 4,2 |
| 3 | 870 | 4 |
| 4 | 1500 | 3,4 |
| 5 | 2300 | 2,5 |
| 6 | 5600 | 1 |
| 7 | 100 | 6,1 |
| 8 | 20 | 8,2 |
| 9 | 5 | 14,6 |

*Решение:*

Сформулируем статистические гипотезы:

1. Нулевая гипотеза – взаимосвязь между позицией сайта и числом посетителей существует.
2. Конкурирующая гипотеза – взаимосвязь между позицией сайта и числом посетителей отсутствует.

Исходя из исходных данных сделаем вспомогательную таблицу:

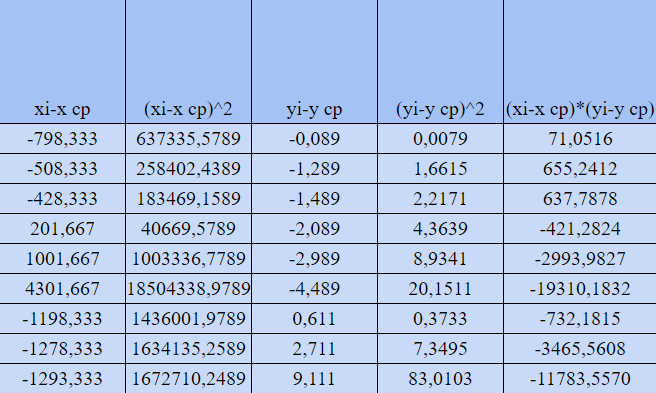


Табл.10. Вспомогательная таблица с промежуточными вычислениями

Найдем значение коэффициента корреляции по формуле (1):



Далее по таблице “Критические значения корреляции rzy Пирсона” определим критические значения для полученного коэффициента корреляции, где k = n - 2 = 9 - 2 = 7, P = 0,05 (уровень значимости):



Для автоматического поиска критического значения по таблице используем функции ПОИСК и ИНДЕКС.

= 0,67 > |-0,65| , следовательно, гипотеза Н1 отвергается и принимается гипотеза H0.

Построим корреляционное поле (рис.8):

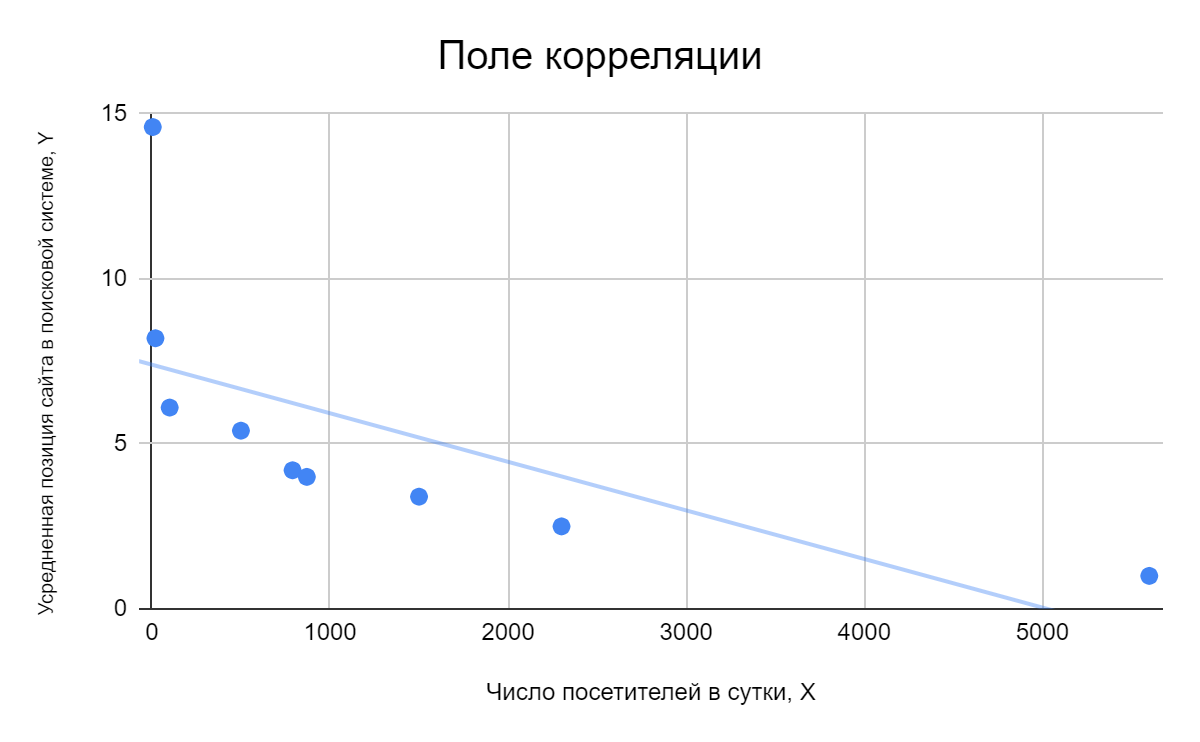


Рис.8. Корреляционное поле

**Ответ**: на уровне значимости 0,05 нулевую гипотезу принимаем.

## **Лабораторная работа №7. Часть 1**

*Выполнил Иванов Н. Р.*

**Цель работы:** Вычисление коэффициентов корреляции.

**Оборудования:** ПК, Microsoft Excel.

*Использованные формулы:*

1. Коэффициент корреляции — параметр, который характеризует степень линейной взаимосвязи между двумя выборками, рассчитывается по формуле:

1. Коэффициент ранговой корреляции Спирмена:

,

где - квадраты разности рангов,

n - число наблюдений.

1. Статистика гипотезы Стьюдента:

*Задание 2:*

Построить корреляционное поле. Сделать предположение о форме и направлении взаимосвязи двух исследуемых показателей. Вычислить коэффициент линейной корреляции Пирсона и коэффициент ранговой корреляции Спирмена.

*Пример 1:*

10 школьникам были даны тесты на наглядно-образное и вербальное мышление. Измерялось среднее время решения заданий теста в секундах. Исследователя интересует вопрос: существует ли взаимосвязь между временем решения этих задач? Переменная X — обозначает среднее время решения наглядно-образных, а переменная Y— среднее время решения вербальных заданий тестов.

| № испытуемых | X | Y |
| --- | --- | --- |
| 1 | 19 | 17 |
| 2 | 32 | 7 |
| 3 | 33 | 17 |
| 4 | 44 | 28 |
| 5 | 28 | 27 |
| 6 | 35 | 31 |
| 7 | 39 | 20 |
| 8 | 39 | 17 |
| 9 | 44 | 35 |
| 10 | 44 | 43 |

*Решение:*

Исходя из исходных данных сделаем вспомогательную таблицу:

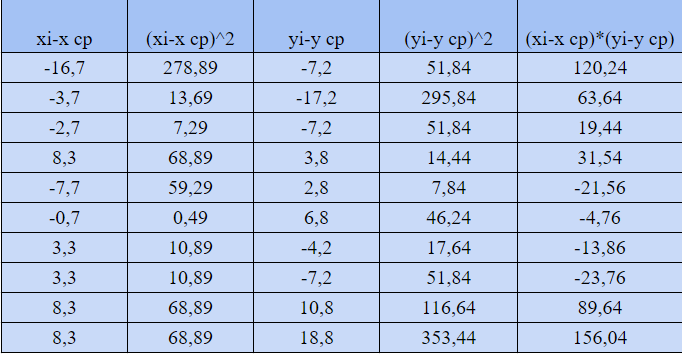


Табл.1. Вспомогательная таблица с промежуточными вычислениями

1. Нулевая гипотеза – взаимосвязь между временем решения существует.
2. Конкурирующая гипотеза – взаимосвязь между временем решения отсутствует.

Найдем значение коэффициента корреляции по формуле (1):



Далее по таблице “Критические значения корреляции rzy Пирсона” определим критические значения для полученного коэффициента корреляции, где k = n - 2 = 10 - 2 = 8, P = 0,05 (уровень значимости):



Для автоматического поиска критического значения по таблице используем функции ПОИСК и ИНДЕКС.

= 0,63 > 0,54 , следовательно, гипотеза Н1 отвергается и принимается гипотеза H0.

Построим корреляционное поле (рис.1):

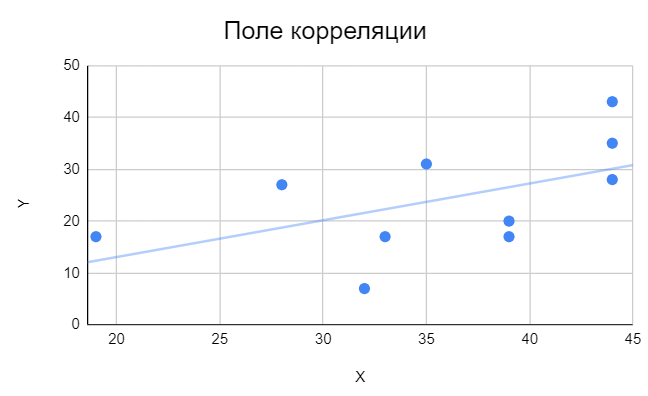


Рис.1. Корреляционное поле

**Ответ:** на уровне значимости 0,05 нулевую гипотезу Н0 принимаем.

*Пример 2:*

Преподавателю и студенту было предложено расположить 10 профессий в порядке их общественной значимости. Ответы перечислены в таблице 1:

***Таблица 1.***

| Оценка преподавателя, xi | Профессии | Оценка студента, yi |
| --- | --- | --- |
| 3 | профессор | 2 |
| 1 | врач | 1 |
| 4 | учитель школы | 7 |
| 2 | директор магазина | 4 |
| 8 | бухгалтер | 5 |
| 6 | банкир | 3 |
| 9 | водитель | 9 |
| 5 | журналист | 8 |
| 10 | ди-джей | 10 |

Какова корреляция рангов между двумя рядами оценок? Одинаково ли мнение преподавателя и студента по этому вопросу?

*Решение:*

Исходя из исходных данных составим вспомогательную таблицу разности рангов, их квадраты и суммы:

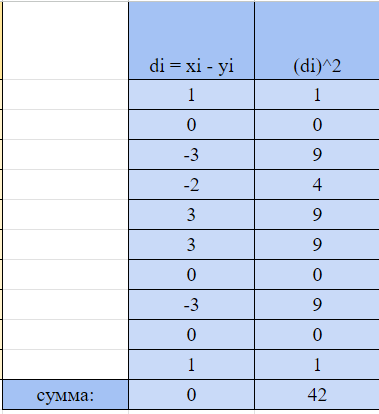


Табл. 2. Вспомогательная таблица с промежуточными вычислениями

Найдем коэффициент ранговой корреляции Спирмена по формуле (2):



Далее необходимо проверить, существует ли положительная корреляционная связь между мнениями преподавателя и студента. Для этого (при n ≥ 10) используем формулу (3) с k = (n – 2) степенями свободы:



1. Нулевая гипотеза – коэффициент корреляции не является статистически значимым ().
2. Конкурирующая гипотеза – существует положительная корреляционная зависимость ().

Далее по таблице “Критические точки t-распределения Стьюдента” определим критическое значение статистики Стьюдента, где k = n - 2 = 10 - 2 = 8, уровень значимости α = 0.05 для односторонней (правосторонней) критической области:



Для автоматического поиска критического значения по таблице используем функции ПОИСК и ИНДЕКС.

t расч > t кр(3,16>1,86). Следовательно, связь между мнениями преподавателя и студента является статистически значимой при 5%-ном уровне значимости.

Построим корреляционное поле (рис.2):

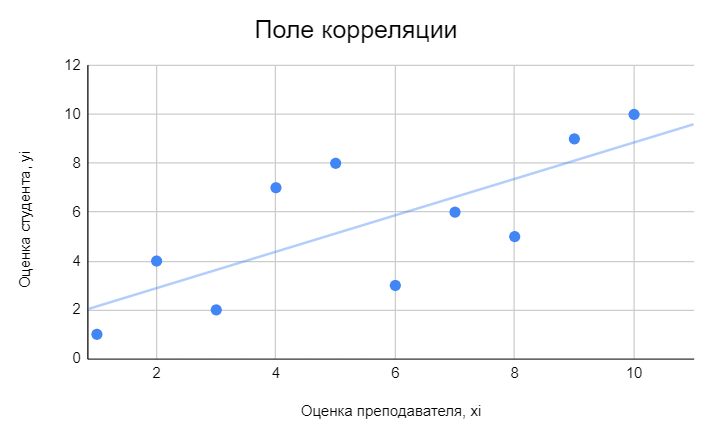


Рис.2. Корреляционное поле

**Ответ:** на уровне значимости 0,05 нулевую гипотезу отвергаем.

*Задание 3:*

Построить корреляционное поле. Сделать предположение о форме и направлении взаимосвязи двух исследуемых показателей. Найти значения коэффициентов ранговой корреляции Спирмена.

*Задача 3.1*

С помощью коэффициента ранговой корреляции установить зависимость между стажем практической работы и временем решения контрольной задачи у 10 программистов на основе следующих данных:

| Номера испытуемых | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Стаж (в мес.) | 32 | 15 | 16 | 18 | 20 | 28 | 21 | 29 | 23 | 17 |
| Время решения (в мин.) | 12 | 24 | 23 | 21 | 20 | 9 | 11 | 10 | 15 | 16 |

*Решение:*

Сформулируем статистические гипотезы:

1. Нулевая гипотеза – коэффициент корреляции не является статистически значимым ().
2. Конкурирующая гипотеза – существует положительная корреляционная зависимость ().

