

РАНГОВЫЕ МЕТОДЫ ИССЛЕДОВАНИЯ СТАТИСТИЧЕСКОЙ СВЯЗИ. КОРРЕЛЯЦИОННЫЕ ПЛЕЯДЫ

RANK METHODS OF STATISTICAL COMMUNICATION RESEARCH. CORRELATIONAL PLEIADES

*Андреев Александр Алексеевич
Тюрнева Татьяна Геннадьевна*

Ключевые слова: корреляционные плеяды; ранговые методы

Рассмотрена задача оптимизации размерности пространства параметров, характеризующих вторичный рынок мототехники и проведён сравнительный анализ полученных результатов. Объектом исследования послужили мотоциклы марки «Honda» и «Yamaha». Данные были получены с сайта <https://moto.drom.ru/> по состоянию на 06.03.2022 года. Рассматриваемыми параметрами стали количественные признаки – X_1 -число просмотров объявления, X_2 - цена мотоцикла, X_3 - год выпуска, X_4 -объем двигателя, X_6 – пробег и качественные признаки – X_5 - тактность двигателя (2-х и 4-х тактные), X_7 - оттенок цвета (тёмный, цветной, светлый) и X_8 - субъективная оценка цены представленной модели (низкая, средняя, высокая). Объем выборок составил $n_1=17$ и $n_2=16$ соответственно для марок «Honda» и «Yamaha».

Для оценки наличия и тесноты зависимости между параметрами использованы методы корреляционного анализа. Ранговая корреляция подразумевает ранжирование данных, то есть присвоение значениям признака определённого ранга. В исследовании был применён метод ранговой корреляции Спирмена. Все коэффициенты вычислены по формуле:

$$\rho = 1 - \frac{6 \sum d^2}{n(n^2-1)},$$

где, d -разность рангов каждой пары значений, n – объём выборки, и записаны в корреляционную матрицу (рис.1). Корреляционная матрица может быть преобразована в корреляционную плеяду, которая позволяет наглядно отображать информацию о корреляциях в виде их структурирования и объединения коррелирующих факторов. Существующие варианты метода корреляционных плеяд являются в действительности несколько упрощёнными эвристическими

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
X_1	1							
X_2	0,54	1						
X_3	0,23	0,18	1					
X_4	0,68	0,61	0,19	1				
X_5	0,5	0,4	0,046	0,8	1			
X_6	0,23	0,002	-0,12	0,61	0,79	1		
X_7	-0,31	-0,36	-0,02	-0,14	0,1	0,14	1	
X_8	-0,18	0,13	-0,14	-0,23	0,09	-0,07	-0,02	1

Рис.1. Кор. матрица для моделей марки «Honda»

версиями более совершенных в математическом плане алгоритмов исследования структуры связей между компонентами многомерного параметра, использующими графы-деревья.

После нахождения всех коэффициентов, оставим только значимые для интерпретации значения и представим их методом корреляционных плеяд при пороге $t=0,3$ (рис. 2):

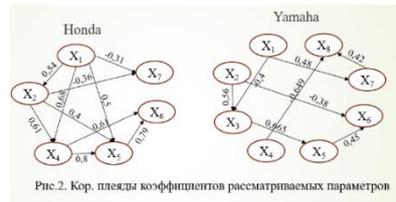


Рис. 2. Кор. плеяды коэффициентов рассматриваемых параметров

Выводы: внутри группы марки «Honda» люди больше склонны просматривать объявление в зависимости от цены, объема двигателя или цвета мотоцикла. Цена устанавливается в зависимости от объема двигателя, его тактности и цвета. Объем двигателя достаточно высоко связан с тактностью и чуть менее с пробегом, вследствие этого и тактность достаточно тесно связана с пробегом. Внутри группы марки «Yamaha» люди больше смотрят на год выпуска и оттенок цвета, где год выпуска имеет обратную связь по отношению к просмотрам. Ценообразование зависит от пробега модели и его года выпуска, чем больше пробег, тем меньше цена. Год производства умеренно связан с тактностью модели, а тактность с пробегом. Оценка цены зависит от объема двигателя и оттенка цвета.