

4.6 Тема. Коэффициент корреляции рангов

Цель. Знакомство с методами вычисления коэффициента ранговой корреляции

Наряду с параметрическими показателями корреляционной связи существуют и непараметрические, или порядковые, показатели, которые позволяют измерять степень сопряженности между признаками независимо от закона распределения и формы связи. Одним из таких показателей является коэффициент корреляции рангов, предложенный Спирменом:

$$r_s = 1 - \frac{6 \sum D^2}{n(n^2 - 1)} \quad (51)$$

где \sum - знак суммирования, D – разность между рангами сопряженных значений признаков x и y , n – число парных наблюдения, или объем выборки.

Ранжируя попарно связанные значения признаков, обращаем внимание на то, как они распределяются относительно друг друга. Если возрастающим значениям одного признака (x) соответствуют значения другого (y), то между ними существует положительная связь, если же при возрастании значений одного признака значения другого последовательно уменьшаются, это укажет на отрицательную связь между ними. При отсутствии корреляции ранжированным значениям одного признака будут соответствовать различные значения другого.

Обозначив ранжированные значения признаков порядковыми номерами (натуральным рядом чисел), нетрудно определить ранги этих значений и по их разности судить о степени зависимости одного признака от изменений другого. При полной связи ранги коррелируемых признаков совпадут и разность между ними будет равна нулю. В таких случаях коэффициент корреляции окажется равным единице. Если же признаки варьируют независимо друг от друга, то коэффициент корреляции рангов будет равен нулю.

Пример. При изучении зависимости между живой массой (кг) и содержанием гемоглобина (% по Сали) у павианов гамадрилов.

№	Масса (x)	Hb% (y)	Ранги рядов		$x_i - y_i = d$	d^2	Расчет ранга	
			x_i	y_i			x_i	y_i
1	17	70	1	1	0,0	0,00	70	1
2	18	74	2,5	3	0,5	0,25	72	2
3	18	78	2,5	7	4,5	20,25	74	3
4	19	72	4,5	2	2,5	6,25	76	4
5	19	77	4,5	5,5	1,0	1,00	77	5,5

6	20	76	6	4	2,0	4,00	77	5,5
7	21	88	7	10	3,0	9,00	78	7
8	22	80	8	8	0,0	0,00	80	8
9	23	77	9	5,5	3,5	12,25	86	9
10	25	86	10	9	1,0	1,00	88	10
Σ						54,00		

Если бы отдельные члены ранжированного ряда не повторялись, их рангами были бы соответствующие порядковые числа натурального ряда. Но так как некоторые варианты повторяются, их рангами будут средние арифметические из соответствующих чисел натурального ряда. Так, например, в ряду X варианты 18 и 19 повторяются дважды и их ранги равны полусумме соответствующих порядковых чисел. Подставляя известные величины в формулу Спирмена, находим

$$r_s = 1 - \frac{6 \cdot 54}{10(10^2 - 1)} = 1 - 0,327 = +0,673.$$

Задание 1. При изучении связи качества меха норок с агрессивностью по поведению отобраны 9 самцов, которые занимают ранги от самого спокойного, т.е. слабой агрессивности (1), до сильного (9). По качеству меха животные расположены также по рангу от худшего (1) до лучшего (9).

Норки А Б В Г Д Е Ж З И

Ранги:

по агрессивности... 1 2 3 4 5 6 7 8 9

по качеству меха... 3 1 2 5 4 8 9 6 7

Можно ли вести селекцию по поведению норок с целью повышения экономической эффективности?

Задание 2. Приводятся ранжированные ряды густоты шерсти овец от меньшей (1) до большей (7) и упитанности от низшей (1) до высшей (7). Что можно сказать о характере взаимосвязи между густотой шерсти и упитанностью овец?

Номера овец..... 20 21 22 23 24 25 26

Ранги:

по густоте шерсти..... 1 2 3 4 5 6 7

по упитанности..... 2 3 4 1 7 4 5

Контрольные вопросы.

1. В каких случаях используется коэффициент ранговой корреляции?

2. В чем заключается ценность непараметрических показателей связи?

3. Приведите формулы вычисления коэффициентов ранговой корреляции.