Определим ранги оценкам стажа xi, располагая оценки в возрастающем порядке и сохраняя связь с оценками времени решения. Ранги пронумеруем от 1 до 10.

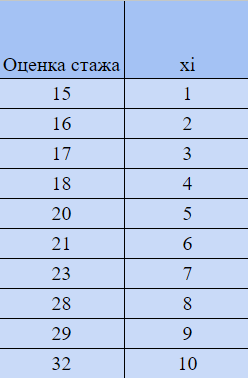


Табл. 3. Вспомогательная таблица с промежуточными вычислениями

Аналогично определим ранги оценкам времени решения yi. И также пронумеруем их от 1 до 10, располагая по возрастанию.

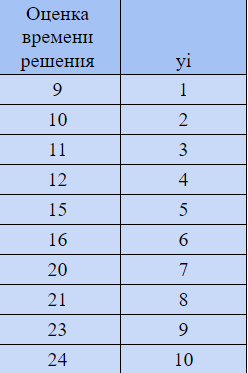


Табл. 4. Вспомогательная таблица с промежуточными вычислениями

Для пояснения процесса вычисления рангов yi возьмем ранг y1. Мы рассматриваем оценку времени решения программиста, который по оценке стажа занимает 1 место (xi = 1; 15 мес.) Теперь найдем оценку времени решения этого программиста (24 мин.) и видим, что это время решения получило ранг yi=10. В сводную таблицу заносим xi = 1, yi=10. Проделываем аналогичные действия для остальных значений и получаем сводную таблицу:

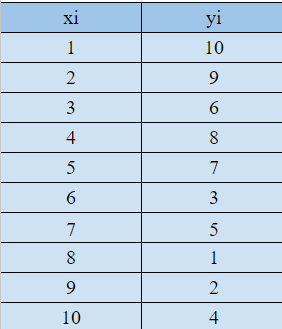


Табл. 5. Сводная таблица результатов

Исходя из полученных данных составим вспомогательную таблицу разности рангов, их квадраты и суммы:

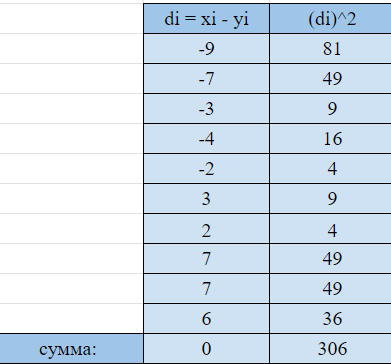


Табл. 6. Вспомогательная таблица с промежуточными вычислениями

Найдем коэффициент ранговой корреляции Спирмена по формуле (2):



Далее необходимо проверить, существует ли положительная корреляционная связь между оценкой стажа и оценкой времени решения задач. Для этого (при n ≥ 10) используем формулу (3) с k = (n – 2) степенями свободы:



Далее по таблице “Критические точки t-распределения Стьюдента” определим критическое значение статистики Стьюдента, где k = n - 2 = 10 - 2 = 8, уровень значимости α = 0.05 для односторонней (правосторонней) критической области:



Для автоматического поиска критического значения по таблице используем функции ПОИСК и ИНДЕКС.

|t расч |> t кр(4,65>1,86). Следовательно, связь между оценкой стажа и оценкой времени решения является статистически значимой при 5%-ном уровне значимости. Также можно сказать, что связь достаточно сильная и обратная по направлению (видно по корреляционному полю).

Построим корреляционное поле (рис.3):

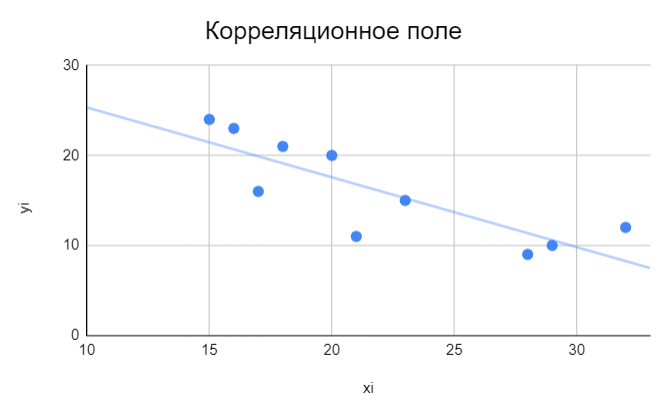


Рис.3. Корреляционное поле

**Ответ:** на уровне значимости 0,05 нулевую гипотезу отвергаем.

*Задача 3.2*

Три арбитра оценили мастерство 10 спортсменов, в итоге были получены три последовательности рангов (в первой строке приведены ранги арбитра А, во второй – ранги арбитра В, в третьей – ранги арбитра С):

| xi | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| yi | 3 | 10 | 7 | 2 | 8 | 5 | 6 | 9 | 1 | 4 |
| zi | 6 | 2 | 1 | 3 | 9 | 4 | 5 | 7 | 10 | 8 |

Определить пару арбитров, оценки которых наиболее согласуются, используя коэффициент ранговой корреляции Спирмена. Построить корреляционное поле.

*Решение:*

Для определения коэффициента корреляции Спирмена найдем разности рангов, их квадраты и сумму для оценок каждой из пар арбитров.

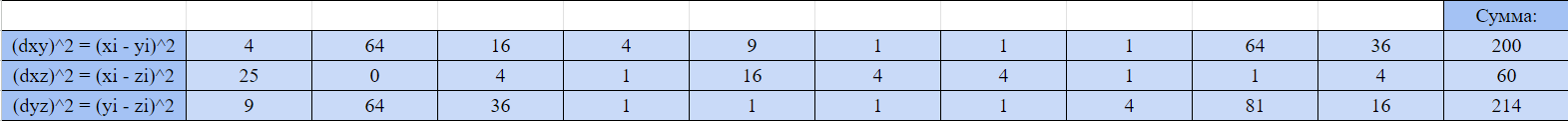


Табл. 7. Вспомогательная таблица с промежуточными вычислениями

Также для каждой пары арбитров вычислим коэффициент Спирмена и определим t-критерий Стьюдента tвыч по формулам (2) и (3). Далее по таблице “Критические точки t-распределения Стьюдента” определим критическое значение статистики Стьюдента, где k = n - 2 = 10 - 2 = 8, уровень значимости α = 0.05 для односторонней (правосторонней) критической области:

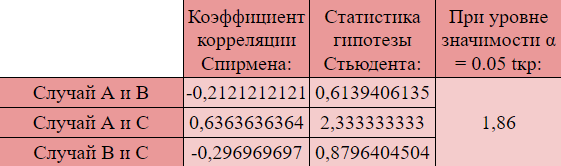


Табл. 8. Таблица с итоговыми результатами вычислений.

Для автоматического поиска критического значения по таблице используем функции ПОИСК и ИНДЕКС.

Построим поле корреляции для графического изображения зависимостей каждой пары арбитров (рис.4-рис.6):

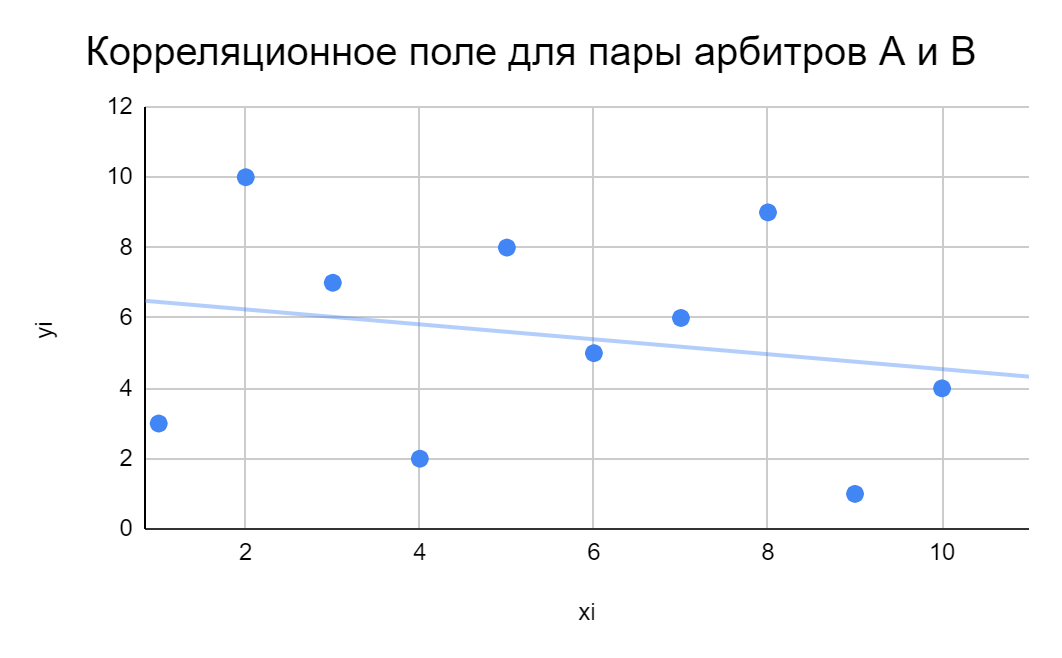


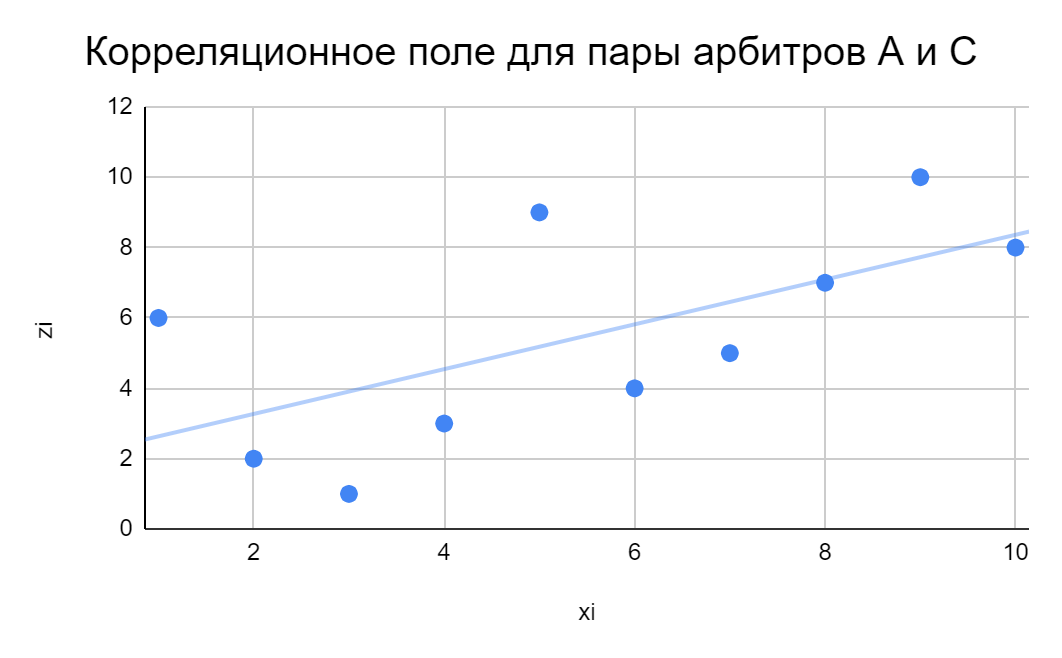
Рис.4. Корреляционное поле

Рис.5. Корреляционное поле

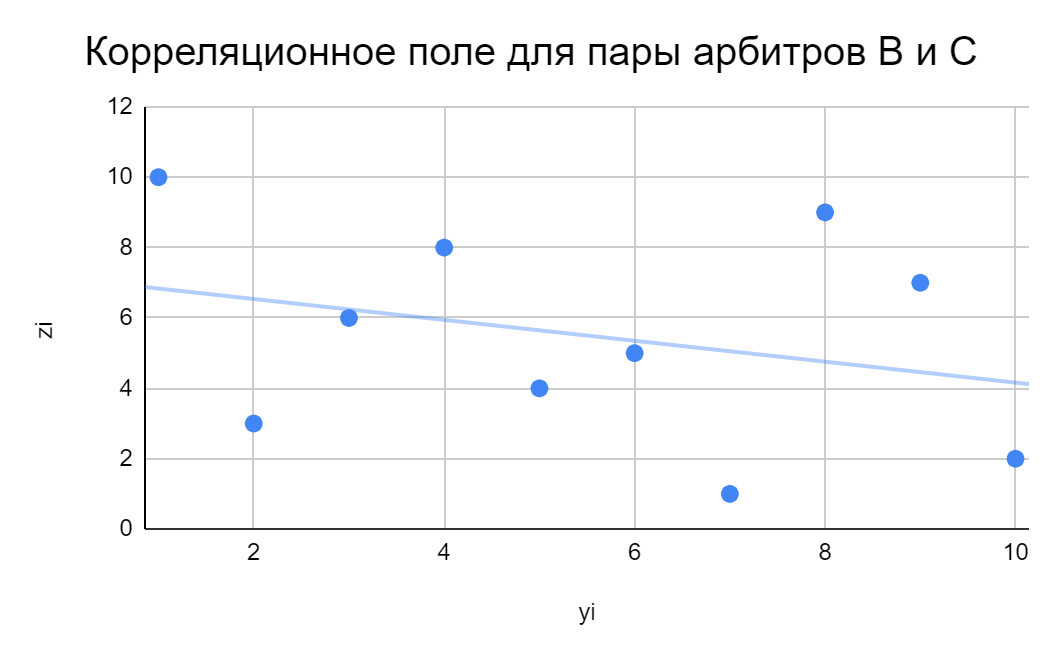


Рис.6. Корреляционное поле

Сравнивая значения с критическим, видим, что наиболее согласуются оценки арбитров А и С. Данный коэффициент ранговой корреляции наибольший по величине.

