# Лабораторные работы по дисциплине Вероятностные модели

Исполнитель: Сидельников Павел

## Вычислительная работа №1

**Вариант: 9**

### Постановка

В таблице представлены данные о производстве всех зерновых культур и производстве пшеницы в России за 1996 – 2002 гг.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Год | 1996 | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 |
| Производство зерновых, млн т | 69,3 | 88,6 | 47,9 | 54,7 | 65,5 | 85,2 | 86,6 |
| Производство пшеницы, млн т | 34,9 | 44,3 | 27,0 | 31,0 | 34,5 | 47,0 | 57,7 |

Вычислите выборочный коэффициент ранговой корреляции Спирмена между показателями. Проверьте нулевую гипотезу об отсутствии орреляционной связи на уровне значимости а = 0,05.

### Решение

Присвоим ранги ri оценкам по производству зерна:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| ri | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|  | 88,6 | 86,6 | 85,2 | 69,3 | 65,5 | 54,7 | 47,9 |

Присвоим ранги si оценкам по производству пшеницы:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| si | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|  | 57,7 | 47,0 | 44,3 | 34,9 | 34,5 | 31,0 | 27,0 |

Найдем ранг s1. Индекс i = 1 указывает, что рассматривается год, за который было произведено больше всего зерна (значение равно 88,6), из условия видно, что за тот же год было произведено 44,3 пшеницы, значение которого указана на 3 месте. Таким образом, ранг s1 = 3.

Найдем ранги:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| ri | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Si | 3 | 1 | 2 | 4 | 5 | 6 | 7 |

Найдем разность рангов:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Ri-si | -2 | 1 | 1 | 0 | 0 | 0 | 0 |

Вычислим сумму квадратов разностей рангов:

S = ∑ = 4 + 1 + 1+ 0 + 0 + 0+ 0= 6

Найдем искомый коэффициент ранговой корреляции Спирмена, учитывая, что n = 7:

Rs = 1 – 6 \* S / (n^3 -n) = 1 – 6 \* 6 / (7^3 - 7) = 1 – 36 / 336 = 1 – 3 / 28 = 0.89

Когда мы уже вычислили выборочный коэффициент ранговой корреляции Спирмена (rs = 1) для данных о производстве зерновых и производстве пшеницы в России за указанный период, мы можем проверить нулевую гипотезу об отсутствии корреляционной связи на уровне значимости α = 0,05:

Вычислите статистику используя точный тест Фишера для коэффициента корреляции:

F = (n - 2) \* rs² / (1 - rs²), где

n - количество наблюдений,

rs - выборочный коэффициент корреляции.

Соответственно:

n = 7, rs = 0,89

F = (7 - 2) \* (0,89)² / (1 – (0,89)²) = 25,78

Сравним значение статистики теста F с критическим значением F, соответствующим критической области.

Вычислим степени свободы числителя и степени свободы знаменателя:

df1 = n1 – 2 = 7 – 2 = 5

df2 = n2 – 2 = 7 – 2 = 5

Для α = 0.05 и степеней свободы df1 = 5 и df2 = 5, критическое значение F составляет примерно 5.05.

### Вывод

Искомый коэффициент равен 0.89, это указывает на наличие положительной корреляции.

Так как значение статистики теста F (25.78) превышает критическое значение F (5.05), мы отвергаем нулевую гипотезу об отсутствии корреляционной связи на уровне значимости α = 0.05. Это свидетельствует о наличии статистически значимой корреляционной связи между производством зерновых и производством пшеницы в России за указанный период.

## Вычислительная работа №2

### Постановка

**Вариант 9**

1. Для данных выборок найти оценки параметров распределения;
2. Взять в качестве выброса одну из вариант выборки. В чѐтных вариантах – наибольшую; в нечѐтных – наименьшую;
3. По критерию Граббса проверить гипотезу, что в выборке нет выбросов, с уровнем значимости а = 0,1;
4. По новой выборке снова найти оценки параметров, сравнить со старыми значениями. Сделать вывод.

|  |  |  |
| --- | --- | --- |
| 8.86 | 2.11 | 4.34 |
| 5.59 | 6.36 | 0.67 |
| 8.40 | 4.43 | 1.97 |
| 4.10 | 7.11 | 8.64 |
| 4.36 | 7.99 | 6.62 |

### Решение

Оценки параметров распределения:

Для каждой выборки можно найти оценки параметров распределения, такие как среднее значение и стандартное отклонение. Для этого выполним следующие шаги:

a) Найдем среднее значение (M) для каждой выборки:

Среднее значение первой выборки: M1 = (8.86 + 5.59 + 8.40 + 4.10 + 4.36) / 5 = 6.262

Среднее значение второй выборки: M2 = (2.11 + 6.36 + 4.43 + 7.11 + 7.99) / 5 = 5.6

Среднее значение третьей выборки: M3 = (4.34 + 0.67 + 1.97 + 8.64 + 6.62) / 5 = 4.48

b) Найдем стандартное отклонение (σ) для каждой выборки:

Стандартное отклонение первой выборки:

Стандартное отклонение второй выборки:

Стандартное отклонение третьей выборки:

Таким образом, мы найдем оценки параметров для каждой выборки.

Выброс:

В четных вариантах мы возьмем наибольшее значение в выборке в качестве выброса, а в нечетных вариантах - наименьшее значение.

В данном случае у нас нечетное количество выборок, поэтому мы возьмем наименьшее значение в качестве выброса.

Наименьшее значение: 2

Проверка гипотезы о наличии выброса с использованием критерия Граббса:

Для выполнения этого шага нам нужно иметь данные о выборке, которая содержит выброс, и вычислить значение критерия Граббса для проверки гипотезы о наличии выброса.

Новая выборка и сравнение оценок параметров:

После удаления выброса из исходной выборки, мы можем создать новую выборку и вычислить новые оценки параметров (среднее значение и стандартное отклонение). Затем мы можем сравнить их со старыми значениями, чтобы сделать выводы о влиянии выброса на оценки параметров.

Пожалуйста, предоставьте данные для новой выборки, и я помогу вам найти новые оценки параметров и сравнить их со старыми значениями.

### Вывод