**Ответ:** Пара арбитров, оценки которых наиболее согласуются, это пара A и C.

*Задание 4:*

*Задача 4.1*

Построить корреляционное поле. Сделать предположение о форме и направлении взаимосвязи двух исследуемых показателей. Найти значение коэффициентов линейной корреляции. Необходимо определить взаимосвязь характеристик: агрессивности и IQ у школьников по полученным данным тестирования.

| № п/п | Данные по агрессивности X\_agr | Данные по IQ Y\_iq |
| --- | --- | --- |
| 1 | 24 | 100 |
| 2 | 27 | 115 |
| 3 | 26 | 117 |
| 4 | 21 | 119 |
| 5 | 20 | 134 |
| 6 | 31 | 94 |
| 7 | 26 | 105 |
| 8 | 22 | 103 |
| 9 | 20 | 111 |
| 10 | 18 | 124 |
| 11 | 30 | 122 |
| 12 | 29 | 109 |
| 13 | 24 | 110 |
| 14 | 26 | 86 |

*Решение:*

Сформулируем статистические гипотезы:

1. Нулевая гипотеза – взаимосвязь между агрессивностью и IQ у школьников существует.
2. Конкурирующая гипотеза – взаимосвязь между агрессивностью и IQ у школьников отсутствует.

Исходя из исходных данных сделаем вспомогательную таблицу:

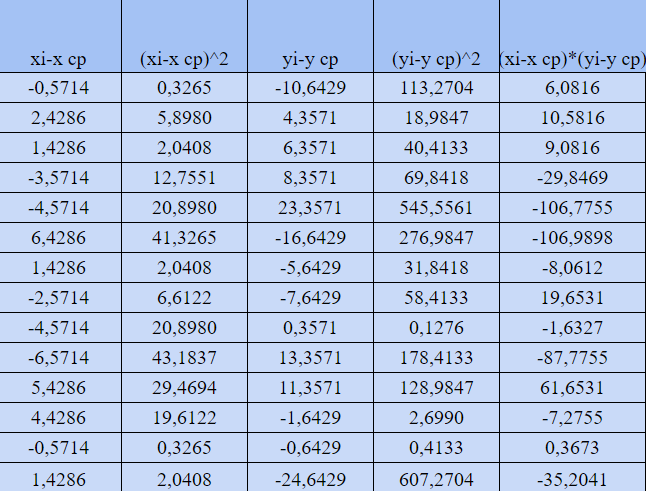


Табл.9. Вспомогательная таблица с промежуточными вычислениями

Найдем значение коэффициента корреляции по формуле (1):



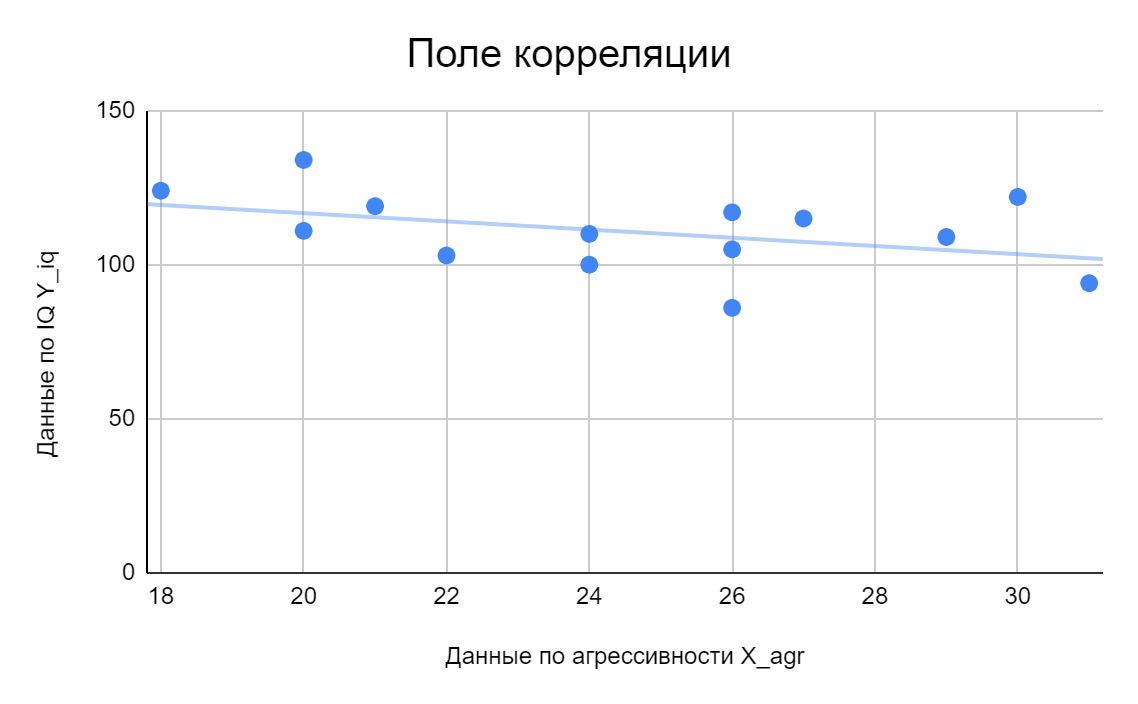
Далее по таблице “Критические значения корреляции rzy Пирсона” определим критические значения для полученного коэффициента корреляции, где k = n - 2 = 14 - 2 = 12, P = 0,05 (уровень значимости):



Для автоматического поиска критического значения по таблице используем функции ПОИСК и ИНДЕКС.

= 0,53 > |-0,42| , следовательно, гипотеза Н1 отвергается и принимается гипотеза H0.

Построим корреляционное поле (рис.7):

Рис.7. Корреляционное поле

**Ответ**: на уровне значимости 0,05 нулевую гипотезу принимаем.

*Задача 4.2*

На основании наблюдений за развивающимся сайтом и изменением его средневзвешенной позиции по основным запросам в поисковой системе необходимо проверить, можно ли говорить о линейной зависимости между позицией сайта и числом посетителей. Построить корреляционное поле.

Исходные данные:

X - число посетителей в сутки;

Y – усредненная позиция сайта в поисковой системе.

В таблице даны значения признаков

| № п/п | Число посетителей в сутки, X | Усредненная позиция сайта в поисковой системе, Y |
| --- | --- | --- |
| 1 | 500 | 5,4 |
| 2 | 790 | 4,2 |
| 3 | 870 | 4 |
| 4 | 1500 | 3,4 |
| 5 | 2300 | 2,5 |
| 6 | 5600 | 1 |
| 7 | 100 | 6,1 |
| 8 | 20 | 8,2 |
| 9 | 5 | 14,6 |

*Решение:*

Сформулируем статистические гипотезы:

1. Нулевая гипотеза – взаимосвязь между позицией сайта и числом посетителей существует.
2. Конкурирующая гипотеза – взаимосвязь между позицией сайта и числом посетителей отсутствует.

Исходя из исходных данных сделаем вспомогательную таблицу:

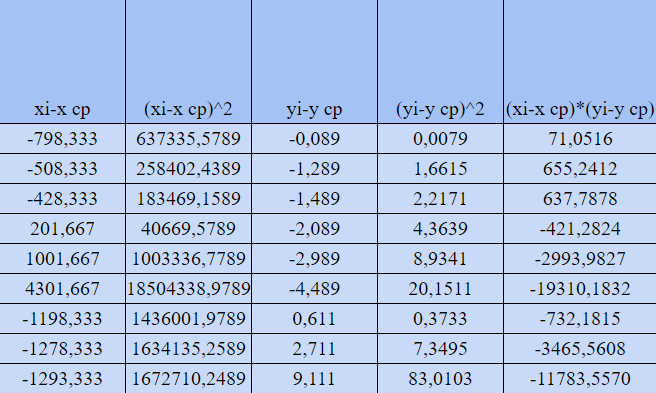


Табл.10. Вспомогательная таблица с промежуточными вычислениями

Найдем значение коэффициента корреляции по формуле (1):



Далее по таблице “Критические значения корреляции rzy Пирсона” определим критические значения для полученного коэффициента корреляции, где k = n - 2 = 9 - 2 = 7, P = 0,05 (уровень значимости):



Для автоматического поиска критического значения по таблице используем функции ПОИСК и ИНДЕКС.

= 0,67 > |-0,65| , следовательно, гипотеза Н1 отвергается и принимается гипотеза H0.

Построим корреляционное поле (рис.8):

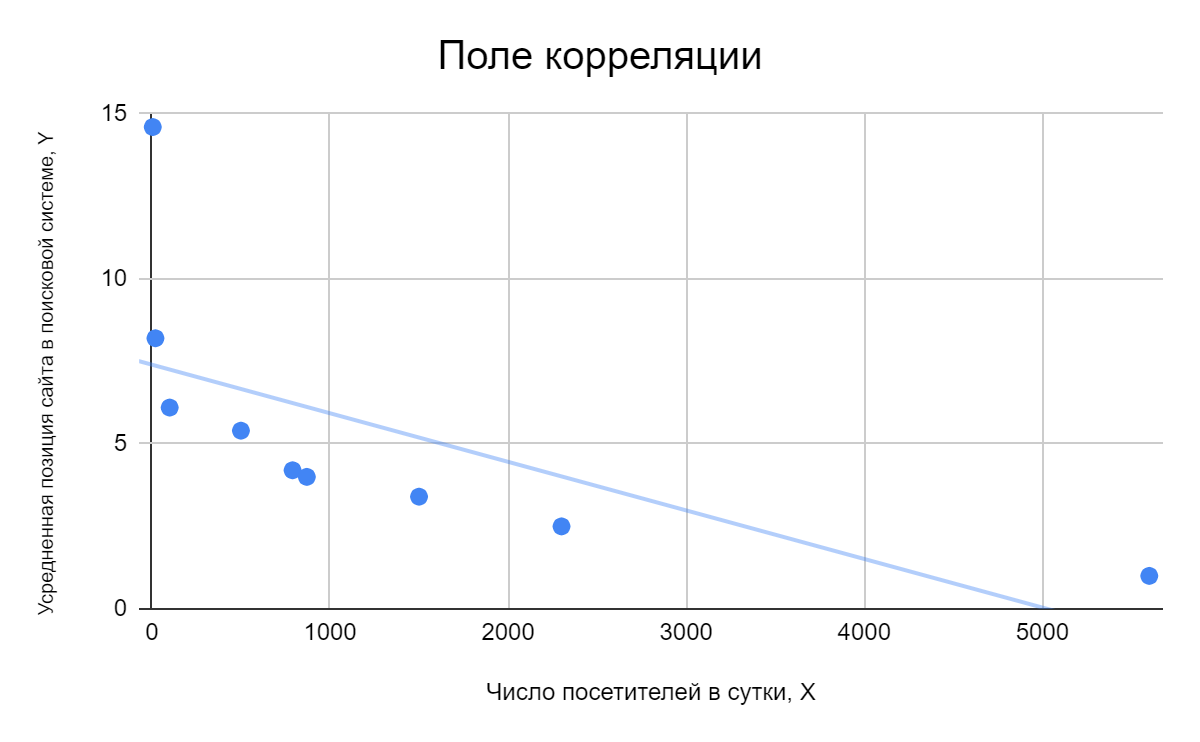


Рис.8. Корреляционное поле

**Ответ**: на уровне значимости 0,05 нулевую гипотезу принимаем.

## **Лабораторная работа №7. Часть 1**

*Выполнил Рожков М. В.*

**Цель работы:** Вычисление коэффициентов корреляции.

**Оборудования:** ПК, Microsoft Excel.

*Использованные формулы:*

1. Коэффициент корреляции — параметр, который характеризует степень линейной взаимосвязи между двумя выборками, рассчитывается по формуле:

1. Коэффициент ранговой корреляции Спирмена:

,

где - квадраты разности рангов,

n - число наблюдений.

1. Статистика гипотезы Стьюдента:

*Задание 2:*

Построить корреляционное поле. Сделать предположение о форме и направлении взаимосвязи двух исследуемых показателей. Вычислить коэффициент линейной корреляции Пирсона и коэффициент ранговой корреляции Спирмена.

*Пример 1:*

10 школьникам были даны тесты на наглядно-образное и вербальное мышление. Измерялось среднее время решения заданий теста в секундах. Исследователя интересует вопрос: существует ли взаимосвязь между временем решения этих задач? Переменная X — обозначает среднее время решения наглядно-образных, а переменная Y— среднее время решения вербальных заданий тестов.

| № испытуемых | X | Y |
| --- | --- | --- |
| 1 | 19 | 17 |
| 2 | 32 | 7 |
| 3 | 33 | 17 |
| 4 | 44 | 28 |
| 5 | 28 | 27 |
| 6 | 35 | 31 |
| 7 | 39 | 20 |
| 8 | 39 | 17 |
| 9 | 44 | 35 |
| 10 | 44 | 43 |

*Решение:*

Исходя из исходных данных сделаем вспомогательную таблицу:

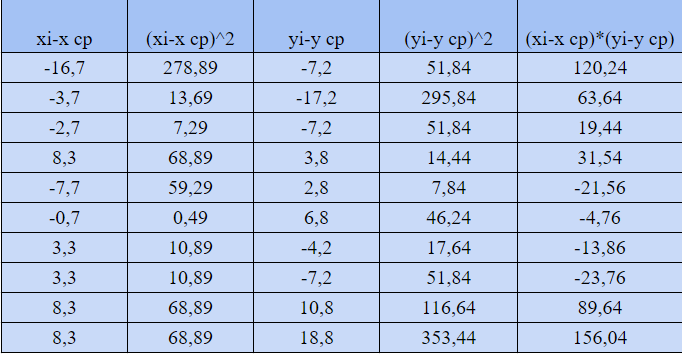


Табл.1. Вспомогательная таблица с промежуточными вычислениями

1. Нулевая гипотеза – взаимосвязь между временем решения существует.
2. Конкурирующая гипотеза – взаимосвязь между временем решения отсутствует.

Найдем значение коэффициента корреляции по формуле (1):



Далее по таблице “Критические значения корреляции rzy Пирсона” определим критические значения для полученного коэффициента корреляции, где k = n - 2 = 10 - 2 = 8, P = 0,05 (уровень значимости):



Для автоматического поиска критического значения по таблице используем функции ПОИСК и ИНДЕКС.

= 0,63 > 0,54 , следовательно, гипотеза Н1 отвергается и принимается гипотеза H0.

Построим корреляционное поле (рис.1):

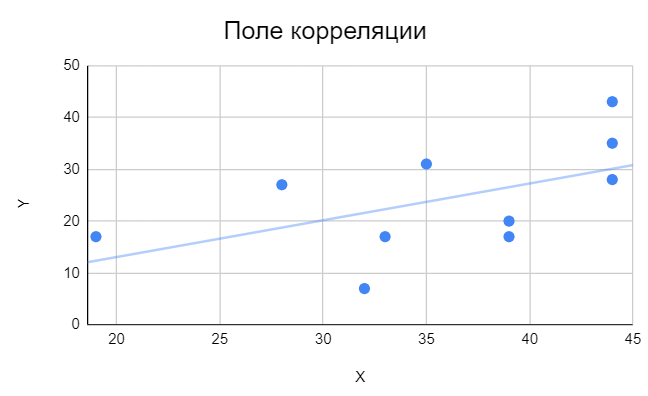


Рис.1. Корреляционное поле

**Ответ:** на уровне значимости 0,05 нулевую гипотезу Н0 принимаем.

*Пример 2:*

Преподавателю и студенту было предложено расположить 10 профессий в порядке их общественной значимости. Ответы перечислены в таблице 1:

***Таблица 1.***

| Оценка преподавателя, xi | Профессии | Оценка студента, yi |
| --- | --- | --- |
| 3 | профессор | 2 |
| 1 | врач | 1 |
| 4 | учитель школы | 7 |
| 2 | директор магазина | 4 |
| 8 | бухгалтер | 5 |
| 6 | банкир | 3 |
| 9 | водитель | 9 |
| 5 | журналист | 8 |
| 10 | ди-джей | 10 |

Какова корреляция рангов между двумя рядами оценок? Одинаково ли мнение преподавателя и студента по этому вопросу?

*Решение:*

Исходя из исходных данных составим вспомогательную таблицу разности рангов, их квадраты и суммы:

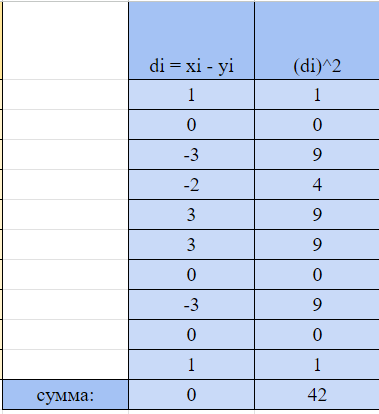


Табл. 2. Вспомогательная таблица с промежуточными вычислениями

Найдем коэффициент ранговой корреляции Спирмена по формуле (2):



Далее необходимо проверить, существует ли положительная корреляционная связь между мнениями преподавателя и студента. Для этого (при n ≥ 10) используем формулу (3) с k = (n – 2) степенями свободы:



1. Нулевая гипотеза – коэффициент корреляции не является статистически значимым ().
2. Конкурирующая гипотеза – существует положительная корреляционная зависимость ().

Далее по таблице “Критические точки t-распределения Стьюдента” определим критическое значение статистики Стьюдента, где k = n - 2 = 10 - 2 = 8, уровень значимости α = 0.05 для односторонней (правосторонней) критической области:



Для автоматического поиска критического значения по таблице используем функции ПОИСК и ИНДЕКС.

t расч > t кр(3,16>1,86). Следовательно, связь между мнениями преподавателя и студента является статистически значимой при 5%-ном уровне значимости.

Построим корреляционное поле (рис.2):

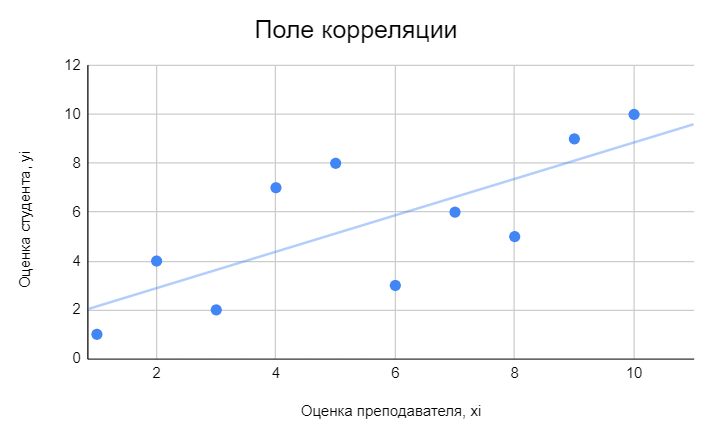


Рис.2. Корреляционное поле

**Ответ:** на уровне значимости 0,05 нулевую гипотезу отвергаем.

*Задание 3:*

Построить корреляционное поле. Сделать предположение о форме и направлении взаимосвязи двух исследуемых показателей. Найти значения коэффициентов ранговой корреляции Спирмена.

*Задача 3.1*

С помощью коэффициента ранговой корреляции установить зависимость между стажем практической работы и временем решения контрольной задачи у 10 программистов на основе следующих данных:

| Номера испытуемых | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Стаж (в мес.) | 32 | 15 | 16 | 18 | 20 | 28 | 21 | 29 | 23 | 17 |
| Время решения (в мин.) | 12 | 24 | 23 | 21 | 20 | 9 | 11 | 10 | 15 | 16 |

*Решение:*

Сформулируем статистические гипотезы:

1. Нулевая гипотеза – коэффициент корреляции не является статистически значимым ().
2. Конкурирующая гипотеза – существует положительная корреляционная зависимость ().

Определим ранги оценкам стажа xi, располагая оценки в возрастающем порядке и сохраняя связь с оценками времени решения. Ранги пронумеруем от 1 до 10.

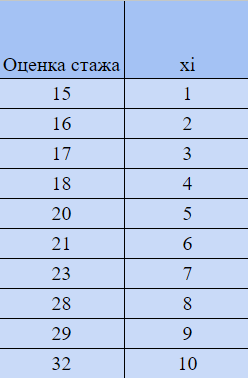


Табл. 3. Вспомогательная таблица с промежуточными вычислениями

Аналогично определим ранги оценкам времени решения yi. И также пронумеруем их от 1 до 10, располагая по возрастанию.

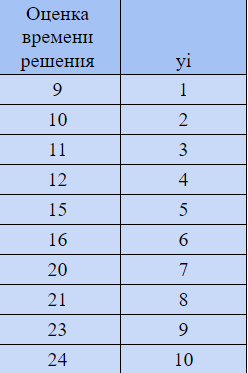


Табл. 4. Вспомогательная таблица с промежуточными вычислениями

Для пояснения процесса вычисления рангов yi возьмем ранг y1. Мы рассматриваем оценку времени решения программиста, который по оценке стажа занимает 1 место (xi = 1; 15 мес.) Теперь найдем оценку времени решения этого программиста (24 мин.) и видим, что это время решения получило ранг yi=10. В сводную таблицу заносим xi = 1, yi=10. Проделываем аналогичные действия для остальных значений и получаем сводную таблицу:

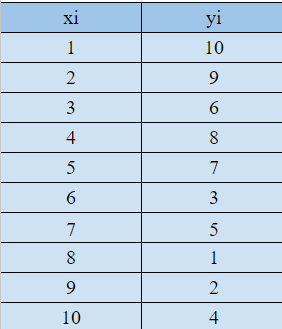


Табл. 5. Сводная таблица результатов

Исходя из полученных данных составим вспомогательную таблицу разности рангов, их квадраты и суммы:

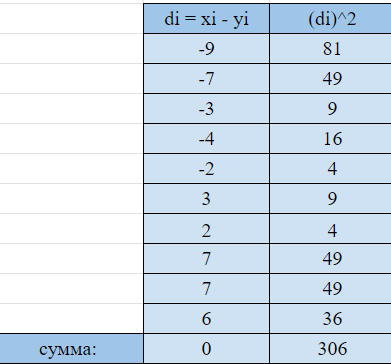


Табл. 6. Вспомогательная таблица с промежуточными вычислениями

Найдем коэффициент ранговой корреляции Спирмена по формуле (2):



Далее необходимо проверить, существует ли положительная корреляционная связь между оценкой стажа и оценкой времени решения задач. Для этого (при n ≥ 10) используем формулу (3) с k = (n – 2) степенями свободы:



Далее по таблице “Критические точки t-распределения Стьюдента” определим критическое значение статистики Стьюдента, где k = n - 2 = 10 - 2 = 8, уровень значимости α = 0.05 для односторонней (правосторонней) критической области:



Для автоматического поиска критического значения по таблице используем функции ПОИСК и ИНДЕКС.

|t расч |> t кр(4,65>1,86). Следовательно, связь между оценкой стажа и оценкой времени решения является статистически значимой при 5%-ном уровне значимости. Также можно сказать, что связь достаточно сильная и обратная по направлению (видно по корреляционному полю).

Построим корреляционное поле (рис.3):

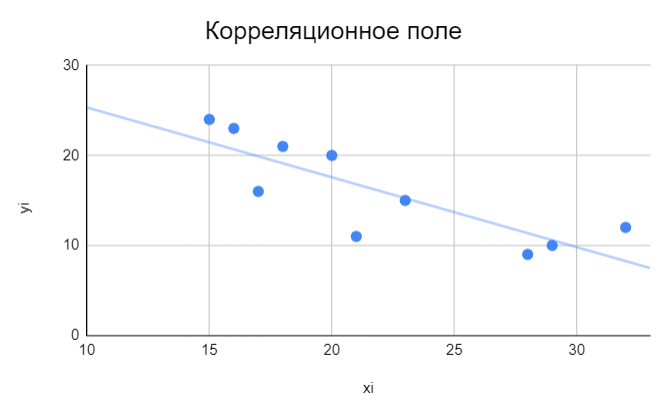


Рис.3. Корреляционное поле

**Ответ:** на уровне значимости 0,05 нулевую гипотезу отвергаем.

*Задача 3.2*

Три арбитра оценили мастерство 10 спортсменов, в итоге были получены три последовательности рангов (в первой строке приведены ранги арбитра А, во второй – ранги арбитра В, в третьей – ранги арбитра С):

| xi | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| yi | 3 | 10 | 7 | 2 | 8 | 5 | 6 | 9 | 1 | 4 |
| zi | 6 | 2 | 1 | 3 | 9 | 4 | 5 | 7 | 10 | 8 |

Определить пару арбитров, оценки которых наиболее согласуются, используя коэффициент ранговой корреляции Спирмена. Построить корреляционное поле.

*Решение:*

Для определения коэффициента корреляции Спирмена найдем разности рангов, их квадраты и сумму для оценок каждой из пар арбитров.

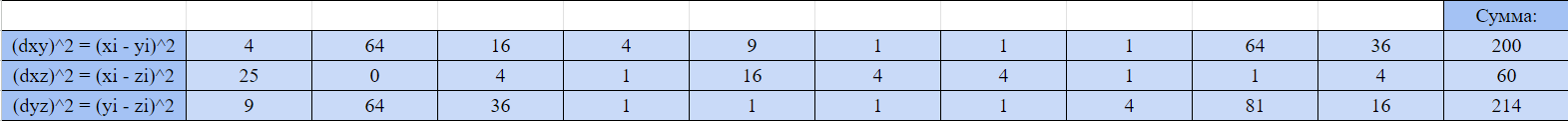


Табл. 7. Вспомогательная таблица с промежуточными вычислениями

Также для каждой пары арбитров вычислим коэффициент Спирмена и определим t-критерий Стьюдента tвыч по формулам (2) и (3). Далее по таблице “Критические точки t-распределения Стьюдента” определим критическое значение статистики Стьюдента, где k = n - 2 = 10 - 2 = 8, уровень значимости α = 0.05 для односторонней (правосторонней) критической области:

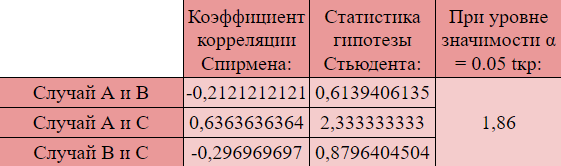


Табл. 8. Таблица с итоговыми результатами вычислений.

Для автоматического поиска критического значения по таблице используем функции ПОИСК и ИНДЕКС.

Построим поле корреляции для графического изображения зависимостей каждой пары арбитров (рис.4-рис.6):

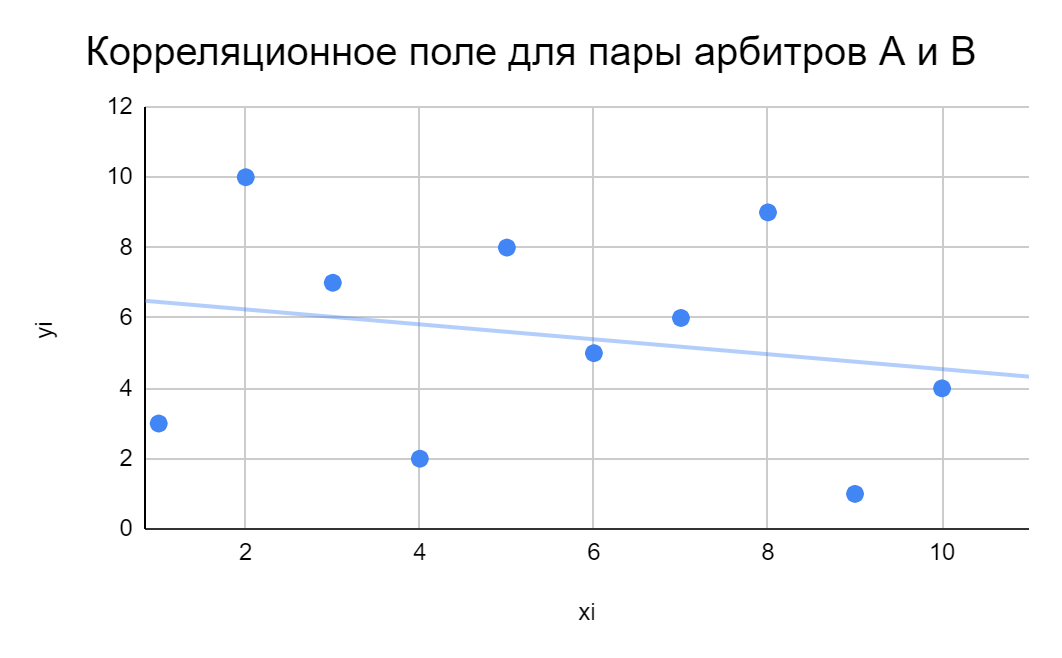


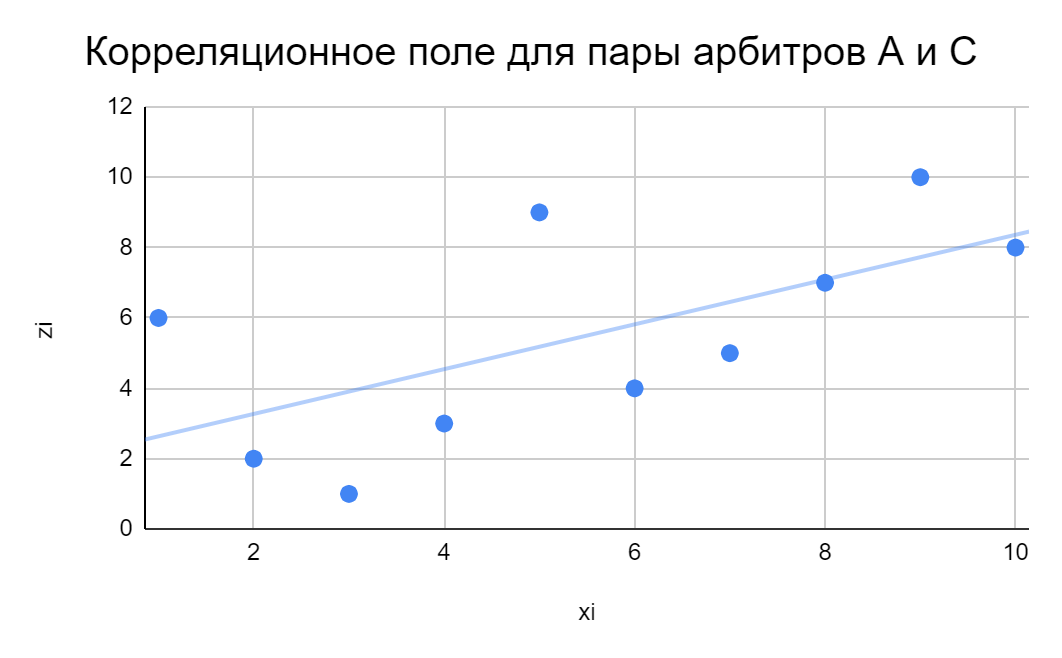
Рис.4. Корреляционное поле

Рис.5. Корреляционное поле

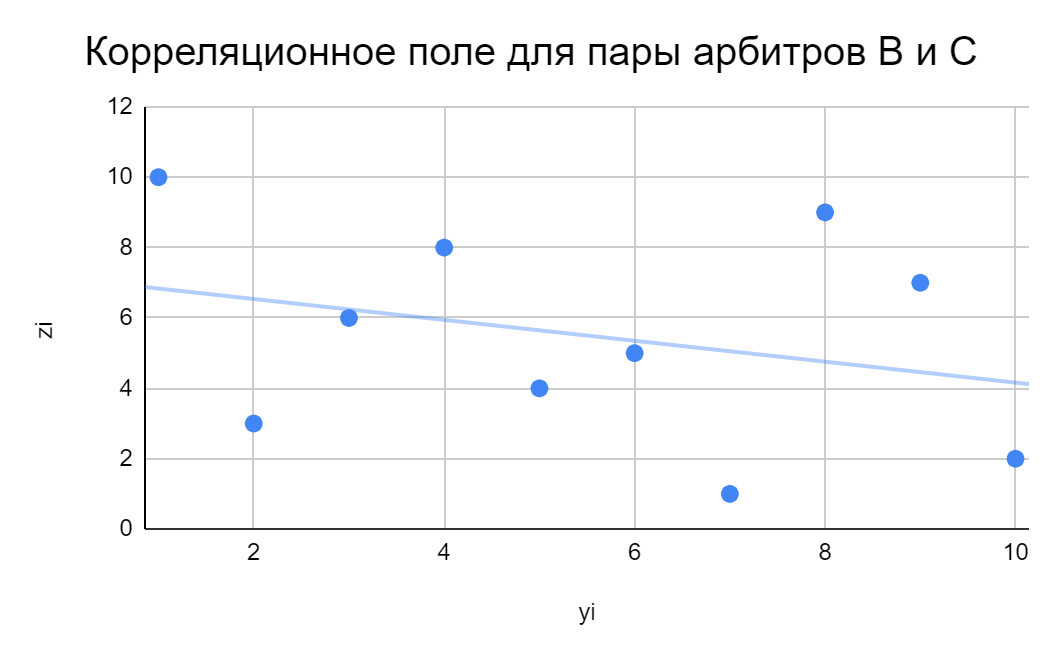


Рис.6. Корреляционное поле

Сравнивая значения с критическим, видим, что наиболее согласуются оценки арбитров А и С. Данный коэффициент ранговой корреляции наибольший по величине.

**Ответ:** Пара арбитров, оценки которых наиболее согласуются, это пара A и C.

*Задание 4:*

*Задача 4.1*

Построить корреляционное поле. Сделать предположение о форме и направлении взаимосвязи двух исследуемых показателей. Найти значение коэффициентов линейной корреляции. Необходимо определить взаимосвязь характеристик: агрессивности и IQ у школьников по полученным данным тестирования.

| № п/п | Данные по агрессивности X\_agr | Данные по IQ Y\_iq |
| --- | --- | --- |
| 1 | 24 | 100 |
| 2 | 27 | 115 |
| 3 | 26 | 117 |
| 4 | 21 | 119 |
| 5 | 20 | 134 |
| 6 | 31 | 94 |
| 7 | 26 | 105 |
| 8 | 22 | 103 |
| 9 | 20 | 111 |
| 10 | 18 | 124 |
| 11 | 30 | 122 |
| 12 | 29 | 109 |
| 13 | 24 | 110 |
| 14 | 26 | 86 |

*Решение:*

Сформулируем статистические гипотезы:

1. Нулевая гипотеза – взаимосвязь между агрессивностью и IQ у школьников существует.
2. Конкурирующая гипотеза – взаимосвязь между агрессивностью и IQ у школьников отсутствует.

Исходя из исходных данных сделаем вспомогательную таблицу:

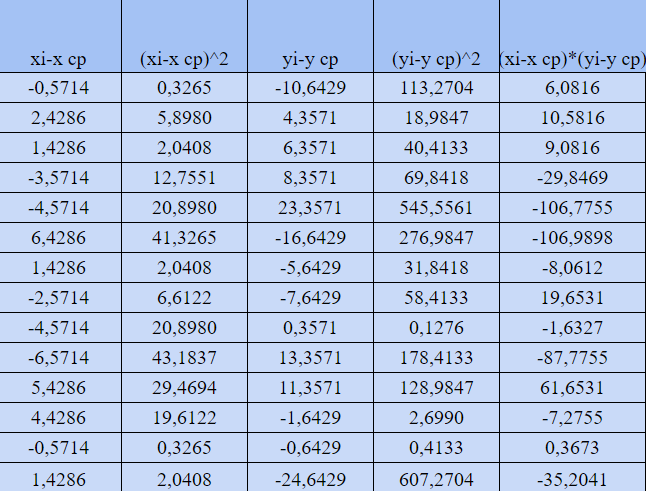


Табл.9. Вспомогательная таблица с промежуточными вычислениями

Найдем значение коэффициента корреляции по формуле (1):



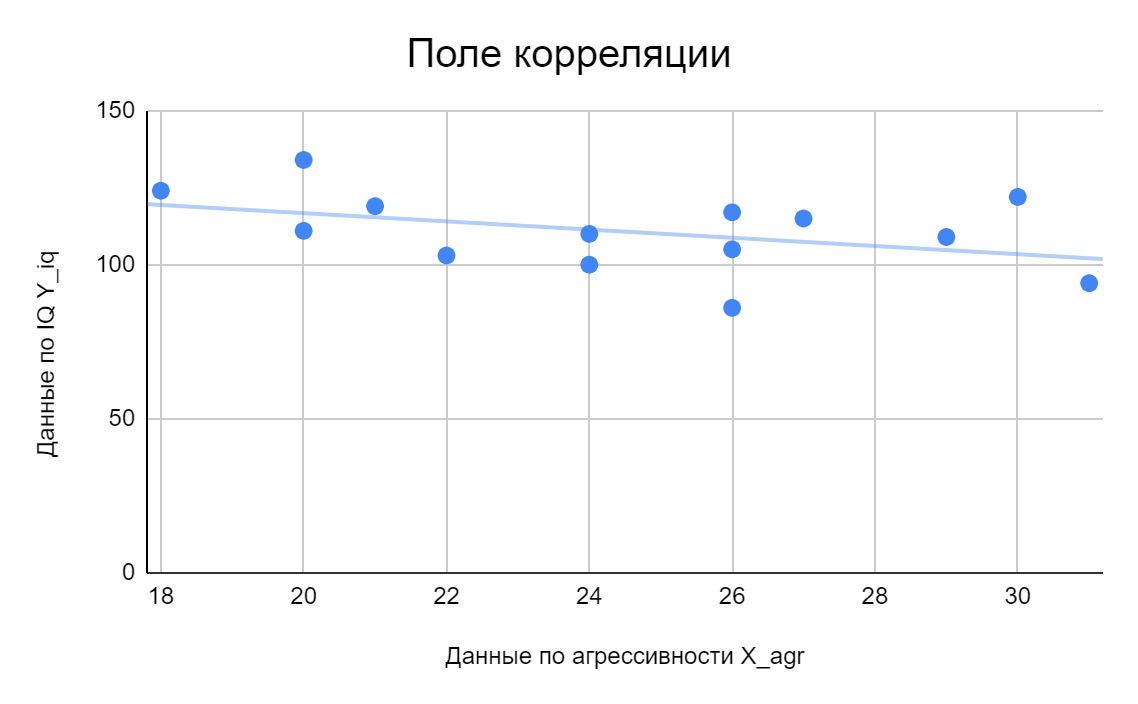
Далее по таблице “Критические значения корреляции rzy Пирсона” определим критические значения для полученного коэффициента корреляции, где k = n - 2 = 14 - 2 = 12, P = 0,05 (уровень значимости):



Для автоматического поиска критического значения по таблице используем функции ПОИСК и ИНДЕКС.

= 0,53 > |-0,42| , следовательно, гипотеза Н1 отвергается и принимается гипотеза H0.

Построим корреляционное поле (рис.7):

Рис.7. Корреляционное поле

**Ответ**: на уровне значимости 0,05 нулевую гипотезу принимаем.

*Задача 4.2*

На основании наблюдений за развивающимся сайтом и изменением его средневзвешенной позиции по основным запросам в поисковой системе необходимо проверить, можно ли говорить о линейной зависимости между позицией сайта и числом посетителей. Построить корреляционное поле.

Исходные данные:

X - число посетителей в сутки;

Y – усредненная позиция сайта в поисковой системе.

В таблице даны значения признаков

| № п/п | Число посетителей в сутки, X | Усредненная позиция сайта в поисковой системе, Y |
| --- | --- | --- |
| 1 | 500 | 5,4 |
| 2 | 790 | 4,2 |
| 3 | 870 | 4 |
| 4 | 1500 | 3,4 |
| 5 | 2300 | 2,5 |
| 6 | 5600 | 1 |
| 7 | 100 | 6,1 |
| 8 | 20 | 8,2 |
| 9 | 5 | 14,6 |

*Решение:*

Сформулируем статистические гипотезы:

1. Нулевая гипотеза – взаимосвязь между позицией сайта и числом посетителей существует.
2. Конкурирующая гипотеза – взаимосвязь между позицией сайта и числом посетителей отсутствует.

Исходя из исходных данных сделаем вспомогательную таблицу:

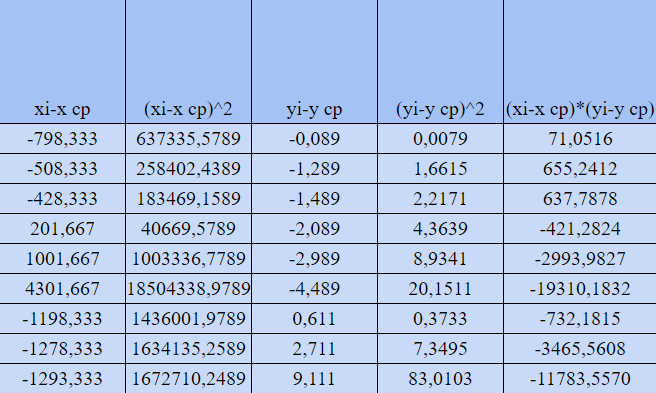


Табл.10. Вспомогательная таблица с промежуточными вычислениями

Найдем значение коэффициента корреляции по формуле (1):



Далее по таблице “Критические значения корреляции rzy Пирсона” определим критические значения для полученного коэффициента корреляции, где k = n - 2 = 9 - 2 = 7, P = 0,05 (уровень значимости):



Для автоматического поиска критического значения по таблице используем функции ПОИСК и ИНДЕКС.

= 0,67 > |-0,65| , следовательно, гипотеза Н1 отвергается и принимается гипотеза H0.

Построим корреляционное поле (рис.8):

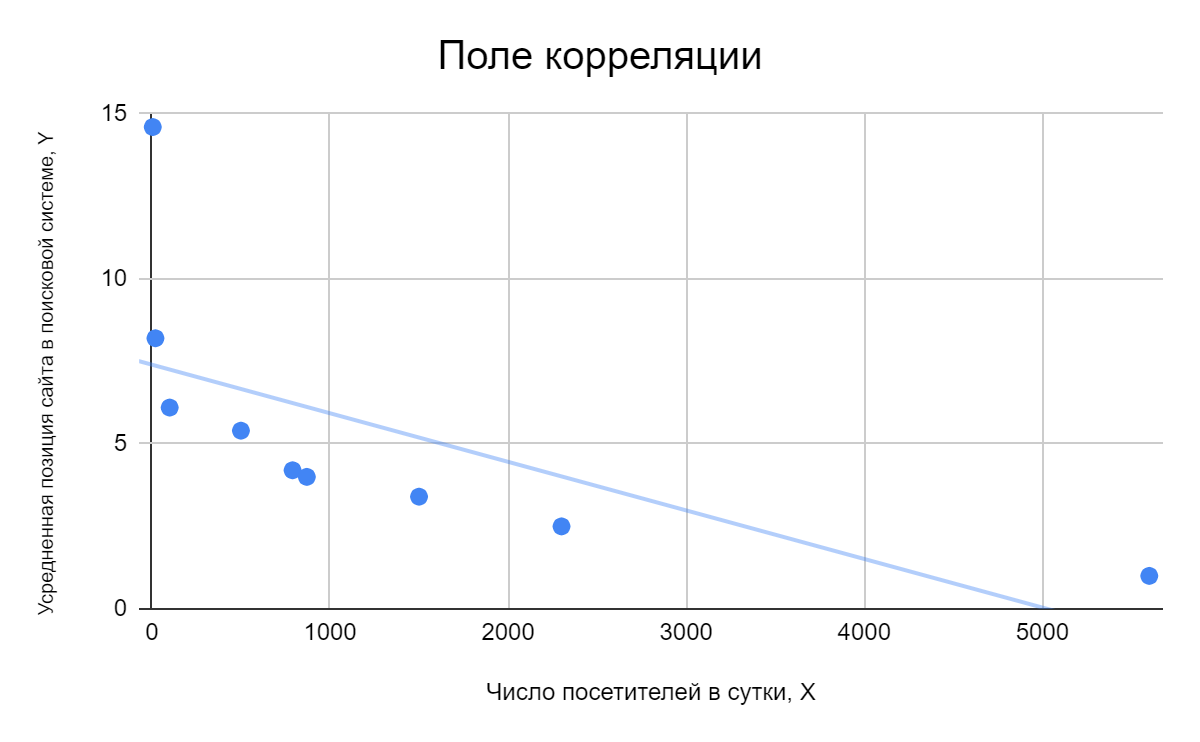


Рис.8. Корреляционное поле

**Ответ**: на уровне значимости 0,05 нулевую гипотезу принимаем.

## **Лабораторная работа №7. Часть 1**

*Выполнил Шардт М. А.*

**Цель работы:** Вычисление коэффициентов корреляции.

**Оборудования:** ПК, Microsoft Excel.

*Использованные формулы:*

1. Коэффициент корреляции — параметр, который характеризует степень линейной взаимосвязи между двумя выборками, рассчитывается по формуле:

1. Коэффициент ранговой корреляции Спирмена:

,

где - квадраты разности рангов,

n - число наблюдений.

1. Статистика гипотезы Стьюдента:

*Задание 2:*

Построить корреляционное поле. Сделать предположение о форме и направлении взаимосвязи двух исследуемых показателей. Вычислить коэффициент линейной корреляции Пирсона и коэффициент ранговой корреляции Спирмена.

*Пример 1:*

10 школьникам были даны тесты на наглядно-образное и вербальное мышление. Измерялось среднее время решения заданий теста в секундах. Исследователя интересует вопрос: существует ли взаимосвязь между временем решения этих задач? Переменная X — обозначает среднее время решения наглядно-образных, а переменная Y— среднее время решения вербальных заданий тестов.

| № испытуемых | X | Y |
| --- | --- | --- |
| 1 | 19 | 17 |
| 2 | 32 | 7 |
| 3 | 33 | 17 |
| 4 | 44 | 28 |
| 5 | 28 | 27 |
| 6 | 35 | 31 |
| 7 | 39 | 20 |
| 8 | 39 | 17 |
| 9 | 44 | 35 |
| 10 | 44 | 43 |

*Решение:*

Исходя из исходных данных сделаем вспомогательную таблицу:

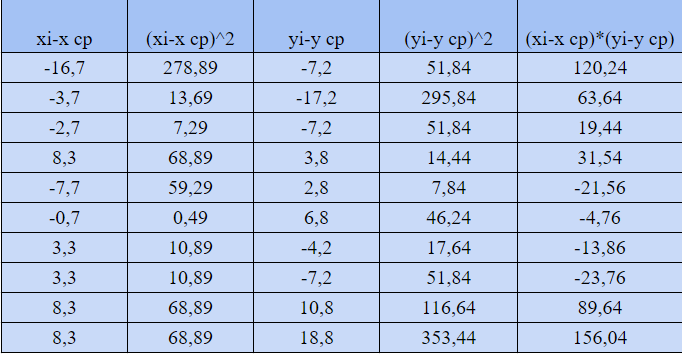


Табл.1. Вспомогательная таблица с промежуточными вычислениями

1. Нулевая гипотеза – взаимосвязь между временем решения существует.
2. Конкурирующая гипотеза – взаимосвязь между временем решения отсутствует.

Найдем значение коэффициента корреляции по формуле (1):



Далее по таблице “Критические значения корреляции rzy Пирсона” определим критические значения для полученного коэффициента корреляции, где k = n - 2 = 10 - 2 = 8, P = 0,05 (уровень значимости):



Для автоматического поиска критического значения по таблице используем функции ПОИСК и ИНДЕКС.

= 0,63 > 0,54 , следовательно, гипотеза Н1 отвергается и принимается гипотеза H0.

Построим корреляционное поле (рис.1):

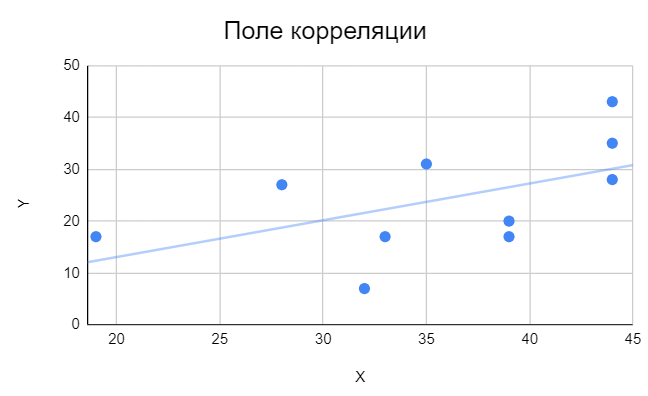


Рис.1. Корреляционное поле

**Ответ:** на уровне значимости 0,05 нулевую гипотезу Н0 принимаем.

*Пример 2:*

Преподавателю и студенту было предложено расположить 10 профессий в порядке их общественной значимости. Ответы перечислены в таблице 1:

***Таблица 1.***

| Оценка преподавателя, xi | Профессии | Оценка студента, yi |
| --- | --- | --- |
| 3 | профессор | 2 |
| 1 | врач | 1 |
| 4 | учитель школы | 7 |
| 2 | директор магазина | 4 |
| 8 | бухгалтер | 5 |
| 6 | банкир | 3 |
| 9 | водитель | 9 |
| 5 | журналист | 8 |
| 10 | ди-джей | 10 |

Какова корреляция рангов между двумя рядами оценок? Одинаково ли мнение преподавателя и студента по этому вопросу?

*Решение:*

Исходя из исходных данных составим вспомогательную таблицу разности рангов, их квадраты и суммы:

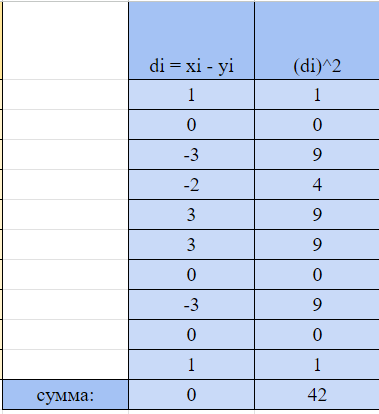


Табл. 2. Вспомогательная таблица с промежуточными вычислениями

Найдем коэффициент ранговой корреляции Спирмена по формуле (2):



Далее необходимо проверить, существует ли положительная корреляционная связь между мнениями преподавателя и студента. Для этого (при n ≥ 10) используем формулу (3) с k = (n – 2) степенями свободы:



1. Нулевая гипотеза – коэффициент корреляции не является статистически значимым ().
2. Конкурирующая гипотеза – существует положительная корреляционная зависимость ().

Далее по таблице “Критические точки t-распределения Стьюдента” определим критическое значение статистики Стьюдента, где k = n - 2 = 10 - 2 = 8, уровень значимости α = 0.05 для односторонней (правосторонней) критической области:



Для автоматического поиска критического значения по таблице используем функции ПОИСК и ИНДЕКС.

t расч > t кр(3,16>1,86). Следовательно, связь между мнениями преподавателя и студента является статистически значимой при 5%-ном уровне значимости.

Построим корреляционное поле (рис.2):

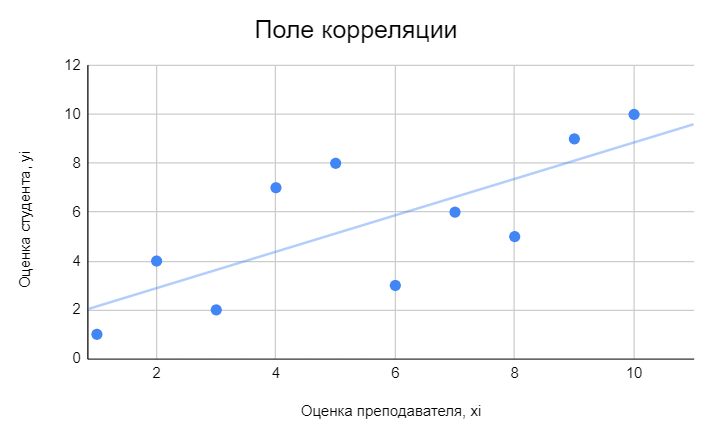


Рис.2. Корреляционное поле

**Ответ:** на уровне значимости 0,05 нулевую гипотезу отвергаем.

*Задание 3:*

Построить корреляционное поле. Сделать предположение о форме и направлении взаимосвязи двух исследуемых показателей. Найти значения коэффициентов ранговой корреляции Спирмена.

*Задача 3.1*

С помощью коэффициента ранговой корреляции установить зависимость между стажем практической работы и временем решения контрольной задачи у 10 программистов на основе следующих данных:

| Номера испытуемых | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Стаж (в мес.) | 32 | 15 | 16 | 18 | 20 | 28 | 21 | 29 | 23 | 17 |
| Время решения (в мин.) | 12 | 24 | 23 | 21 | 20 | 9 | 11 | 10 | 15 | 16 |

*Решение:*

Сформулируем статистические гипотезы:

1. Нулевая гипотеза – коэффициент корреляции не является статистически значимым ().
2. Конкурирующая гипотеза – существует положительная корреляционная зависимость ().

Определим ранги оценкам стажа xi, располагая оценки в возрастающем порядке и сохраняя связь с оценками времени решения. Ранги пронумеруем от 1 до 10.

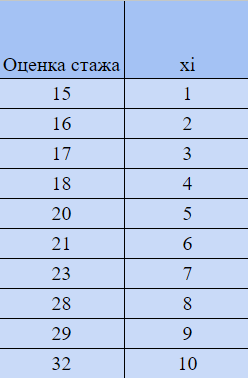


Табл. 3. Вспомогательная таблица с промежуточными вычислениями

Аналогично определим ранги оценкам времени решения yi. И также пронумеруем их от 1 до 10, располагая по возрастанию.

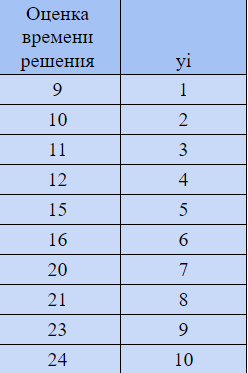


Табл. 4. Вспомогательная таблица с промежуточными вычислениями

Для пояснения процесса вычисления рангов yi возьмем ранг y1. Мы рассматриваем оценку времени решения программиста, который по оценке стажа занимает 1 место (xi = 1; 15 мес.) Теперь найдем оценку времени решения этого программиста (24 мин.) и видим, что это время решения получило ранг yi=10. В сводную таблицу заносим xi = 1, yi=10. Проделываем аналогичные действия для остальных значений и получаем сводную таблицу:

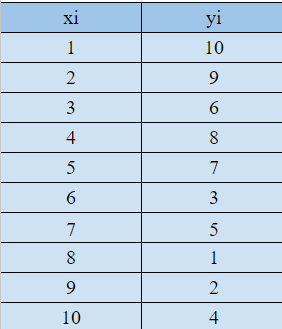


Табл. 5. Сводная таблица результатов

Исходя из полученных данных составим вспомогательную таблицу разности рангов, их квадраты и суммы:

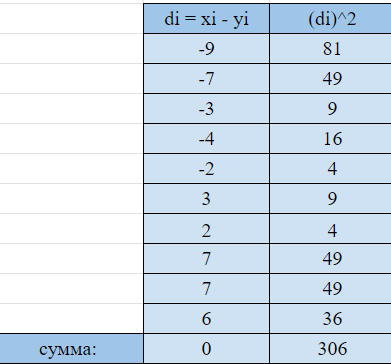


Табл. 6. Вспомогательная таблица с промежуточными вычислениями

Найдем коэффициент ранговой корреляции Спирмена по формуле (2):



Далее необходимо проверить, существует ли положительная корреляционная связь между оценкой стажа и оценкой времени решения задач. Для этого (при n ≥ 10) используем формулу (3) с k = (n – 2) степенями свободы:



Далее по таблице “Критические точки t-распределения Стьюдента” определим критическое значение статистики Стьюдента, где k = n - 2 = 10 - 2 = 8, уровень значимости α = 0.05 для односторонней (правосторонней) критической области:



Для автоматического поиска критического значения по таблице используем функции ПОИСК и ИНДЕКС.

|t расч |> t кр(4,65>1,86). Следовательно, связь между оценкой стажа и оценкой времени решения является статистически значимой при 5%-ном уровне значимости. Также можно сказать, что связь достаточно сильная и обратная по направлению (видно по корреляционному полю).

Построим корреляционное поле (рис.3):

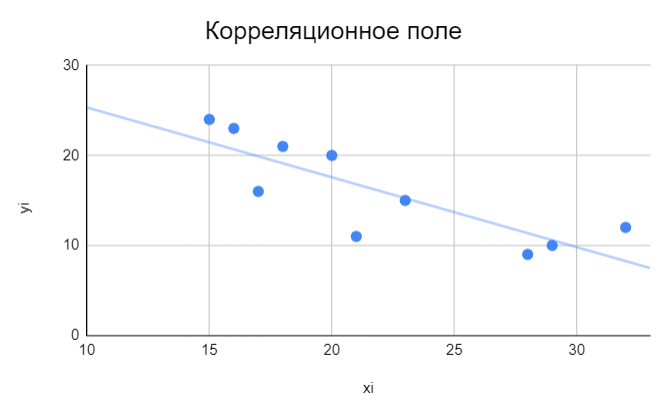


Рис.3. Корреляционное поле

**Ответ:** на уровне значимости 0,05 нулевую гипотезу отвергаем.

*Задача 3.2*

Три арбитра оценили мастерство 10 спортсменов, в итоге были получены три последовательности рангов (в первой строке приведены ранги арбитра А, во второй – ранги арбитра В, в третьей – ранги арбитра С):

| xi | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| yi | 3 | 10 | 7 | 2 | 8 | 5 | 6 | 9 | 1 | 4 |
| zi | 6 | 2 | 1 | 3 | 9 | 4 | 5 | 7 | 10 | 8 |

Определить пару арбитров, оценки которых наиболее согласуются, используя коэффициент ранговой корреляции Спирмена. Построить корреляционное поле.

*Решение:*

Для определения коэффициента корреляции Спирмена найдем разности рангов, их квадраты и сумму для оценок каждой из пар арбитров.

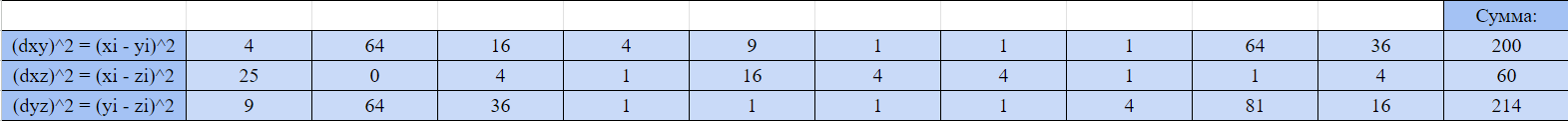


Табл. 7. Вспомогательная таблица с промежуточными вычислениями

Также для каждой пары арбитров вычислим коэффициент Спирмена и определим t-критерий Стьюдента tвыч по формулам (2) и (3). Далее по таблице “Критические точки t-распределения Стьюдента” определим критическое значение статистики Стьюдента, где k = n - 2 = 10 - 2 = 8, уровень значимости α = 0.05 для односторонней (правосторонней) критической области:

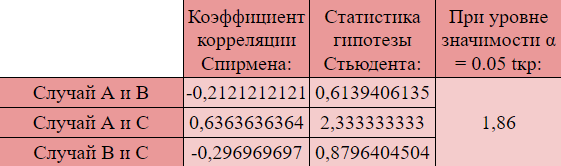


Табл. 8. Таблица с итоговыми результатами вычислений.

Для автоматического поиска критического значения по таблице используем функции ПОИСК и ИНДЕКС.

Построим поле корреляции для графического изображения зависимостей каждой пары арбитров (рис.4-рис.6):

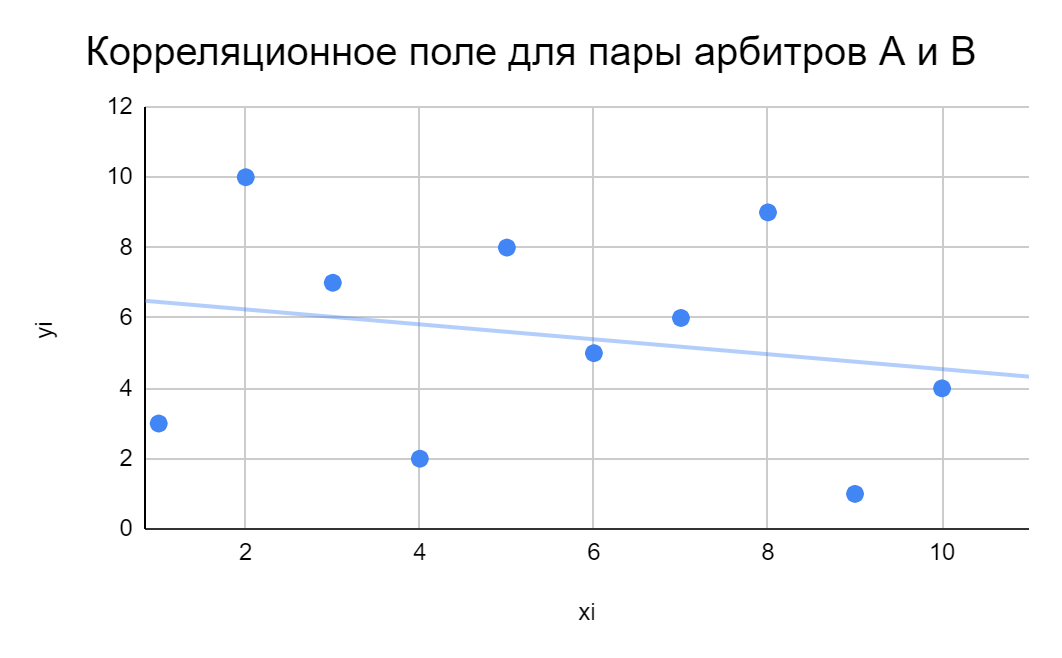


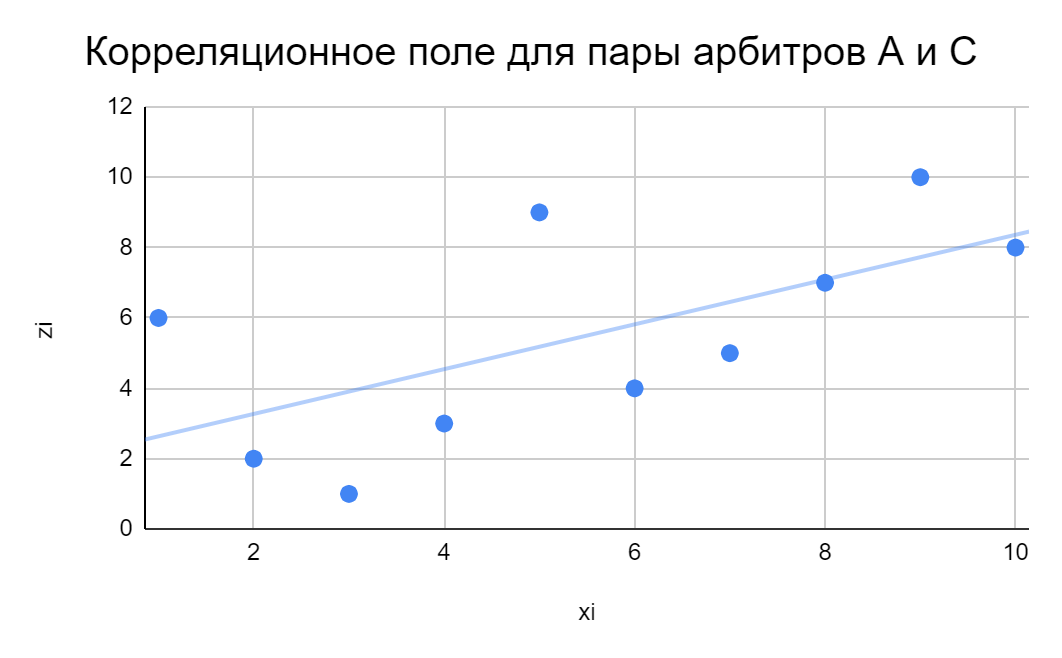
Рис.4. Корреляционное поле

Рис.5. Корреляционное поле

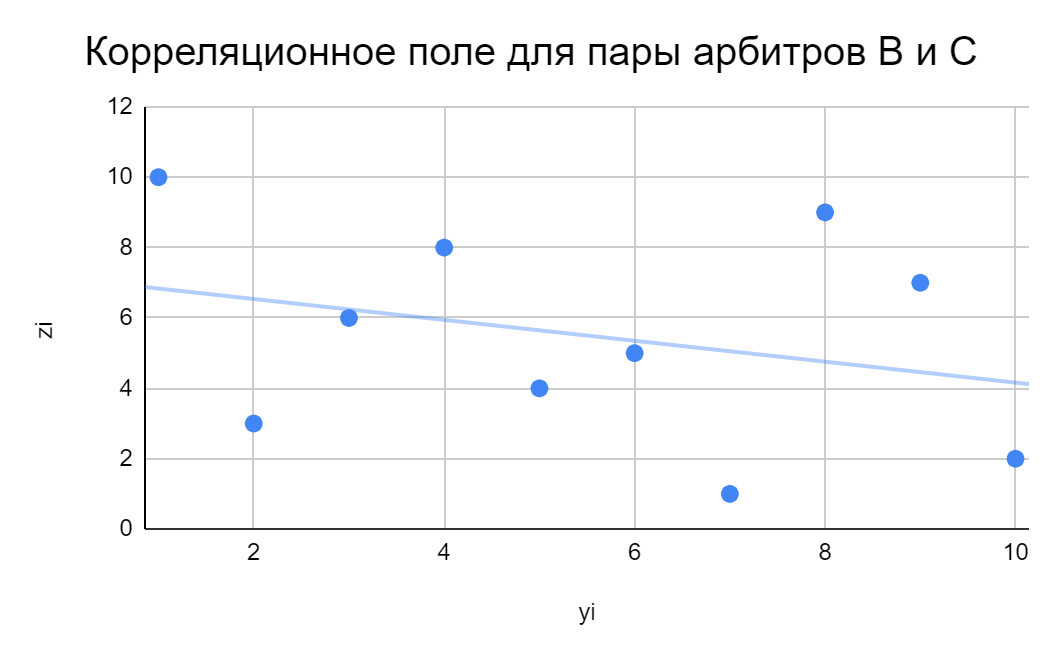


Рис.6. Корреляционное поле

Сравнивая значения с критическим, видим, что наиболее согласуются оценки арбитров А и С. Данный коэффициент ранговой корреляции наибольший по величине.

**Ответ:** Пара арбитров, оценки которых наиболее согласуются, это пара A и C.

*Задание 4:*

*Задача 4.1*

Построить корреляционное поле. Сделать предположение о форме и направлении взаимосвязи двух исследуемых показателей. Найти значение коэффициентов линейной корреляции. Необходимо определить взаимосвязь характеристик: агрессивности и IQ у школьников по полученным данным тестирования.

| № п/п | Данные по агрессивности X\_agr | Данные по IQ Y\_iq |
| --- | --- | --- |
| 1 | 24 | 100 |
| 2 | 27 | 115 |
| 3 | 26 | 117 |
| 4 | 21 | 119 |
| 5 | 20 | 134 |
| 6 | 31 | 94 |
| 7 | 26 | 105 |
| 8 | 22 | 103 |
| 9 | 20 | 111 |
| 10 | 18 | 124 |
| 11 | 30 | 122 |
| 12 | 29 | 109 |
| 13 | 24 | 110 |
| 14 | 26 | 86 |

*Решение:*

Сформулируем статистические гипотезы:

1. Нулевая гипотеза – взаимосвязь между агрессивностью и IQ у школьников существует.
2. Конкурирующая гипотеза – взаимосвязь между агрессивностью и IQ у школьников отсутствует.

Исходя из исходных данных сделаем вспомогательную таблицу:

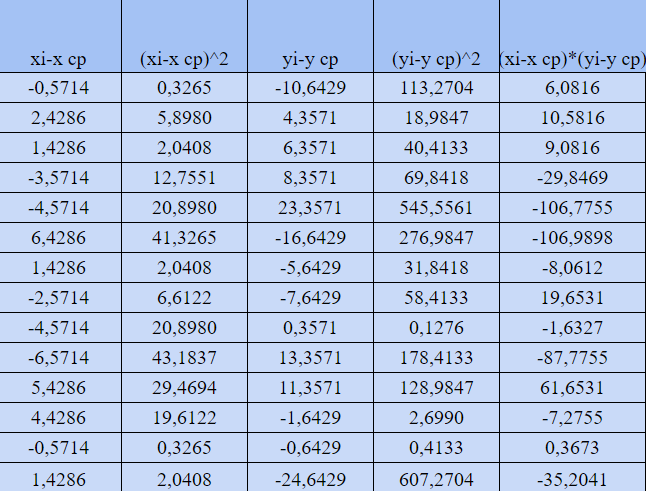


Табл.9. Вспомогательная таблица с промежуточными вычислениями

Найдем значение коэффициента корреляции по формуле (1):



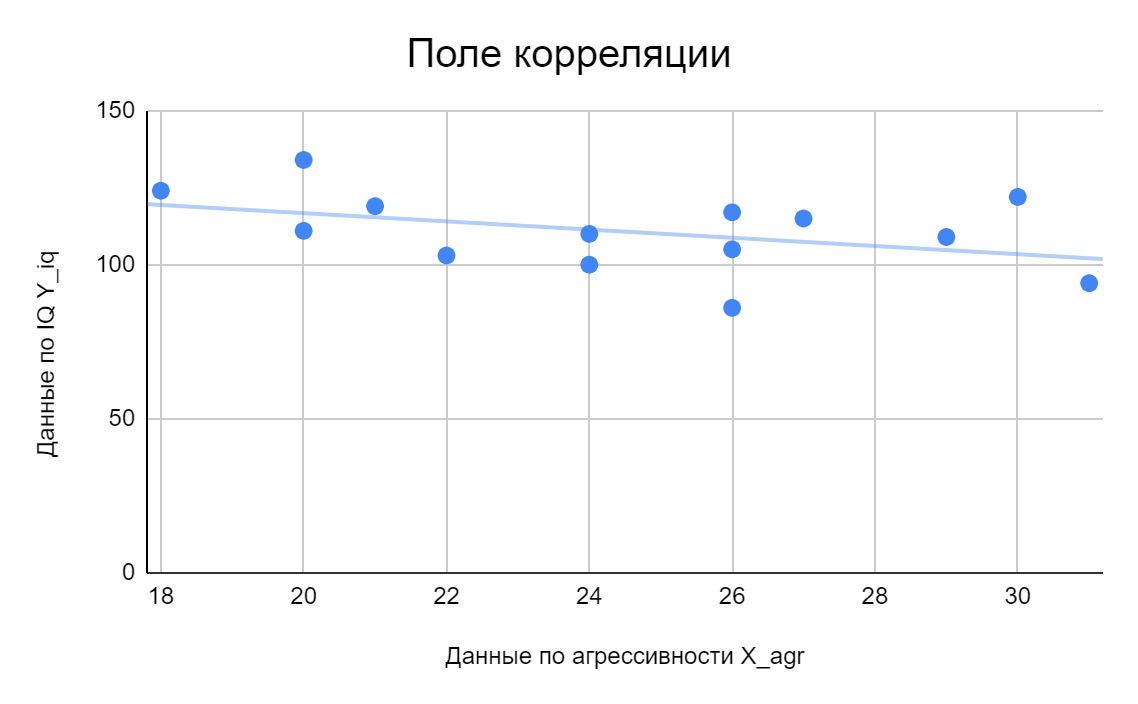
Далее по таблице “Критические значения корреляции rzy Пирсона” определим критические значения для полученного коэффициента корреляции, где k = n - 2 = 14 - 2 = 12, P = 0,05 (уровень значимости):



Для автоматического поиска критического значения по таблице используем функции ПОИСК и ИНДЕКС.

= 0,53 > |-0,42| , следовательно, гипотеза Н1 отвергается и принимается гипотеза H0.

Построим корреляционное поле (рис.7):

Рис.7. Корреляционное поле

**Ответ**: на уровне значимости 0,05 нулевую гипотезу принимаем.

*Задача 4.2*

На основании наблюдений за развивающимся сайтом и изменением его средневзвешенной позиции по основным запросам в поисковой системе необходимо проверить, можно ли говорить о линейной зависимости между позицией сайта и числом посетителей. Построить корреляционное поле.

Исходные данные:

X - число посетителей в сутки;

Y – усредненная позиция сайта в поисковой системе.

В таблице даны значения признаков

| № п/п | Число посетителей в сутки, X | Усредненная позиция сайта в поисковой системе, Y |
| --- | --- | --- |
| 1 | 500 | 5,4 |
| 2 | 790 | 4,2 |
| 3 | 870 | 4 |
| 4 | 1500 | 3,4 |
| 5 | 2300 | 2,5 |
| 6 | 5600 | 1 |
| 7 | 100 | 6,1 |
| 8 | 20 | 8,2 |
| 9 | 5 | 14,6 |

*Решение:*

Сформулируем статистические гипотезы:

1. Нулевая гипотеза – взаимосвязь между позицией сайта и числом посетителей существует.
2. Конкурирующая гипотеза – взаимосвязь между позицией сайта и числом посетителей отсутствует.

Исходя из исходных данных сделаем вспомогательную таблицу:

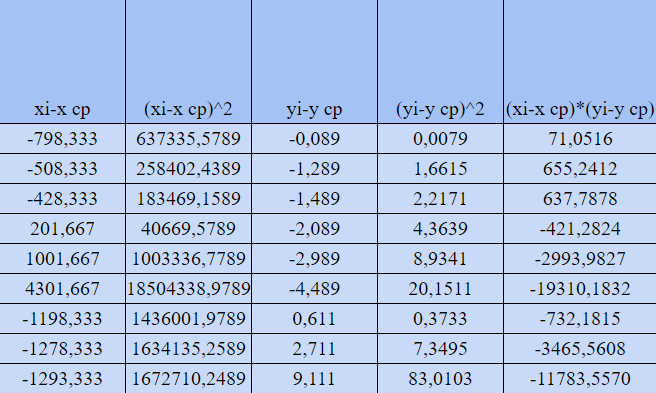


Табл.10. Вспомогательная таблица с промежуточными вычислениями

Найдем значение коэффициента корреляции по формуле (1):



Далее по таблице “Критические значения корреляции rzy Пирсона” определим критические значения для полученного коэффициента корреляции, где k = n - 2 = 9 - 2 = 7, P = 0,05 (уровень значимости):



Для автоматического поиска критического значения по таблице используем функции ПОИСК и ИНДЕКС.

= 0,67 > |-0,65| , следовательно, гипотеза Н1 отвергается и принимается гипотеза H0.

Построим корреляционное поле (рис.8):

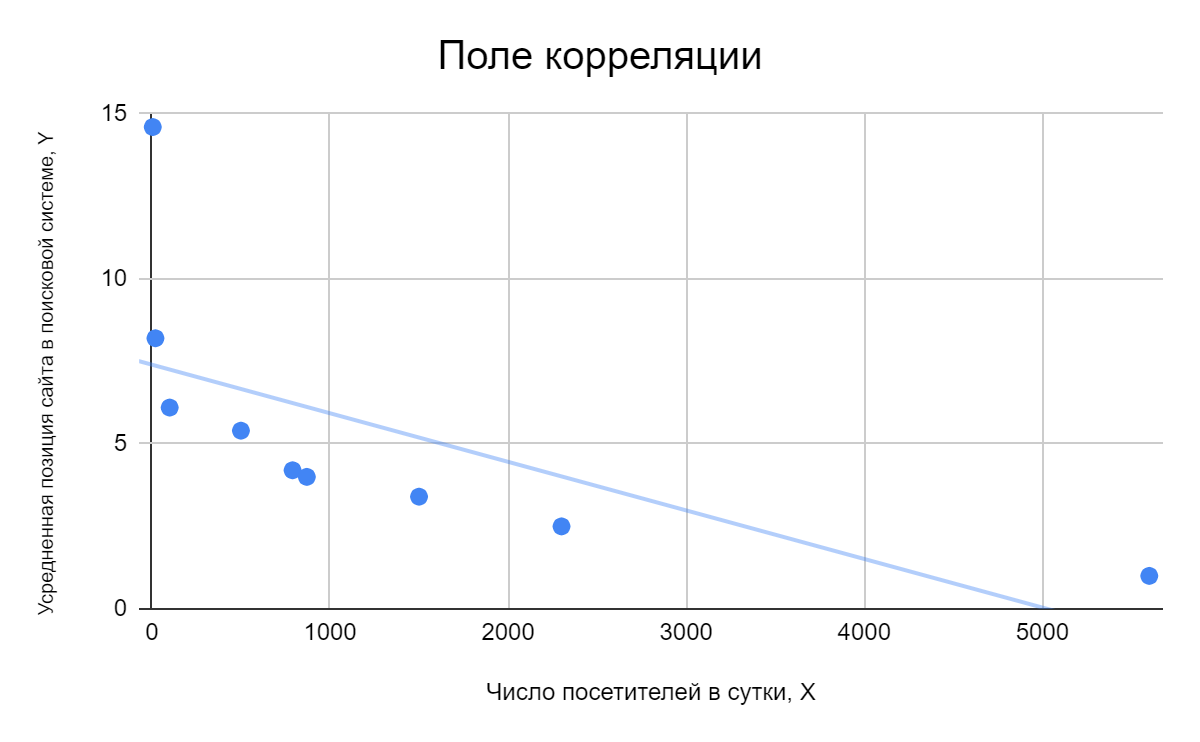


Рис.8. Корреляционное поле

**Ответ**: на уровне значимости 0,05 нулевую гипотезу принимаем.